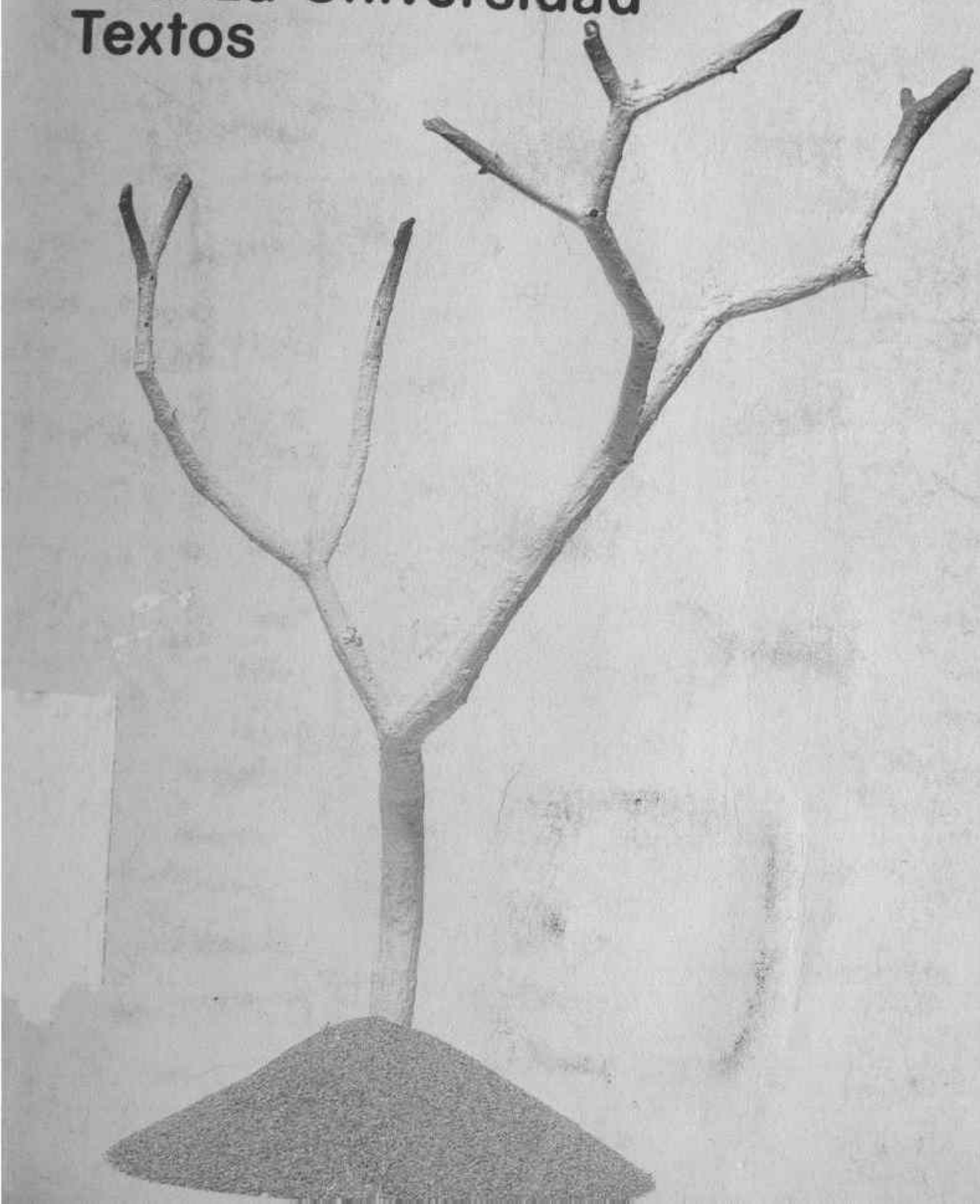
C. Ulises Moulines

Exploraciones metacientíficas

Alianza Universidad Textos



Alianza Universidad. Textos

C. Ulises Moulines

Exploraciones metacientíficas

Estructura, desarrollo y contenido de la ciencia

451 T
4AM-I
VERDE
M

Prólogo de Jesús Mosterín
Prefacio del autor

Alianza Editorial

125679

B67
M6.8
c.8

© C. Ulises Moulines
© Alianza Editorial, S. A., Madrid, 1982
Calle Milán; 38. ☎ 200 00 45
I.S.B.N.: 84-206-8038-9
Depósito legal: M. 985-1982
Compuesto en Fernández Ciudad, S. L.
Impreso en Hijos de E. Minuesa, S. L.
Ronda de Toledo, 24 - Madrid-5
Printed in Spain

A Adriana

INDICE

Prólogo de Jesús Mosterín 11

Prefacio del autor 19

Recordatorio de símbolos lógicos y matemáticos 23

1. Consideraciones metafilosóficas 27

1.1. Blanco, negro, gris: contra el extremismo filosófico. . . . 31

1.2. ¿Qué hacer en filosofía de la ciencia? Una alternativa en catorce puntos 40

2. Análisis estructural de la ciencia 61

2.1. Hacia un nuevo concepto de teoría empírica 63

2.2. Un programa de reconstrucción estructural de las teorías físicas 74

2.3. Forma y función de los principio-guía en las teorías físicas 88

2.4. Redes teóricas 108

2.5. La red teórica de la mecánica clásica de partículas . . . 117

2.6. La red teórica de la termodinámica de los sistemas simples 132

2.7. Aplicación aproximativa de teorías empíricas 164

2.8. Relaciones interteóricas 191

2.9. Aproximación interteórica: el caso Kepler-Newton. . . . 204

2.10. Marcos teóricos: el ejemplo de la termodinámica 225

3. Génesis y desarrollo de programas científicos 249

3.1. Forma y contenido de las revoluciones científicas: el caso de la mecánica newtoniana 251
3.2. Los fundamentos metodológicos de la filosofía natural de Isaac Newton 262
3.3. La evolución de la mecánica newtoniana 278
3.4. La epistemología fisiológica de Hermann von Helmholzt 297
3.5. La génesis del positivismo en su contexto científico ... 305

4. Hacia una ontología de la ciencia 325

4.1. Ontosemántica de las teorías científicas 329
4.2. Los conceptos de materia y sus dificultades 346
4.3. Los problemas del materialismo 357

Bibliografía 367

PROLOGO DE JESUS MOSTERIN

El hacer es más importante biológicamente, más fundamental y más antiguo que el reflexionar sobre lo que hacemos. La gramática es parasitaria del habla, y no a la inversa. Mucho antes de que a nadie se le ocurriese reflexionar sobre el habla y tratar de explicitar la gramática de la lengua, la gente ya llevaba cientos de miles de años hablando. Los mejores conversadores, oradores y escritores no han solido tener ni idea de gramática. Y maldita la falta que les hubiera hecho.

La ciencia es una actividad mucho más reciente que el habla. Y la reflexión sobre la ciencia, la filosofía de la ciencia, es también una actividad parasitaria de la ciencia misma y muy posterior a ella. Desde luego la mayor parte de los científicos no saben nada de filosofía de la ciencia, y la mayor parte de los filósofos de la ciencia no aportan nada al progreso de la misma. Esto no tiene nada de lamentable. Los habitantes de nuestro planeta somos muchos y hay sitio para todos, para los que hacen, para los que reflexionan sobre lo que los otros hacen e incluso para los que tratan de hacer y reflexionar a la vez.

La ciencia es una complejísima actividad humana, un fenómeno cultural que depende de una enorme cantidad de factores biológicos, psicológicos, sociales, económicos, técnicos, legales, políticos, ideológicos, etc. Millones de humanes se encuentran involucrados en el proceso científico, como profesores o como estudiantes, como investigadores o como administradores, como descubridores de nuevas ideas o como autores de libros de texto, buscando fondos, criticando artículos, poniendo a prueba hipótesis, concediendo becas, progra-

mando computadoras, observando gorilas en la selva... Miles de millones de dólares se gastan (y en parte malgastan) cada año en el mantenimiento de esta actividad, en edificios, pizarras, ciclotrones, sueldos... Una cascada ininterrumpida de informes, actas, tesis, revistas y libros científicos sale constantemente de las imprentas y posibilita, complica y en cualquier caso manifiesta la amplitud alcanzada por la empresa científica en nuestro tiempo. Y en las mentes de los participantes surgen las preguntas y las respuestas, la curiosidad y la envidia, el gozo del saber y el descubrir –que, según algunos filósofos, es el más grande gozo que le es dado a la especie humana– y la frustración del fracaso y del choque contra los propios límites. Todo esto y mucho más es la ciencia. Por eso la reflexión sobre la ciencia no puede ser tarea de una sola disciplina, sino de muchas. La historia de la ciencia, la sociología de la ciencia, la economía de la ciencia, la psicología de la ciencia, la filosofía de la ciencia, etc.

La actividad científica produce muchos resultados: viajes, salarios, negocios, cambios de actitudes, esfuerzos, sorpresas, inquietudes, prestigios, ilusiones, úlceras de estómago, etc. Entre otros, la ciencia produce un tipo de saber distinto al saber común encardinado en el lenguaje ordinario y en la experiencia cotidiana, un saber más sistematizado, regimentado, explicitado, un saber de mucho mayor alcance y precisión. Para ello se introducen nuevos conceptos, se aplican nuevos métodos, se formulan definiciones, hipótesis, leyes, y, en las ciencias más avanzadas, se construyen teorías. El estudio de este arsenal de resultados conceptuales o intelectuales de la actividad científica constituye el objeto propio de la filosofía de la ciencia (en griego, *epistemología*). Así como el hablar nos permite habérnoslas con un mundo mucho más amplio que el del mero percibir, así también el teorizar nos abre nuevas posibilidades que no tenía el mero hablar. El teorizar, el construir, cambiar y manejar teorías, constituye la culminación intelectual de la empresa científica. Y el estudio de las teorías está por ello mismo en el foco central de atención de la filosofía de la ciencia.

La epistemología tiene una larga prehistoria, algunos de cuyos capítulos describe Moulines en este libro (véase *Los fundamentos metodológicos de la filosofía natural de Isaac Newton*, *La epistemología fisiología de Hermann von Helmholtz* y III. 5), pero su historia es relativamente breve. En efecto, la filosofía de la ciencia como disciplina establecida y profesionalizada apenas si tiene medio siglo de existencia. Limitándonos a la epistemología de los últimos treinta años, podríamos dividir su desarrollo en tres etapas, que *grosso modo* coinciden con las tres décadas que abarca: la etapa de los años cincuenta, la de los años sesenta y la de los años setenta. Esta división es trivial y arbitraria, pero no peor que sus posibles alternativas.

La epistemología de los años cincuenta representa la culminación de la filosofía clásica de la ciencia, la filosofía iniciada en Austria y

Alemania a finales de los años veinte y continuada luego en Inglaterra y Estados Unidos por Reichenbach, Carnap, Popper y Hempel, entre otros. Esta filosofía clásica de los años cincuenta era enormemente rica y coherente, pero albergaba también sonadas polémicas y discusiones. Casi todos estaban de acuerdo en la distinción entre conceptos observacionales y conceptos teóricos, pero había grandes diferencias de opinión sobre el rol de estos últimos en la ciencia, que era interpretado de modo opuesto por realistas, operacionalistas y nominalistas. Todos estaban de acuerdo en la concepción de las teorías como conjuntos de enunciados organizados axiomáticamente, y casi todos aceptaban como ejemplar la dilucidación por Hempel de las nociones de explicación y predicción científica. Sin embargo, la contrastación empírica de las teorías con los hechos provocaba profundas diferencias de opinión entre el inductivismo de Carnap y el refutacionismo de Popper. Al final de los años cincuenta se resume esta filosofía clásica de la ciencia en una serie de libros que van apareciendo a principios de los sesenta, como *The Structure of Science* (1961)*, de Ernest Nagel. Esta epistemología clásica constituye uno de los momentos más creativos de la historia de la filosofía. Por primera vez filósofos conocedores de la ciencia de su tiempo y armados con el utillaje formal del cálculo lógico elaboran de un modo profesional y coherente una filosofía precisa y sistemática de la ciencia. Pero a finales de los años cincuenta esta concepción chocaba con sus propias limitaciones. Su deseo de aplicar un formalismo lógico excesivamente rígido (el cálculo deductivo) llevaba a muchos epistemólogos a concentrarse en el análisis de ejemplos sencillos y fácilmente formalizables («todos los cuervos son negros»), pero alejados de la praxis científica, y a aplicar esquemas de refutación y confirmación de teorías apriorísticos y poco realistas.

La epistemología de los años sesenta surge precisamente como una reacción historicista frente a ese estado de cosas. La filosofía clásica de la ciencia se verá ahora acusada de ser excesivamente simplista, de ser más una filosofía de la ciencia ficción que de la ciencia real. En la ciencia real las teorías no se construyen por inducción ni se destruyen por la regla lógica del *modus tollens* a partir de un solo contraejemplo. La historia y la observación del presente nos muestran que las teorías se mantienen, a pesar de que no se cumplan en diversos casos, a pesar de que no resulten aplicables a ciertos dominios. Los filósofos de la ciencia que llevaban la voz cantante en los años sesenta sabían muy poco de lógica, pero mucho de historia. Muchos de ellos eran de hecho historiadores de la ciencia, que se habían metido a filósofos para criticar la imagen deformada de la ciencia que (en su opinión) estaban dando los filósofos clásicos. La acción crítica de los filósofos historicistas de la ciencia de los años

* *La estructura de la ciencia*, trad. cast., Buenos Aires, Paidos, 1968.

sesenta –Hanson, Toulmin, Kuhn, etc.– fue sumamente oportuna y contribuyó poderosamente a despertar a la epistemología clásica del «sueño dogmático» o al menos de la complacencia en que ésta se había instalado. El problema con los filósofos historicistas de los años sesenta es que su postura era básicamente negativa, crítica. Señalaban los defectos de la concepción clásica, pero no sabían ofrecer soluciones. Eran capaces de narrar (por algo eran historiadores) brillantemente casos concretos, pero carecían del vigor intelectual y de la capacidad para el pensamiento abstracto y preciso que había caracterizado a los filósofos clásicos. En definitiva, el historicismo resultó ser una moda efímera, aunque su impacto en la filosofía de la ciencia será sin duda duradero y saludable. Quizás el más influyente de los filósofos-historiadores fue Thomas S. Kuhn. Su libro *The Structure of Scientific Revolutions* (1962)** marcaría como ningún otro todas las discusiones epistemológicas de los años sesenta. Típicamente se trata de un libro brillante, jugoso, sugestivo, lleno de ejemplos históricos y de observaciones críticas acertadas sobre la epistemología clásica. Pero la construcción filosófica positiva que contiene es muy endeble. Su concepto central, el de paradigma, es una noción confusa y jabonosa, una metáfora más que un concepto preciso. El mismo Kuhn acabaría reconociendo su insuficiencia en el epílogo a la segunda edición de la obra citada (1970). En su último libro, *Black-Body Theory and the Quantum Discontinuity*, 1894-1912 (1978)***, Kuhn vuelve por sus fueros de historiador puro de la ciencia, analizando con todo detalle los inicios de la teoría cuántica en torno a Max Planck. A pesar de que se trataba de un caso que ni pintado para desplegar la concepción romántica y heroica de la ciencia que Kuhn había hecho famosa en los años sesenta, Kuhn renuncia ahora por completo a ella y las palabras «revolución científica», «paradigma», «ciencia normal», «ciencia revolucionaria», etc., no aparecen ni una sola vez en su voluminoso libro.

La epistemología de los años setenta (y de lo poco que llevamos de los ochenta) marca una vuelta de las aguas a sus cauces normales, una cierta superación de la crisis historicista de los sesenta y un nuevo inicio, que recoge la tradición de la filosofía clásica, pero asume también las críticas y orientaciones de los historiadores. Esta filosofía actual de la ciencia se diferencia de la clásica en su atención mucho mayor al análisis y reconstrucción de conceptos y teorías realmente usadas por la comunidad científica (en este sentido puede decirse que ha aceptado y asumido la crítica historicista). Pero se diferencia de la epistemología historicista de los años sesenta en el decidido empeño

---

** *La estructura de las revoluciones científicas*, trad. cast. de Agustín Contín, México, FCE, 1971.

*** Thomas S. Kuhn, *La teoría del cuerpo negro y la discontinuidad cuántica*, 1894-1912. Versión española de M. Paredes Larrucea. Alianza Editorial, 1980.

de no quedarse en la mera narración de casos concretos, sino de proceder a la construcción de esquemas abstractos con que dilucidar la estructura y el funcionamiento de las teorías, usando para ello todo el arsenal de herramientas formales que la lógica y la matemática ponen a su disposición. De todos modos conviene señalar un cambio. Mientras los epistemólogos clásicos sólo disponían de los lenguajes formales y los cálculos deductivos, que son herramientas conceptuales de manejo engorroso y complicado, los epistemólogos de los años sesenta utilizan sobre todo el instrumentario conceptual de la teoría de conjunto y de modelos, mucho más flexible y de más fácil manejo.

En 1970 aparece la obra de P. Suppes *Set-Theoretical Structures in Science* (un curso multicopiado llamado a ejercer gran influencia) y G. Ludwig publica *Deutung des Begriffs «Physikalische Theorie» und Grundlegung der Hilbertraumstruktur der Quantenmechanik durch Hauptsätze des Messens*. En 1971 se publican *Foundations of Measurement*, de Krantz, Luce, Suppes y Tversky, y *The Logical Structure of Mathematical Physics*, de J. Sneed. No hay más que echar una ojeada a estas obras para ver que una nueva etapa se inicia, la etapa en la que todavía estamos.

A pesar de su juventud, en Ulises Moulines se aprecia bien la influencia que han dejado las tres etapas que acabamos de describir, y que constituyen como otros tantos estratos de su obra.

En su época de estudiante Ulises Moulines se interesaba apasionadamente por la filosofía clásica de la ciencia, incluso por sus ya lejanos orígenes, por las empresas intelectuales ya abandonadas por sus propios iniciadores, como las reconstrucciones del mundo a partir de los datos sensibles (y la lógica) intentadas por Mach, Russell, Carnap y Goodman. Fruto de este interés fue su libro *La estructura del mundo sensible*, escrito en 1971 y publicado en 1973, una obra ajena a toda moda, precisa, sólida, audaz y vigorosa como las creaciones del círculo de Viena. El permanente interés de Moulines por los orígenes y precedentes de la filosofía clásica de la ciencia se manifiesta también en algunos capítulos de este libro, como en el dedicado a *La génesis del positivismo en su contexto histórico*.

La crisis historicista de los años sesenta y las ideas de Kuhn también tuvieron su impacto en Moulines. Este impacto se aprecia en su interés genuino por el aspecto diacrónico, evolutivo, de las teorías científicas, manifiesto en toda la tercera parte de este libro y especialmente en el capítulo dedicado a *La evolución de la mecánica newtoniana*, en que los puntos de vista histórico y formal se funden armoniosamente para proporcionarnos una visión sinóptica y rigurosa de la evolución de la mecánica clásica de partículas hasta fines del siglo XVIII. La influencia de Kuhn es patente en varios lugares, como en el capítulo *Forma y contenido de las revoluciones científicas: el caso de la mecánica newtoniana*, donde la aportación de Newton a la mecánica se encaja con buena voluntad en la jerga kuhniana de las

revoluciones científicas y de los componentes de la matriz disciplinaria (el sustituto kuhniano de 1969 para la anterior noción de paradigma). También en la segunda parte la influencia kuhniana está presente, a veces incluso demasiado presente (en mi opinión, claro), como en la propuesta de incluir las comunidades científicas dentro de la estructura de la teoría, contenida en el capítulo *Redes teóricas*.

La filosofía de la ciencia de los años setenta es, con mucha diferencia, la que más ha influido en Moulines. Una vez acabados sus estudios en Barcelona, acudió a Munich, donde conoció a J. Sneed, uno de los iniciadores de la nueva epistemología, a W. Stegmüller, un filósofo clásico de la ciencia entonces recién convertido al sneedismo, y a W. Balzer, un compañero de Ulises. Desde entonces Sneed, Stegmüller, Moulines y Balzer han trabajado conjuntamente y han propagado y hecho avanzar el programa de investigación epistemológica del primero. Moulines se doctoró bajo la dirección de Stegmüller y ha publicado artículos conjuntos con Sneed y Balzer. La influencia de Sneed en todo este libro, sobre todo en su segunda parte, es tan evidente que no necesita ser subrayada. Moulines ha escrito también la recensión más detallada del libro de Sneed. La influencia de los otros epistemólogos de los setenta es también visible. Así, el capítulo sobre *Aplicación aproximativa de teorías empíricas* (uno de los más logrados y originales de este libro) se basa en la idea de G. Ludwig de aplicar la estructura matemática de uniformidad a la dilucidación de la noción de aproximación, noción importantísima, pero completamente descuidada por la epistemología clásica. Y del interés de Moulines por la obra de Suppes es buen testimonio su artículo, escrito en colaboración con Sneed, sobre la filosofía de la física de Suppes, sin duda el estudio más completo sobre la misma.

Ulises Moulines no sólo ha sido influido por la filosofía de la ciencia de los años setenta, sino que a su vez él ha influido sobre ella. Esta influencia se ha ejercido dentro del contexto de la escuela Sneed-Stegmüller-Moulines-Balzer, pero dentro de ese contexto su influencia ha sido grande. Uno de los puntos flacos de esa escuela consistía en que, desde sus inicios, su análisis se centró en una sola teoría seria: la mecánica clásica. Una de las aportaciones de Moulines ha consistido en reconstruir un segundo caso real de aplicación de los esquemas sneedianos: el caso de la termodinámica de sistemas simples en equilibrio. Esta aportación se refleja aquí en *La red teórica de la termodinámica de los sistemas simples* (sin duda uno de los capítulos más importantes del libro) y en *Marcos teóricos: el ejemplo de la termodinámica*. Incluso en el análisis de la mecánica clásica Moulines ha aportado novedades dentro de la escuela sneediana. La principal de ellas consiste en su nueva visión del segundo principio de Newton como principio-guía. Esta concepción se encuentra expuesta en *Forma y función de los principios-guía en las teorías físicas*, un capítulo de dudosa precisión formal, pero de gran vigor, profundidad

y originalidad. Naturalmente quien se queda en los meros formalismos o en la mera historia nunca tendrá los problemas formales con que tropieza Moulines al introducir su noción de principios-guía. Que no rehúya esos problemas, aun cuando no logre dominarlos del todo, y que no renuncie ni a su profunda intuición ni al intento de formalizarla, le honra a él y constituye un reto para nosotros, sus lectores, a seguir avanzando por ese camino.

Este libro es básicamente un libro de filosofía de la física. La filosofía de la matemática, de la biología, etc., no aparecen aquí. Dentro de la filosofía de la física es fundamentalmente un libro de filosofía de la física clásica, sobre todo de la mecánica y la termodinámica clásicas. Las teorías del electromagnetismo, de la mecánica cuántica, de la relatividad especial y general, etc., tampoco son aquí analizadas. Pero desde un punto de vista filosófico parece adecuado proceder a un análisis de las teorías de la física clásica para afinar los instrumentos que más tarde, quizás, se aplicarán a otras teorías más complejas. De hecho todavía no hemos logrado una comprensión suficiente de ninguna teoría, y avanzar en la clarificación de la estructura de alguna teoría física ya representa un gran progreso. A algún lector le puede parecer que, desde el punto de vista del área cubierta, este libro da poco. Pero ningún libro de la nueva epistemología da más. Todo está aún *in fieri* y la hora de las grandes síntesis *à la Nagel* todavía no ha sonado.

Una de las consecuencias del estado embrionario de la actual filosofía de la ciencia estriba en su dispersión en diversas escuelas, parroquias o programas de investigación, que, si bien tienen mucho en común (los métodos formales conjuntistas y de teoría de modelos, la atención detallada a teorías reales, etc.), discrepan con frecuencia en sus esquemas y terminología. La escuela de Sneed-Stegmüller-Moulines-Balzer tiene su propia jerga y sus propios esquemas, que no todo el mundo considera como los más simples, precisos o adecuados, pero que ellos aplican con tenacidad un poco escolástica, en la que sin embargo se esconde la promesa de la coherencia interna y del progreso continuado. De todos modos Moulines también sabe escribir con una gracia y un desparpajo poco frecuentes en la prosa filosófica, como el lector puede comprobar leyendo los dos últimos capítulos de este libro, ligeros, inteligentes e irónicos.

Mi amigo, antiguo discípulo y actual maestro Ulises Moulines, es el filósofo de la ciencia más descollante de la joven generación latinoamericana e hispánica. Leyendo este libro tan estimulante, que resume gran parte de su labor en los años setenta, he aprendido mucho y lo he pasado muy bien. Ojalá los demás lectores saquen tanto provecho y placer de su lectura como yo.

Moyá, junio de 1981.

JESÚS MOSTERÍN.

PREFACIO DEL AUTOR

El presente libro no es un tratado. No se pretende en él exponer una visión sistemática y completa de la filosofía de la ciencia, ni siquiera de un área particular dentro de ella. Dado el actual estado «fluido» de la disciplina, no parece razonable aventurarse en empresas totalizantes y sistematizadoras –y menos por parte de un sólo autor–. Más bien, lo que se ofrece en este libro es un *collage* de indagaciones metacientíficas que, en diversas direcciones y desde distintos puntos de apoyo, el autor ha emprendido en los últimos siete años. Algunos temas han sido tratados con más detalle que otros, algunos resultados conseguidos aparecen como más firmes que otros, pero en ningún caso se ha pretendido decir la última palabra. Muchos filósofos de nuestra época han intentado, justificada pero infructuosamente, hacer ver a los científicos el carácter provisional, problemático, incierto, de su ciencia; pero pocos han sido plenamente conscientes de que lo mismo se aplica, y aún en mayor medida, a su propia filosofía.

Las diversas e inconclusas líneas de investigación que aparecen en este libro no se hallan, por supuesto, aisladas unas de otras. Se entrecruzan y complementan, no para constituir un edificio acabado, pero sí para formar una madeja que esperamos tenga algún valor para apresar ciertos aspectos de esto que llamamos «ciencia». Estas diversas líneas se pueden agrupar en tres haces distintos, aunque relacionados: la explicitación de las estructuras sincrónicas subyacentes a las teorías científicas; aspectos diacrónicos concretos de la ciencia y de su metodología, y el análisis de ciertos supuestos ontológicos. A estos tres grupos de investigaciones se anteponen unas «consideraciones metafilosóficas» de carácter general.

Parte del material contenido en este libro ha sido publicado previamente en forma de artículos, en castellano, inglés o alemán. En todos estos textos, sin embargo, se han hecho ciertas modificaciones. En algunos casos las modificaciones se han limitado a cambios de terminología, estilo o notación; en otros, se trata de revisiones sustanciales. Algunos capítulos, finalmente, se publican enteramente aquí por primera vez. A continuación se dan las referencias de los capítulos que contienen material previamente publicado:

— El capítulo 1.2, «Qué hacer en filosofía de la ciencia. Una alternativa en catorce puntos», apareció como artículo en la revista *Crítica* (México), núm. 32, en agosto de 1979.
— El capítulo 2.1, «Hacia un nuevo concepto de teoría empírica», corresponde aproximadamente a un artículo que con el mismo título se publicó en la revista *Convivium* (Barcelona), núm. 39, en 1973.
— El capítulo 2.2, «Un programa de reconstrucción estructural de las teorías físicas», apareció en la *Revista Latinoamericana de Filosofía* (Buenos Aires), núm. III, 2, en julio de 1977.
— El capítulo 2.3, «Forma y función de los principios-guía en las teorías físicas», corresponde a «Cuantificadores existenciales y principios-guía en las teorías físicas», *Crítica* (México), núm. 29, agosto de 1978.
— El capítulo 2.6, «La red teórica de la termodinámica de los sistemas simples», está basado en la traducción, revisión y ampliación de «A Logical Reconstruction of Simple Equilibrium Thermodynamics», *Erkenntnis*, 9, I, 1975.
— El capítulo 2.7, «Aplicación aproximativa de teorías empíricas», es una traducción y revisión de «Approximate Application of Empirical Theories: A General Explication», *Erkenntnis*, 10, II, 1976.
— El capítulo 2.9, «Aproximación interteórica: el caso Kepler-Newton», es traducción y revisión de «Intertheoretic Approximation: The Kepler-Newton Case», *Synthese*, núm. 45, 1980.
— El capítulo 3.1, «Forma y contenido de las revoluciones científicas: el caso de la mecánica newtoniana», fue publicado en la antología *La filosofía y las revoluciones científicas* (comp. por A. Sánchez Vázquez *et al.*), UNAM, México, 1979.
— El capítulo 3.2, «Los fundamentos metodológicos de la filosofía natural de Isaac Newton», se publicó en *Dianoia* (México), 1976.
— El capítulo 3.3, «La evolución de la mecánica newtoniana», está basado en parte de «Theory-nets and the Evolution of Theories: The Example of Newtonian Mechanics», *Synthese*, 41, 1979.
— El capítulo 3.4, «La epistemología fisiológica de Hermann von Helmholtz», es traducción y ampliación de una primera versión en alemán, «Helmholtz' sinnesphysiologische Erkenntnistheorie», publicada en *Epistemologische und soziale Probleme der Wissenschaft-*

sentwicklung im frühen 19. *Jahrhundert* (antología editada por la Universidad de Bielefeld), Bielefeld, 1979, y de una versión posterior en inglés, «Hermann von Helmholtz: A Physiological Approach to the Theory of Knowledge», publicado en *Epistemological and Social Problems of the Sciences in the Early Nineteenth Century* (comp. por H. N. Jahnke y M. Otte), Reidel, 1980.

— El capítulo 3.5, «La génesis del positivismo en su contexto científico», se publicó originalmente en *Dianoia*, México, 1975.

— El capítulo 4.1, «Ontosemántica de las teorías científicas», es una ampliación y revisión del artículo que con el mismo título apareció en *Teorema*, X, 1, Valencia, 1980.

— El capítulo 4.3, «Los problemas del materialismo», apareció bajo el título «Por qué no soy materialista» en *Crítica*, 26, México, 1977.

El resto del libro es completamente inédito en el momento de escribir estas líneas.

Las siguientes personas han dejado su huella, con sus observaciones críticas, en el material presentado en diversos capítulos de este libro: Wolfgang Balzer (Munich), William Craig (Berkeley), Adolfo García de la Sierra (Morelia), Ignacio Jané (México), Jesús Mosterín (Barcelona), J. Daniel Quesada (Barcelona), Erhard Scheibe (Gotinga) y Kurt Stützer (Heidelberg). No es de ellos la culpa si sus observaciones no han sido aprovechadas suficientemente. A Jesús Mosterín y J. Daniel Quesada debo especial agradecimiento por haber leído el manuscrito con tanto cuidado. A la señora Gabriela Vázquez le agradezco su paciente ayuda al mecanografiar una gran parte del manuscrito.

México, D. F., noviembre de 1980.

RECORDATORIO DE SIMBOLOS LOGICOS Y MATEMATICOS

Aunque en este libro se ha procurado mantener a un nivel mínimo el grado de complicación formal en la exposición de los diversos temas, en algunos capítulos ha sido necesario usar nociones técnicas de las ciencias formales. En los casos en que estas nociones están poco divulgadas fuera del área de especialistas que trabajan con ellas, se han explicado con cierto detalle en el lugar pertinente (es el caso, por ejemplo, de nociones como las de «gradiente» y «uniformidad»). La mayor parte del aparato formal utilizado, sin embargo, pertenece al acervo común de lógica, teoría elemental de conjuntos y cálculo elemental del que dispone el público universitario en general. No parece, pues, apropiado explicar dicho aparato en este libro. El lector que tenga alguna dificultad de comprensión frente a él puede acudir a los numerosos libros de texto pertinentes.

No obstante, a modo de recordatorio y dado que la notación utilizada para estas nociones formales no siempre es uniforme, se resumen aquí brevemente los símbolos utilizados y su significado.

De la lógica:

— Si P es un predicado cualquiera, $\hat{P}$ indica su *extensión;* o sea, la clase de objetos que poseen la propiedad expresada por P.

— Si p es un enunciado cualquiera, $\neg p$ es su *negación;* $\wedge$ indica la *conjunción* de enunciados, $\vee$ la *disyunción*, $\rightarrow$ el *condicional* y $\leftrightarrow$ el *bicondicional.*

$\forall x$ es el *cuantificador universal* y $\exists x$ el *cuantificador existencial.* $\exists!x$ es la abreviación para «existe un x y *sólo uno...*».

De la teoría de conjuntos:

— Utilizaremos los términos «conjunto» y «clase» como sinónimos.

— «$a \in A$» significa «a es *elemento* del (o pertenece al) conjunto A».

— $\{a, b, c, ...\}$ indica el conjunto consistente en los elementos a, b, c, ...

— Si A es un conjunto, $\|A\|$ indica su *cardinalidad*; o sea, el número de elementos de A.

— $\varnothing$ es el *conjunto vacío*.

— Si A y B son dos conjuntos, $A \cup B$ es su *unión*, $A \cap B$ su *intersección* y $A - B$ su *diferencia*.

$A \subseteq B$ indica que A es un *subconjunto* de B.

$A \subset B$ indica que A es un *subconjunto propio* de B; es decir, que A es subconjunto de B, pero no a la inversa.

Si A es un conjunto, $\mathscr{P}(A)$ es el *conjunto potencia* de A; es decir, el conjunto de todos los subconjuntos de A.

Dada la sucesión de conjuntos $A_1, ..., A_n$, $\bigcup_{i=1}^{n} A_i$ es la abreviación para $A_1 \cup ... \cup A_n$. Si $\bar{A} = \{A_1, ..., A_n\}$ es un conjunto de conjuntos, entonces $\cup \bar{A}$ indica la «gran unión» de $\bar{A}$; es decir, la unión de todos los conjuntos A_i que son elementos de $\bar{A}$. *Mutatis mutandis* vale para la intersección.

— $\langle a, b \rangle$ es el *par ordenado* de los elementos a y b. $\langle a_1, ..., a_n \rangle$ es el *tuplo ordenado* o secuencia de los n elementos $a_1, ..., a_n$. Para indicar que un elemento a_i aparece dentro del tuplo $\langle a_1, ..., a_n \rangle$ usaremos la abreviación $a_i \hat{\in} \langle a_1, ..., a_n \rangle$. (Nótese que, formalmente, es una relación distinta, aunque definible a partir de la relación conjuntista fundamental $\in$.)

Si A y B son dos conjuntos (iguales o distintos), $A \times B$ indica su *producto cartesiano*, es decir, el conjunto de todos los pares ordenados $\langle a_i, b_i \rangle$ tales que $a_i \in A$ y $b_i \in B$. Si $A_1, ..., A_n$ son n conjuntos (iguales o distintos), $A_1 \times ... \times A_n$ indica el producto cartesiano de los mismos; es decir, el conjunto de todos los n-tuplos $\langle a_1, ..., a_n \rangle$ con $a_i \in A_i$. En el caso de que todos los n A_i sean iguales, escribiremos para abreviar A^n en vez de $A \times ... \times A$ n veces.

— Si R es una relación *diádica*, entonces $D_I(R)$ indica su *dominio* (el conjunto de los primeros términos de la relación) y $D_{II}(R)$ su *contradominio* o *recorrido* (el conjunto de los segundos términos). Lo mismo vale, naturalmente, para las funciones, que, como es sabido, se pueden considerar como casos particulares de relaciones. Si f es una función, A su dominio y B su contradominio, a menudo lo indicaremos así: $f: A \rightarrow B$.

Del cálculo:

— $\mathbb{N}$ indica el conjunto de los *números naturales*, $\mathbb{R}$ el conjunto de los *números reales*, $\mathbb{R}^+$ el conjunto de los *números reales positivos*, $\mathbb{R}^3$ el *espacio euclídeo tridimensional*. A los elementos de $\mathbb{R}^3$ (que son triplos de números reales) los llamamos «vectores». Si $r=\langle x, y, z\rangle$ es un vector, $|r|$ es su *magnitud;* o sea, $|r|=\sqrt{x^2+y^2+z^2}$.

— Si a y b son números reales, $/a-b/$ es el *valor absoluto* de su diferencia.

— $\sum_{i=1}^{n} a_i$ indica la suma de una serie de n números reales a_i. Cuando no está fijado el número n, escribimos simplemente $\sum_i a_i$.

— Si f es una función monádica real diferenciable, Df representará la función correspondiente. Si f es una función de n argumentos, $f(a_1, ..., a_n)$, y es diferenciable en cada uno de ellos, entonces $D_{a_i}f$ representa la *derivada parcial* de f respecto al argumento a_i.

1. CONSIDERACIONES METAFILOSOFICAS

Es característico de cualquier tipo de empresa filosófica el cuestionarse a sí misma. Los filósofos no suelen limitarse a realizar su labor filosófica, sino que parece como si sintieran la necesidad de preguntarse frecuentemente por la naturaleza de su propia actividad. La *reflexión* filosófica sobre la misma filosofía suele denominarse «metafilosofía». Parece que la metafilosofía es un acompañante inevitable, o al menos natural, de la filosofía. A este principio (metametafilosófico) no escapa la filosofía de la ciencia. También el filósofo de la ciencia puede plantearse con sentido ciertas preguntas acerca de la naturaleza y propósitos propios de su empresa filosófica específica. Incluyo en esta primera parte dos capítulos de carácter básicamente metafilosófico. El primero pretende ser no sólo relevante para la filosofía de la ciencia, sino para la filosofía en general; aun así, las ideas contenidas en él se ponen a prueba ante un ejemplo tomado de la filosofía de la ciencia: el de la distinción entre conceptos teóricos y observacionales.

El segundo se refiere más específicamente a la filosofía de la ciencia. Propone un programa metodológico con varios apartados, un programa que algunos probablemente consideran ajeno a lo que normalmente se ha entendido por filosofía de la ciencia; sin embargo, no sólo creo haber dado una justificación general de este programa, sino que los ensayos incluidos en los capítulos subsiguientes pueden considerarse como una implementación del mismo. (Es frecuente que un programa general se conciba *después* de que partes del mismo ya se han llevado a cabo; así ha sido en este caso.)

1.1. BLANCO, NEGRO, GRIS: CONTRA EL EXTREMISMO FILOSOFICO

125679

1. Entre las muchas peculiaridades de la especie humana se cuenta su notable capacidad recursiva. Por ejemplo, los hombres no sólo pueden hacer la guerra, sino también la guerra a la guerra, la guerra a los que hacen la guerra a la guerra, etc. Asimismo, no sólo podemos hablar de las cosas, sino además hablar sobre el hablar de las cosas, y así sucesivamente. Como es sabido, una de las formas en que se habla sobre el hablar de las cosas es la forma filosófica. Se hace filosofía *del* lenguaje común, filosofía *de* la ciencia, filosofía *de* la moral, etc., donde lenguaje común, ciencia, moral, son formas de hablar de las cosas. Pero dada la general capacidad recursiva de los hombres, no es de extrañar que algunos también empiecen a hablar sobre esa forma especial de hablar sobre el hablar de las cosas a la que llamamos filosofía. Esto, a su vez, puede hacerse de varias maneras (sociología de la filosofía, historia de la filosofía, etc.). Una de esas maneras es la manera filosófica. Entonces se hace filosofía de la filosofía o, si se prefiere, «metafilosofía». (Por supuesto, el nombre es lo de menos.)

Aquellos que miran la filosofía con malos ojos, o por lo menos con escepticismo, pensarán sin duda que si hacer filosofía es ya de por sí una actividad dudosa, hacer metafilosofía es simplemente el colmo; pensarán que no es más que «rizar un rizo», el cual ya en sí mismo es «rizar el rizo». Sea cual sea nuestra valoración de este asunto, creo que este «rizar el rizo» filosófico es inevitable –lógica y psicológicamente–. Una vez constituida la física, era inevitable que alguien se pusiera a filosofar sobre la física y luego a filosofar sobre la filosofía de la física, y análogamente en otros campos. Sólo una sociedad muy

bruta y totalitaria podría impedirlo. En circunstancias mínimamente favorables es inevitable filosofar e inevitable también filosofar sobre la filosofía. Tan inevitable como hablar sobre el hablar sobre el hablar sobre... Sirva esto de justificación para el contenido de este capítulo: consideraciones filosóficas sobre la filosofía.

En muchos casos el «hablar de *x*» y el «hablar sobre el hablar de *x*» están bien deslindados, son actividades con distintas reglas y distintos objetivos. En cambio, es difícil separar netamente la metafilosofía de la filosofía, como también resulta cada vez más difícil separar la metamatemática de la matemática. Filosofía y matemática son actividades que inevitablemente, por su propia naturaleza y conveniencia, se ven llevadas a rizar el rizo.

Una de las tareas más típicas de la filosofía es establecer buenas distinciones conceptuales para aclarar las cosas. El presente capítulo contiene un par de reflexiones sobre cómo hay que ejecutar esa tarea. Es, por tanto, un artículo metafilosófico. Pero de lo que he dicho en el párrafo anterior se desprende que no será posible no hacer, al mismo tiempo, un poco de filosofía. Concretamente trataré de esbozar muy brevemente (aquí no puede ser de otro modo) la forma en que ciertas consideraciones metafilosóficas pueden guiarnos en la tarea de establecer determinadas distinciones filosóficas muy discutidas.

2. Al principio metafilosófico central que voy a establecer para empezar lo llamaré «*Principio de la Relevancia de las Distinciones Graduales*» (para abreviar: «Principio RDG»). Este principio tiene dos caras: una positiva y otra negativa. En su forma positiva, el principio RDG puede enunciarse así:

«*Son filosóficamente relevantes las distinciones conceptuales que atienden sólo a diferencias de grado y no a diferencias absolutas en el objeto o dominio de estudio.*»

Y en su forma negativa:

«*Son filosóficamente peligrosas, y frecuentemente perniciosas, las distinciones conceptuales tajantes que pretenden determinar supuestas diferencias absolutas en el objeto o dominio de estudio.*»

El principio RDG, tanto en su enunciación positiva como en la negativa, es un principio-guía sumamente general, y a mi entender válido en cualquier campo del discurso filosófico; y no sólo del filosófico, sino también del científico en general, con la posible excepción de la lógica y la matemática.

Es un vicio muy extendido en las discusiones filosóficas de todos los tiempos y de todas las corrientes ignorar o incluso rechazar

tácitamente el mencionado principio. Dos son las actitudes filosóficas posibles, contrapuestas, que se derivan de ese rechazo. O bien se adopta una actitud de «o-blanco-o-negro», en la que el filósofo absolutiza una distinción y pretende establecer dicotomías tajantes con su aparato conceptual; o bien (generalmente por reacción) se adopta la actitud contrapuesta de «todo-gris», en la que se niega que se puedan hacer distinciones válidas en el tema discutido.

Ambas actitudes están abundantemente ejemplificadas en la filosofía del siglo XX. Menciono sólo tres ejemplos. En epistemologia se intentó durante cierto tiempo establecer una dicotomía absoluta entre «ciencia» y «metafísica»; cuando se observaron dificultades graves en ese intento, algunos filósofos tomaron la actitud opuesta, e igualmente extrema, de sostener que no es posible distinguir de modo alguno entre ciencia y metafísica. En filosofía del lenguaje, a la tesis absolutista extrema de que en *todo* lenguaje y en *toda* disciplina es posible trazar una neta dicotomia entre enunciados analíticos y sintéticos sucedió la antítesis relativista extrema de que la distinción entre lo analítico y lo sintético no tiene sentido en *ningún* lenguaje ni en *ninguna* disciplina. Finalmente, en filosofía de la ciencia estamos contemplando cómo de la distinción, que algunos autores consideran absoluta, entre términos teóricos y observacionales se pasa al rechazo, no menos absoluto, de dicha distinción.

En las últimas décadas se ha notado cierta tendencia general a pasar de las filosofías de «o-blanco-o-negro» a las de «todo-gris»; pero esto no tiene por qué ser más que una moda pasajera, puesto que las segundas son tanto o más estériles que las primeras y, en cualquier caso, igual de precipitadas.

Este permanente conflicto entre filosofías de trazo totalmente nítido y filosofías de dibujo totalmente difuso puede ser considerado por algunos como un interesante ejemplo de «proceso dialéctico» y, por otros, como ilustración de la sentencia de que «la historia de la filosofía es la historia de una lucha de temperamentos». En cualquier caso, y con independencia del valor que tales consideraciones puedan tener para la historia del pensamiento, lo que me parece claro es que ambas actitudes filosóficas son arbitrarias y paralizantes en la mayoría de los casos.

Los peligros de ambas actitudes filosóficas se evitan si se adopta el principio RDG. En efecto, en su formulación negativa, RDG contradice al filósofo de «o-blanco-o-negro»: ¡desconfiemos de las distinciones absolutas y tajantes!; mientras que su formulación positiva es contraria a la actitud «todo-gris»: es conveniente hacer distinciones conceptuales, aun cuando sepamos que obedecen a gradaciones; o, dicho de otro modo: que una diferencia sea de grado no implica que no sea importante.

A veces ocurre que, una vez establecida una distinción más o menos clara y operativa entre diversos ítems, se observa que ella

obedece a la variación continua de cierta propiedad subyacente. Esto permite comprender, por un lado, que la distinción primitiva es de carácter gradual y no tajante; pero, por otro lado, en vez de negar la validez de la distinción, la justifica. Así, por ejemplo, la distinción entre rojo y azul es bastante clara y útil en la vida cotidiana, por lo que quizá se le ocurriera a algún filósofo de «o-blanco-o-negro» sostener que hay una diferencia «absoluta y esencial» entre lo rojo y lo azul. El progreso de la óptica refutó a ese filósofo mostrando que la diferencia entre el rojo y el azul obedece a un cambio gradual y continuo en la frecuencia de ciertas ondas electromagnéticas. Ahora bien, no por ello el óptico se pasa al bando del filósofo de «todo-gris» y afirma que no existe una verdadera diferencia entre los colores; las distinciones cromáticas, a pesar de reconocerse como graduales, siguen considerándose relevantes, no sólo en la vida cotidiana, sino también en la física (en la espectrografía, por ejemplo).

No todos los filósofos han sido adversos al principio RDG. Algunos, quizá bastantes, lo han asumido implícitamente. Se me ocurren, sin pensar mucho, Aristóteles y el segundo Wittgenstein. Entre nuestros contemporáneos, Ferrater Mora es un ejemplo notable. Creo que su larga práctica en los análisis filosóficos y metafilosóficos lo ha confirmado como seguidor constante del principio RDG *avant la lettre*. Y quizá incluso podría decirse que este principio es un elemento importante de su metodología integracionista. Especialmente en *El ser y la muerte* y en *El ser y el sentido*, Ferrater Mora ha mostrado cómo algunas dicotomías tajantes hechas tradicionalmente en filosofía pueden reinterpretarse sensata y operativamente como «conceptos-límite» que en sí mismos no se refieren a ninguna realidad, pero que sirven de polos de referencia entre los cuales sí pueden hacerse distinciones graduales con referente real.

Creo que Ferrater Mora aplica implícitamente la formulación negativa del principio RDG cuando afirma: «... las polaridades sirven justamente para demarcar y situar las realidades efectivas. Así, por ejemplo, nada es ni "pura materia" ni "puro espíritu"; nada es o "absolutamente sujeto" o "absolutamente objeto"; nada es o "puramente mecánico" o "puramente orgánico", etc. Los predicados "es material", "es espiritual", "es mecánico", "es orgánico", etc., no son, pues, en ningún caso predicados absolutos, son expresiones por medio de las cuales se "sitúan" ontológicamente diversos tipos de entidades, esto es, por medio de las cuales tales tipos de entidades se cualifican ontológicamente en términos de "más" o "menos". Según ello, puede decirse que algunas entidades son "más materiales" o "menos materiales", "más orgánicas" o "menos orgánicas", "más conscientes" o "menos conscientes" que otras» (*El ser y la muerte*, *Obras selectas*, II, p. 301).

Este pasaje, y en particular el último párrafo, podrían entenderse en el sentido de que al negarse la validez absoluta de las dicotomías,

se propusiera un «puro gradualismo», típico de los filósofos de «todo-gris». Sin embargo, ésta sería, según creo, una mala interpretación del integracionismo. Ferrater Mora es consciente de la otra cara de la medalla, es decir, de la formulación positiva de RDG cuando más abajo nos advierte ante un ejemplo concreto: «No propongo la confusa tesis de que la materia es espiritual y el espíritu es material o, en general, que cualquier cosa es todo –o que todo es cualquier cosa–. Si X es un trozo de materia, y esto es lo que X es, entonces es innegable que X no es espiritual o consciente. Pero aun negar conciencia a X es considerar, o poder considerar, X desde el punto de vista de la conciencia» (*op. cit.*, p. 302). Es decir, la metodología integracionista no nos lleva, en este caso concreto, a echar por la borda la distinción entre lo «material» y lo «espiritual» (quizá podríamos hacerlo, pero por *otras razones*), sino que permite, e incluso justifica, operar con ella, siempre y cuando seamos conscientes de que se trata de una distinción relativa a cierto punto de vista o, si se prefiere, a cierta *conceptuación de la realidad*.

3. Porque debería estar claro (aunque muchas veces se olvide) que las distinciones conceptuales no las *descubrimos*, sino que las *forjamos*. El dominio que estudiamos no está en sí mismo dividido en tales o cuales categorías, sino que somos nosotros los que así lo dividimos, de acuerdo con cierto aparato conceptual y con ciertos objetivos; del mismo modo como «el libro de la Naturaleza» *no está* «escrito en caracteres matemáticos», ni en ningún otro tipo de caracteres, sino que somos nosotros quienes *lo escribimos* en latín o en álgebra tensorial. Esto debería ser una perogrullada filosófica por lo menos desde Kant, pero muchas discusiones sobre determinadas distinciones conceptuales parecen mostrar que es una perogrullada ignorada. Así, se discute apasionadamente sobre si «hay» enunciados analíticos o no, sin que suela estar claro previamente desde qué punto de vista queremos hacer la distinción y por qué la queremos hacer. O bien, corren ríos de tinta sobre la distinción «teórico-observacional», sin que se aclare antes si vamos a hacerla dentro de la psicología del conocimiento humano, dentro de la teoría del conocimiento en general o dentro de la filosofía de la ciencia *sensu stricto* (tres campos muy distintos); y ya dentro de la filosofía de la ciencia, no suele especificarse si la distinción pretende ser de carácter semántico (atendiendo a los significados de los términos), o bien pragmático (atendiendo a la función que cumplen los términos en una teoría científica). O, en fin, en la famosa controversia sobre las relaciones entre el «ser» y el «deber ser» habría que tener en cuenta que no es lo mismo hacer una distinción entre enunciados descriptivos y prescriptivos, que hacerla entre enunciados descriptivos y valorativos –dos puntos de vista distintos que se suelen confundir.

Cuando emprendemos (inventamos, forjamos, construimos, pero nunca: descubrimos) una clasificación de cierto dominio en categorías, pretendemos acercarnos, implícita o explícitamente, al modelo ideal de una partición de un conjunto en clases de equivalencia. Debería estar claro después de la discusión del principio RDG que este modelo ideal es peligroso, porque su exactitud nos puede empujar hacia la trampa de las dicotomías absolutas. Este peligro es evitable si tomamos el modelo de la partición en clases de equivalencia como lo que es: un ideal, pocas veces realizable, que actúa como concepto-límite en el sentido de Ferrater Mora. Pero no es éste el peligro que quería hacer notar ahora. Hay otro peligro, del que no es responsable el modelo ideal en sí mismo, pero que puede aclararse en sus términos.

Sabemos que la mejor forma de obtener una partición en clases de equivalencia sobre un dominio dado consiste en determinar una relación de equivalencia en ese dominio. Ahora bien, también sabemos que, para un mismo dominio, es posible determinar diversas relaciones de equivalencia a la vez (por tanto, diversas particiones), varias de las cuales pueden ser igual de interesantes para estructurar el dominio de alguna forma. No hay razón alguna para que no las tengamos en cuenta todas a la vez. Es más, lo conveniente para una mejor comprensión conceptual de la situación es determinar todas las relaciones de equivalencia que parezcan interesantes desde algún punto de vista para «partir» el dominio en cuestión. Esto que es válido para la matemática pura lo es, si cabe, todavía más para las distinciones empíricas y también, no se olvide, para las filosóficas.

Es frecuente que varias de las relaciones de equivalencia definibles sobre un mismo dominio tengan un cierto «aire de familia» entre sí, se parezcan en algún sentido. *Intuitivamente* pueden parecer *la misma* relación. Pero aquí suele engañarnos la intuición, por lo menos en el caso de las distinciones filosóficas. El filósofo debe precaverse de este peligro y analizar con cuidado la relación de equivalencia con la que pretende categorizar el dominio estudiado. Si hace esto, generalmente observará que lo que había tomado por *una* relación de equivalencia es, en realidad, un conglomerado de *varias*, que hay que desbrozar.

Para aclarar las consideraciones anteriores es conveniente analizar con cierto detalle un ejemplo clásico: la controvertida distinción entre conceptos teóricos y observacionales.

Del principio RDG se desprende que la estrategia razonable en este caso sería tratar de determinar grados o, mejor dicho, niveles sucesivos de observacionalidad o, correspondientemente, de teoricidad. Los tres enunciados siguientes:

(1) «Esa estrella brilla menos intensamente que esa otra»;

(2) «esa estrella está más alejada de la Tierra que esa otra», y

(3) «debido al efecto Doppler puede conjeturarse que esa estrella está unos mil años-luz más alejada de la Tierra que esa otra»,

aunque pertenezcan a un mismo campo de nuestra experiencia, poseen distinto nivel de teoricidad. Esto es obvio para cualquiera que no sea un esnob filosófico. Por mucha que sea la «carga teórica» contenida en (1), ésta es menor que la contenida en (2) y muchísimo menor que la contenida en (3). El conjunto de personas que pueden manejar con soltura el enunciado (3) es un subconjunto propio del conjunto que puede manejar (2), y éste es, a su vez, un subconjunto propio de los que manejan (1). No considero filosóficamente fructífero tratar de inventarse argumentos más o menos divertidos y *épatants* para mostrar que tal distinción de niveles no es posible (cuando todos sabemos que la diferencia existe de hecho); como tampoco es fructífero absolutizar esa diferencia y sostener que revela algo epistémicamente esencial y «profundo». Lo que hay que hacer es ver a qué criterio (s) responden las distinciones entre los sucesivos niveles de teoricidad (u observacionalidad), cómo funcionan, cómo se relacionan entre sí.

Pero además, y previamente, hay que averiguar si la intuición primaria de la distinción teórico-observacional no confunde aspectos que deberían considerarse por separado. Y, en efecto, treinta años después de acaloradas discusiones sobre dicha dicotomía, a Bar-Hillel se le ocurrió por primera vez hacer notar[1] que muchas de las controversias se debían simplemente a la confusión de dos modos distintos de dicotomizar los términos empíricos: (a) en las categorías de «observacional» y «no-observacional»; (b) en las categorías de «teórico» y «no-teórico». Sin duda, ambas dicotomías estarán relacionadas de algún modo, por complicado que sea (y esta relación es algo que valdría la pena para estudiar); pero, en cualquier caso, no hay razón alguna para pensar que ambas dicotomías deban coincidir ni intensional ni extensionalmente. La primera pertenece propiamente a la teoría del conocimiento en general; la segunda, a la filosofía de las ciencias avanzadas.

Podemos ir más lejos que Bar-Hillel. Incluso la distinción entre conceptos teóricos y no-teóricos puede resultar ser un saco donde metemos cosas distintas, y esto porque hay diversas formas de entender el término «teoría». Cuando Hanson, por ejemplo, sostiene la tesis de que todo concepto, por «observacional» que parezca, está «cargado de teoría», esto no nos dice gran cosa si no sabemos previamente cómo hay que entender «teoría» en ese contexto. Ni Hanson ni la mayoría de los adalides de «todo-teórico» han explicado nunca claramente cuál es la noción de teoría que tienen en mente. Sólo Feyerabend, en un pasaje de *Problems of Empiricism*, I, hace notar, como de pasada, que por "teoría" entiende «cualquier cosa»:

[1] En su ensayo «Neorelism vs. Neopositivism: A Neo-Pseudo Issue» (1970).

cualquier sistema conceptual, conjunto de creencias, mito, ideología, etc.[2]. Ahora bien, si se entiende por "teoría" cualquier cosa, entonces la distinción entre conceptos teóricos y no-teóricos tiene efectivamente poco sentido, pero la contrapartida es que la tesis de la «carga teórica universal» deviene ininteresante por trivial.

Si, por el contrario, se entiende por "teoría" alguna porción concreta (y previamente explicitada) de alguna de las disciplinas científicas avanzadas que se exponen en los libros de texto de los estudiantes de ciencias actualmente, esta noción de teoría y la distinción concomitante resultarán mucho mejor acotadas y, según creo, también más interesantes. En tal caso, la distinción "teórico-no-teórico" no posee el carácter universal que han asumido tanto sus defensores como sus detractores habituales, sino que es una distinción relativa a cierta teoría *T* dada. Los criterios de *T*-teoricidad propuestos por Sneed en *The Logical Structure of Mathematical Physics*, por Tuomela en *Theoretical Concepts* y por mí (de manera aún informal) en *A Logical Reconstruction of Simple Equilibrium Thermodynamics* son todos, a pesar de sus diferencias respectivas, similares en el sentido apuntado de determinar la distinción "teórico-no-teórico" dentro de teorías concretas. En términos muy generales podríamos decir que todo este grupo de criterios de teoricidad se caracteriza por definir los conceptos *T*-teóricos como aquellos que son *específicos*, en algún sentido precisable, de la teoría (precisada) *T*.

Resumiendo, podemos afirmar que a la distinción apresurada entre conceptos observacionales y teóricos subyacen por lo menos tres relaciones de equivalencia distintas:

(a) «el concepto *x* posee el mismo grado de observacionalidad que el concepto *y*»;
(b) «el concepto *x* posee el mismo grado de teoricidad que el concepto *y*» ("teoricidad" aquí entendida en el sentido feyerabendiano de «teoría = cualquier cosa»);
(c) «el concepto *x* es tan específico de la teoría *T* como el concepto *y*».

Si se nos permite emplear una jerga algo más técnica, podemos decir que la distinción intuitiva "teórico-observacional" resulta ser, a fin de cuentas, una distinción «tridimensional» (pues son por lo menos tres los puntos de vista posibles para establecerla). Cada una de estas «dimensiones» debe ser analizada por separado, para averiguar si efectivamente es lo bastante interesante y fructífera como para que nos decidamos a operar con ella.

---

[2] «...the term 'theory' will be used in a wide sense, including ordinary beliefs, myths, religious beliefs, etc. In short, any sufficiently general point of view concerning matter of fact will be termed a "theory".» (P. K. Feyerabend, «Problems of Empiricism», nota 3, página 219.)

En general, una análisis cuidadoso de muchas diferencias conceptuales determinadas intuitivamente nos revelará que éstas, además de ser graduales (o «por niveles»), son «multidimensionales» (o «en distintas direcciones»). Es conveniente, pues, seguir la siguiente receta, a la que podemos denominar esta vez «*Principio de la Multidimensionalidad de las Distinciones Intuitivas*» («Principio MDI»):

«*Dada una distinción conceptual intuida presistemáticamente, hay que analizarla ulteriormente para determinar en cuántas "dimensiones" o direcciones distintas puede ser formulada sistemáticamente.*»

El principio MDI, aunque en rigor distinto del principio RDG, va asociado a él en espíritu, si no en la letra. Ambos principios pueden considerarse instancias de un super-principio metafilosófico general, al que podríamos llamar «*Principio de los Pies de Plomo Filosóficos*», y que podría resumirse en el lema: antes de hacer aseveraciones filosóficas contundentes, sondear bien el panorama conceptual en todos sus recovecos. Este es un principio análogo a los que, en sus respectivos campos, siguen el buen detective que investiga un crimen, el buen paleógrafo que transcribe un texto antiguo y, en general, cualquier persona sensata que intenta aclarar algo que está oscuro y que, para lograrlo, procede con cautela.

Habrá, sin duda, algunos (muchos) a quienes el Principio de los Pies de Plomo Filosóficos les parezca fastidioso, aburrido y, en definitiva, plúmbeo. Son aquellas personas que se han acercado a la filosofía en busca de arrolladoras y grandiosas visiones del tipo «o-blanco-o-negro» o «todo-gris», y no en busca de «lo que es el caso». A estas personas sólo se les puede recordar que un análisis cuidadoso en filosofía (lo mismo que en cualquier otra disciplina) no tiene por qué no ser pesado y aburrido para todo aquel que no esté metido en el asunto hasta el cuello.

1.2. ¿QUE HACER EN FILOSOFIA DE LA CIENCIA? UNA ALTERNATIVA EN CATORCE PUNTOS

0. El presente capítulo es de carácter programático. En él no se expone ninguna teoría ni se presentan resultados de la investigación sobre algún tema particular. Su única pretensión es esbozar las líneas generales de un programa para hacer filosofía de la ciencia en el futuro. Se trata de una propuesta alternativa. Con ello quiero decir que hay otras formas de hacer filosofía de la ciencia, formas con una larga y densa tradición, que son incompatibles o, por lo menos, ajenas al estilo que se propone a continuación. Ello no implica que este programa pretenda excluir las demás formas. Creo que es bueno que haya varios programas alternativos en filosofía de la ciencia. A lo único a lo que aspira el presente es a su derecho de existencia. Por lo demás, como es natural, el programa propuesto intenta atraer al máximo número posible de practicantes. Por eso se presenta a sí mismo como una forma de hacer filosofía de la ciencia más fructífera, más flexible y más realista que formas alternativas. Dejo a aquellos de mis colegas que estén predispuestos en contra suya la tarea de mostrar que no es así. Es previsible que las propuestas que siguen choquen con ciertas preconcepciones más o menos generalizadas de lo que es o debe ser la filosofía de la ciencia. Sin embargo, también creo que es justo concederles a los partidarios del presente programa, a pesar de todas las críticas que se les pueda hacer, un cierto margen de tiempo para que lo implementen: «Por sus frutos los conoceréis.»

La motivación para escribir lo que sigue tiene un doble origen: es en parte «institucional», en parte «ideológica». La motivación institucional es la siguiente. En varios países de habla hispana se desarrollan

125679

actualmente, o van a desarrollarse pronto, proyectos para el estudio avanzado y para la investigación sistemática en la filosofía de la ciencia a nivel de diversas instituciones universitarias o parauniversitarias. Por tanto, parece el momento adecuado para presentar ante el público universitario una visión de cómo uno cree que hay que realizar esa tarea. En otras palabras, para los filósofos de la ciencia hispanohablantes es el momento de «poner las cartas sobre la mesa».

En cuanto a la motivación «ideológica», también parece éste ser un buen momento para un «arreglo de cuentas» con la filosofía de la ciencia tal como se ha venido practicando en los últimos tiempos. Esta se halla actualmente en un estado de flujo, que en algunos casos raya con el caos. Hace veinte o treinta años estaba bastante claro (al menos en los países anglosajones) lo que era hacer filosofía de la ciencia, cuáles eran sus métodos, los problemas básicos y las soluciones que se podían esperar. En los años recientes ciertamente ha crecido el número de profesionales dedicados a la filosofía de la ciencia, el número de publicaciones, congresos y coloquios internacionales, el número de Universidades y de países donde se estudia sistemáticamente; pero la contrapartida ha sido un sentimiento creciente de desorientación y confusión. En tal situación no se trata de volver a alguna especie de monolitismo escolástico; de lo que se trata es de reflexionar pausadamente sobre lo que uno hace y/o pretender hacer y, si es posible, tomar alguna decisión sobre el camino a seguir, con relativa independencia del camino que decidan tomar los demás (si es que deciden tomar alguno y no dejarse llevar por las olas de la moda). Dicho brevemente, en tal situación es conveniente un poco de reflexión metafilosófica. Las páginas que siguen contienen algunas reflexiones de este tipo.

Antes de empezar con el tema propiamente dicho, quisiera hacer una advertencia terminológica. «Filosofía de la ciencia» es el nombre de una disciplina académica existente a la que también podría llamarse, y se la ha llamado, de otras maneras. En los países de habla alemana se la llama «Teoría de la ciencia» («Wisseschaftstheorie»), aunque muy recientemente empieza a divulgarse, por influencia del inglés, el término «Wissenschaftsphilosophie». En Francia se usa indistintamente «Epistémologie» y «Philosophie des Sciences» (en plural). En España y algunos países hispanoamericanos el nombre oficial es «Metodología», aunque está cayendo en desuso. Todos estos calificativos tienen sus ventajas e inconvenientes. Pero «Filosofía de la ciencia» sigue siendo el más divulgado, tanto en español como en inglés. Quizás más adecuados que todos los términos anteriores (por razones de contenido que veremos inmediatamente) serían «Metaciencia», «Metateoría de la ciencia» o incluso «Cientología». Pero dado que éstos aún suenan muy extraños, cuando no esotéricos, he preferido dejar las cosas como están y usar el consabido «Filosofía de

la ciencia». En definitiva, el nombre es lo de menos. Pasemos a la substancia.

1. Permítaseme empezar con una atrevida definición. Quisiera caracterizar esa disciplina que llamamos filosofía de la ciencia como una *teorización sobre teorizaciones.* Y quisiera que se tomara esta caracterización como equivalente a esta otra: como una *interpretación de interpretaciones* de la realidad; o bien, con un poco más de precisión, como la construcción de esquemas interpretativos de carácter filosófico con el fin de entender esos esquemas interpretativos de la realidad que llamamos «teorías científicas». Naturalmente, esta caracterización de la filosofía de la ciencia como teorización sobre teorizaciones o, alternativamente, como interpretación de interpretaciones, todavía es muy pobre en contenido, casi vacía, puesto que no he definido mis conceptos de teorización e interpretación. Trataré de estas nociones en seguida. Pero antes de ello quisiera observar que la caracterización propuesta, a pesar de su aparente vaciedad, contiene ya una sugerencia acerca de lo que la filosofía de la ciencia puede hacer y no puede hacer. En efecto, si aceptamos los términos de la caracterización anterior, se desprende que la filosofía de la ciencia es una actividad intelectual de *segundo nivel*, es decir, una actividad intelectual cuyos objetos son ellos mismos ya el resultado de una actividad intelectual previa. La filosofía de la ciencia es un típico producto de la *capacidad recursiva* de los seres humanos. El tipo de actividad que se aplica recursivamente en este caso es lo que llamo teorización o, alternativamente, interpretación.

Claro que como no he explicado qué entiendo por teorización ni por interpretación, lo único que he dicho hasta ahora es que la filosofía de la ciencia es una actividad intelectual del tipo x que se aplica a otras actividades intelectuales del mismo tipo x. Ahora se trata de acotar esta x, es decir, las nociones de teorización e interpretación.

Bien, debo confesar de buenas a primeras, y a despecho de provocar la decepción del lector, que no tengo ninguna definición formal y tajante de dichas nociones que ofrecer. Sin embargo, lo que sí puedo ofrecer es una aproximación intuitiva a ellas.

En primer lugar, entiendo por teorización la producción deliberada de ciertas entidades que llamamos «teorías». Más abajo trataré con detalle del concepto de teoría en general. Ahora sólo quisiera añadir que entiendo por teorías esquemas conceptuales interpretativos que, supuestamente, nos permiten a los seres humanos entender las cosas que «ocurren por ahí». Por lo tanto, caracterizo la noción de teorización en función de la de interpretación –con lo cual hemos reducido un poco el problema definitorio: ya sólo nos queda por elucidar el concepto de interpretación–. Así, pues, en nuestro intento

de caracterizar la naturaleza de la filosofía de la ciencia, el concepto de interpretación resulta clave.

2. Ahora tendría que decir lo que entiendo por «interpretación». Bien, no quisiera usar este término en un sentido filosófico técnico. No soy especialista en hermenéutica –la disciplina que, supuestamente, se ocupa de la noción de interpretación general–. Mi noción de interpretación no pretende ser ni profunda ni refinada. No es ni «germánica» ni «gálica»; ni tampoco, creo, «anglosajona», pues, que yo sepa, los filósofos anglosajones no se han ocupado mucho de esta cuestión. La noción de interpretación introducida aquí ha surgido, de manera natural y paulatina, de la confrontación de mi actividad (y la de algunos de mis colegas) como filósofo(s) de la ciencia con un dogma, muchas veces implícito, pero en todo caso ampliamente difundido, de la filosofía del lenguaje: lo que podemos llamar el dogma de la dicotomía *descriptivo-normativa* o (equivalentemente, para nuestros propósitos) *descriptivo-evaluativa*. De acuerdo con esta dicotomía, cualquier concepto, enunciado o sistema de enunciados que tenga sentido es o bien descriptivo o bien normativo (o, equivalentemente, prescriptivo-evaluativo), y ambas categorías se supone que son mutuamente excluyentes.

Al asumir esta dicotomía como punto de partida indiscutible, muchos filósofos de la ciencia que han reflexionado sobre su propia tarea se han preguntado si el estudio filosófico de la ciencia debería considerarse como un conjunto de descripciones «neutrales» del contenido de la ciencia o más bien como un conjunto de normas a priori que los científicos deben seguir *ad maiorem gloriam scientiae*. Algunos, como Nagel, Harré y muchos «historicistas», han tendido a favorecer el punto de vista «descriptivista»; otros, como los popperianos y los constructivistas alemanes han tomado una actitud claramente normativista; otros, en fin, como Sneed, se han preguntado si no sería posible combinar los componentes descriptivos y normativos en una especie de *coktail* a diversas proporciones, según el gusto de cada cual. Todos ellos, a pesar de sus diferencias mutuas, han asumido como principio tácito que todo lo que uno, como filósofo, puede hacer con respecto a la ciencia es o bien describirla o bien darle normas de procedimiento (¡la famosa «metodología»!).

Creo que esta dicotomía es engañosa y que ha sido perjudicial para la recta comprensión de muchas actividades intelectuales, entre ellas no sólo la filosofía de la ciencia, sino la ciencia misma. Las categorías «descriptivo» y «normativo» (o «prescriptivo») se pueden aplicar con sentido a muchas parcelas del discurso humano, pero a otras sólo de una manera *ad hoc*, inverosímil. En particular, a las formas de discurso que llamamos «teóricas» no se les puede aplicar propiamente. Por ejemplo, puedo describir lo que veo en el cielo de

noche diciendo que *hay* un gran número de puntos luminosos en diferentes posiciones y con diversos grados de luminosidad, y que algunos de ellos cambian de posición en noches sucesivas de manera más regular que otros. Esto es una descripción. (En rigor, ni siquiera esto es una descripción «pura», pero *pace* Hanson se aproxima mucho al ideal de una descripción pura.) Sin embargo, cuando digo que algunos de esos puntos luminosos *influyen* en nuestro destino y carácter, o que *cumplen* la ley de la gravitación universal, estoy utilizando una forma de discurso que pertenece a otra categoría que la anterior. Ya no se trata de descripciones, en ningún sentido interesante y plausible de «descripción». Es forzado tratar de meter esas formas de discurso en el mismo saco que la primera. Y no veo ninguna razón, más que un prejuicio, para persistir en ese empeño. Creo que es mucho más natural y plausible decir en estos casos que estamos no describiendo, sino «interpretando» nuestros datos. El resultado de tal interpretación es una *teoría;* en nuestro ejemplo, una teoría astrológica y una teoría mecánica, respectivamente. Por otro lado, también puedo hacer otro tipo de cosas en mi discurso. Por ejemplo, puedo redactar un código que prescriba cómo los ciudadanos de un país deben comportarse ante determinadas situaciones. Esto es un discurso normativo. Pero también puedo considerar la conducta de esa gente como el resultado de ciertas constricciones institucionales, sociales y económicas. En tal caso ya no estoy estableciendo normas, sino interpretando su conducta dentro del marco de alguna teoría (por ejemplo, sociológica o económica).

En vez de tratar de aplicar forzadamente la mencionada dicotomía a todos los casos que se nos presenten, creo que es mucho más adecuado admitir que existe por lo menos una tercera categoría semántica, además de descripción y prescripción, a la que podemos llamar «interpretación» y que no puede reducirse ni a descripciones ni a prescripciones. Una tarea importante a realizar por la filosofía analítica del lenguaje sería la de proporcionar una elucidación formal detallada de esta «tercera» categoría y de las reglas semánticas que la rigen. No he visto nada por el estilo en la literatura de la filosofía analítica. En todo caso, creo que habría que trabajar más en esa dirección. Ahora bien, de momento al menos no seré yo quien se ocupe de esta cuestión con detalle, pues mis intereses inmediatos son otros. Dejo la tarea de elucidar el concepto de interpretación a filósofos del lenguaje mejor pertrechados. Por mi parte, y aunque sea de forma provisional, me contento con la idea intuitiva, extraída de ejemplos particulares de actividades de filósofos y científicos, de que, además de descripciones y prescripciones, hay otra cosa que podemos llamar interpretaciones. De manera muy tentativa podría «definir» una interpretación de cierto dominio de objetos de conocimiento como la «incrustación» en este dominio, de modo consciente y deliberado, de un aparato conceptual elaborado previamente que nos

permite «reconstruir» ese dominio (mejor dicho, ciertos aspectos del mismo). Soy perfectamente consciente de que tal «definición» aún deja mucho que desear en cuanto a claridad y precisión; pero por lo menos creo que es suficiente para mostrar que aquí no nos sirve la semántica usual de las descripciones y las prescripciones. Y por supuesto la presente concepción de una interpretación no pretende ser enteramente original. El lector habrá oído ya en ella un eco kantiano. No me molesta este eco. Probablemente, con su concepto de síntesis y su teoría de las categorías, Kant barruntó algo parecido. Pero ahora no quisiera perderme en el laberinto de una discusión de la epistemología kantiana. Para los propósitos subsiguientes basta con que aceptemos esa noción intuitiva de interpretación y la idea de que una teoría es precisamente el resultado de una tal incrustación de un aparato conceptual sobre un dominio.

3. Hay muchas clases de teorías acerca de lo dado en la experiencia humana. Hay teorías mitológicas, metafísicas, ideológicas. También hay teorías que llamamos «científicas». No quisiera extenderme mucho aquí sobre la pregunta insidiosa de cómo podemos diferenciar las teorías científicas de las no-científicas. Y ello por dos razones: una, porque no creo que se pueda dar una respuesta general, a priori y tajante a dicha pregunta, sino que en todo caso se le podrán dar ciertas respuestas mínimamente satisfactorias una vez se haya reunido mucho más material en el análisis de todo tipo de teorías de lo que es el caso hoy día; otra, que no considero que esa pregunta sea tan central para la filosofía de la ciencia como se suele creer. A mi entender, el famoso «problema de la demarcación» entre ciencia y no-ciencia, que tanto apasionó a los filósofos «clásicos» de la ciencia, sólo es relevante para una filosofía de la ciencia de tipo normativo. Es evidente que hay diferencias notables, pongamos por caso, entre la cosmogonía bíblica y la teoría de la evolución de Darwin. Pero no es menos evidente que también hay diferencias notables entre la cosmogonía bíblica y la budista, o bien entre la teoría de la evolución y la mecánica cuántica. Y tratar de concebir un criterio según el cual unas diferencias sean «más esenciales» que otras según algún prejuicio favorito me parece poco fructífero y, en todo caso, demasiado prematuro: la verdad es que aún sabemos muy poco de la estructura y función de las diversas clases de teorías.

De momento, lo único que hemos de tener en cuenta para poder *hablar sobre* la ciencia es el hecho *institucional* indubitable de que hay una serie de teorías que se aprenden, propagan y usan en instituciones universitarias o parauniversitarias y en centros de investigación. Un grupo importante de estas teorías, al menos las que se aprenden, propagan y usan en las Facultades, Institutos y Centros llamados de «Ciencias Naturales», y además algunas de las que pertenecen a otras

áreas (como parte de los Centros de Economía, Psicología y Lingüística), tienen cierto innegable «aire de familia» entre sí. Este aire de familia no tiene por qué ser de momento más que una intuición vaga y provisional. Pero es enteramente suficiente para nuestros propósitos: podemos convenir en llamar a este importante grupo más o menos conexo de teorías «teorías científicas» y concentrarnos entonces en su estudio. Pues bien, la filosofía de la ciencia consiste en teorizaciones «del segundo nivel» cuyo objeto de estudio (de interpretación) son las teorías de ese grupo seleccionado.

4. Las teorías científicas son un subconjunto propio del conjunto global de las teorías sobre la realidad. Y hemos dicho que las teorías son el producto de la actividad discursiva que llamamos interpretación. Ahora bien, las interpretaciones sobre la realidad son fenómenos culturales (son actividades humanas que se dan dentro de ciertas formas socio-culturales). Por tanto, las teorías científicas son productos culturales. Por tanto, su estudio, que llamamos filosofía de la ciencia, forma parte (una parte importante, dada la importancia del fenómeno «ciencia» para nuestra sociedad) de las llamadas «ciencias de la cultura» o, como también se las llama, «Humanidades». Aunque las divisiones académico-administrativas no deben tomarse demasiado en serio, pueden tener algún fundamento en su objeto de estudio. En la medida en que se crea conveniente aceptar la división entre las tres grandes áreas disciplinarias, «Humanidades», «Ciencias», «Tecnología», está claro que debemos adscribir la filosofía de la ciencia a las «Humanidades», puesto que estudia productos culturales (humanos). Este punto me parece tan obvio que casi consideraría ridículo mencionarlo si no fuera porque casi nunca se hace, y porque por experiencia sé que algunos de mis colegas reaccionan con cierto asombro y hasta repudio cuando se les dice que el lugar propio de la filosofía de la ciencia está entre las ciencias de la cultura; que, por su objeto de estudio, la filosofía de la ciencia está más emparentada con disciplinas como la Historia del Arte, la Crítica Literaria y la Antropología Cultural que con la Física, la Biología y la Psicología, pongamos por caso.

Creo que puede haber tres motivos principales por los que esta clasificación de la filosofía de la ciencia entre las ciencias de la cultura puede causar cierta extrañeza cuando no rechazo. Los tres motivos se basan a su vez en tres tipos distintos de razonamientos. Los tres tienen una parte de fundamento, pero también en gran parte son infundados, y en cualquier caso están embrollados por confusiones metodológicas.

Primer razonamiento: la filosofía de la ciencia no pertenece al ámbito general de las ciencias de la cultura, sino a cada disciplina en particular. La filosofía de la física pertenece a la física, la filosofía de la

biología pertenece a la biología, etc. O, resumido en un *slogan* que frecuentemente se oye: «La filosofía de la ciencia deben hacerla los propios científicos.» (Por «científicos» se suelen entender aquí los «científicos naturales».) La confusión conceptual que está en la base de este *slogan* es, naturalmente, la confusión, o no-diferenciación, de los dos niveles conceptuales en que puede moverse una actividad interpretativa: interpretación directa de los fenómenos (naturales o sociales) frente a interpretación de las interpretaciones. Tomado literalmente, este razonamiento es simplemente grotesco: es como decir que la gramática latina sólo la pueden analizar los romanos, o que sólo un papú puede estudiar la religión de los papúes. El pequeño fundamento real que tiene esta actitud es el siguiente. Muchos científicos temen que el estudio filosófico de su ciencia degenere, por ignorancia, en una especie de «ciencia-ficción», en una fantasía sobre la verdadera naturaleza de esa ciencia. Este peligro es real (aunque se ha exagerado más de la cuenta). La manera de obviarlo, por supuesto, es hacer un estudio concienzudo de las teorías científicas en cuestión. Ningún filósofo serio pondrá esto en duda. Pero no hay que olvidar que analizar concienzudamente una teoría científica no es lo mismo que practicarla o aplicarla. No es necesario (y seguramente ni siquiera conveniente) hacer lo segundo para poder hacer lo primero, del mismo modo como no es necesario, ni siquiera conveniente, comulgar cada domingo para analizar la estructura del dogma católico. El hecho de ser practicante de la ciencia (o religión, o lo que sea) en cuestión no implica que uno esté en mejor posición para poder interpretarla filosóficamente. Muchos científicos, cuando se han puesto a filosofar sobre su propia ciencia, han producido gran cantidad de banalidades, simplificaciones y hasta barbaridades. (Basta leer algunas de las cosas que han escrito, por ejemplo, Heisenberg sobre «su» física, Monod sobre «su» biología y Skinner sobre «su» psicología.)

Segundo razonamiento: asimilar la filosofía de la ciencia a las ciencias de la cultura entraña el peligro del sociologismo, es decir, de la reducción del análisis de la ciencia a la consideración de las instituciones y estructuras sociales que la posibilitan, de su función social, ignorando sistemáticamente los problemas de contenido y estructura interna de las teorías científicas. Este peligro es real, muy real, en verdad. Resulta tangible en los voluminosos acopios de trivialidades, vaguedades y digresiones sin fundamento empírico que han producido muchos de los «sociólogos del conocimiento» y de los epistemólogos marxistas, entre otros. Más recientemente hemos visto cómo incluso un autor tan perspicaz como Kuhn no ha logrado evitar siempre el canto de sirenas del sociologismo. Ahora bien, el remedio a esta situación no consiste en arrojar la filosofía de la ciencia fuera del dominio de las ciencias de la cultura, al que naturalmente pertenece, sino en tener un concepto más amplio y articulado de estas últimas, que no implique, entre otras cosas, la identificación de la filosofía de la

ciencia con la sociología de la ciencia. Sobre este punto volveré más adelante.

Tercer razonamiento: la filosofía de la ciencia, si se practica «en serio», debe usar las herramientas precisas de las matemáticas y las ciencias naturales en vez de las divagaciones puramente verbales e incontrolables típicas de las Humanidades. Esta actitud, que se suele describir como «cientificismo», traduce un menosprecio por las ciencias culturales, a las que se ve como un caos conceptual irremediable, y el intento de «salvar» a la filosofía de la ciencia de ese caos. Concuerdo con esta actitud en el punto de que la filosofía de la ciencia debe ser lo más rigurosa posible, y para ello debe hacer el máximo uso *plausible* de los instrumentos formales de la lógica y las matemáticas. El odio visceral de algunos de los actuales filósofos de la ciencia hacia todo instrumento formal de análisis conceptual puede tener su origen en problemas psicológicos del autor en cuestión o en modas académicas, pero en cualquier caso no tiene mayor justificación que su contrapartida, el exceso de formalismos irrelevantes que se dio hace un par de décadas.

El rechazo de toda formalización en el estudio de la ciencia tiene su origen en los ataques sistemáticos al positivismo lógico que se generalizaron a principios de los años sesenta. Sea cual sea la valoración última que se haga de tales ataques, lo cierto es que muchos de esos autores críticos la emprendieron contra el positivismo lógico no sólo por «positivista», sino también por «lógico», es decir, por el uso que hacía de las herramientas de la lógica y las matemáticas. Con ello tales autores, evidentemente, confundieron «la gimnasia con la magnesia». Pues lógica y matemáticas no son monopolio de ninguna concepción filosófica particular, sino el mejor instrumento del que disponemos para hablar con claridad y precisión sobre las cosas, especialmente las «cosas» conceptuales. La lección de rigor que nos han dado los grandes clásicos de la filosofía de la ciencia (Russell. Reichenbach, Carnap y tantos otros) al usar métodos formales de reconstrucción, no debería ser olvidada por las nuevas generaciones de filósofos de la ciencia. Quisiera subrayar en este punto que, aunque ya no concordemos con su visión filosófica particular de la ciencia, con sus supuestos y conclusiones, considero inexcusable el estudio de esos «clásicos» como entreno intelectual para formular teorías filosóficas de la ciencia que no desemboquen en una «noche donde todos los gatos son pardos» (como es mucho de lo que se produce actualmente en nuestra disciplina). Ellos nos mostraron con el ejemplo la gran cantidad de cosas interesantes que se pueden hacer con instrumentos lógico-matemáticos, y no hay que echar por la borda ese útil arsenal simplemente porque ya no nos satisfaga su concepción empirista o positivista de la ciencia.

Ahora bien, este alegato en favor del uso de instrumentos formales no implica que haya que evitar el «contagio» de la filosofía de la

ciencia con otras áreas de las ciencias de la cultura. El *desideratum* sería hacer a *todas* las interpretaciones de segundo nivel lo más rigurosas posible, y si se hallan en situación de usar construcciones formales, tanto mejor. De hecho, una porción aún minoritaria, pero cada vez más importante, de las ciencias humanas ha emprendido ya este camino. En el estudio del lenguaje y de las formas de organización económica y social se hace un uso cada vez más frecuente de modelos lógicos y matemáticos, más refinados, en realidad, que los usados por la filosofía estándar de la ciencia.

5. No veo, pues, ningún peligro fundado en asimilar la filosofía de la ciencia a las ciencias de la cultura, siempre que esté claro (como debería estar) lo que significa y lo que no significa tal asimilación. En este punto es importante hacer notar lo siguiente. La interpretación filosófica de las teorías científicas evidentemente no es el único modo posible de teorizar sobre ese fenómeno que llamamos ciencia. En vez de interpretar las teorías científicas *qua* teorías (que es la misión de la filosofía de la ciencia tal como la entiendo) podríamos dedicarnos al estudio de las instituciones sociales o de las personas que las producen. En tal caso, nos dedicaríamos a otra clase de estudio teórico de la ciencia; por ejemplo, a sociología de la ciencia o psicología de la ciencia. Espero que se interprete bien esta diferenciación. No estoy tratando de negar o negligir las múltiples, importantes y fructíferas conexiones entre la filosofía de la ciencia por una parte y la sociología o la psicología de la ciencia por otra. Considero incuestionable la existencia de esas conexiones y creo que este punto hoy día es ya un lugar común. Sin embargo, también debemos ser plenamente conscientes de que la filosofía de la ciencia y los estudios sociológicos de la ciencia no pueden identificarse. Sean cuales sean las importantes conexiones entre ambas disciplinas, sus objetos de estudio son esencialmente distintos. Los estudios sociológicos de la ciencia tratan de entidades localizadas espaciotemporalmente, tales como personas o instituciones académicas; la filosofía de la ciencia trata de *teorías*, y éstas son entidades abstractas en el sentido de no estar localizadas espaciotemporalmente, al menos en el mismo sentido en que lo están organismos o entidades sociales. (A lo sumo, podemos decir que están localizadas en un sentido «vicario» respecto de organismos y objetos sociales.) El filósofo de la ciencia estudia las producciones científicas de un modo similar a como el matemático estudia números, figuras geométricas y otras entidades matemáticas, y de un modo más similar aún a como el *meta*matemático estudia los sistemas formales de la matemática pura: a saber, como *estructuras abstractas*. Por esta razón, la filosofía de la ciencia posee ciertas semejanzas de familia con la matemática pura y con la metamatemática (aunque tampoco se puede identificar con ellas, claro).

He dicho más arriba que las teorías científicas son objetos culturales producidos por cierta gente en determinados contextos socioculturales y que, por esta razón, el estudio filosófico de las teorías científicas, la filosofía de la ciencia, debe incluirse en el dominio de las ciencias de la cultura. Puede parecer que esto se halla en contradicción directa con el carácter abstracto, casi platónico, que acabo de postular para la filosofía de la ciencia. Sin embargo, no creo que haya aquí ninguna contradicción real. Es una visión demasiado estrecha del contenido de las ciencias de la cultura considerar que *sólo* se ocupan de entidades localizadas espaciotemporalmente. Las teorías científicas son precisamente un ejemplo de entidades *culturales abstractas* (en el sentido de no-localizadas).

6. Por otro lado, existe otro sentido de «abstracto» en el que la filosofía de la ciencia no es (o no debe ser) abstracta. El hecho de que el objeto de estudio de la filosofía de la ciencia sea abstracto en el sentido que acabamos de elucidar no implica que el *método* usado por el filósofo en su estudio deba ser abstracto en el sentido de apriorístico, sin tener en cuenta los datos concretos que nos proporciona el *corpus* de las ciencias. Por el contrario, quisiera subrayar, incluso bajo riesgo de sonar nuevamente contradictorio, que el *procedimiento* del filósofo de la ciencia debería ser concreto en el sentido de enfrentarse a teorías científicas realmente existentes que se puedan identificar concretamente. El filósofo de la ciencia no puede construir una metateoría general de la ciencia a menos que base su intento en el análisis de ejemplos concretos de teorías científicas. Para evitar la tentación de desarrollar una especie de filosofía de la ciencia-ficción, debe tomar ejemplos concretos de teorías como sus puntos de referencia. De lo contrario, la filosofía de la ciencia acabaría siendo, en el mejor de los casos, una empresa puramente normativa, en el peor, una novela. Muchos de los fallos de la concepción clásica o estándar de la filosofía de la ciencia se originaron en el hecho de que, con demasiada frecuencia, procedió en términos demasiado generales y apriorísticos, con lo que resultó más prescriptiva que interpretativa.

La exigencia de que la filosofía de la ciencia proceda más concretamente en sus análisis significa esencialmente que las representaciones existentes de teorías científicas particulares deben tomarse muy en serio. Ahora bien, la cuestión es: ¿A dónde hemos de dirigir nuestra mirada para encontrar esas representaciones concretas? Bien, la respuesta es: principalmente en libros y artículos, del presente lo mismo que del pasado. Los *libros de texto* juegan un papel especialmente importante en este respecto. El filósofo de la ciencia desprejuiciado, no-normativo, que desea entender los aspectos esenciales de las disciplinas científicas existentes, se ve llevado naturalmente a un análisis cuidadoso y detallado y a una reconstrucción de los libros de texto correspondientes.

125679

Debo hacer notar aquí que uso el término «libro de texto» en un sentido amplio, que no solamente cubre textos elementales de los primeros cursos de una disciplina, sino también obras avanzadas o históricamente famosas, en cualquier caso sistemáticas. Como ilustración del presente sentido de «libro de texto» se podrían mencionar, para el caso de la filosofía de la física, obras tales como la *Mecánica analítica* de Lagrange, los *Principios de la Mecánica Cuántica* de Dirac y la *Termodinámica* de Callen.

Contrariamente a una opinión muy extendida, creo que los libros de texto son esenciales a los objetivos del filósofo de la ciencia –no porque opine que contienen profundas ideas filosóficas, sino simplemente porque en general son la expresión más adecuada de una teoría dada en su actual estadio evolutivo–. Tan pronto como las teorías han alcanzado un cierto grado de madurez y sistematicidad, normalmente resultan estandardizadas en un libro de texto y transmitidas a la totalidad de la comunidad científica bajo esa forma. Las generaciones más jóvenes de científicos se familiarizan primero con los conceptos, principios e instrumentos de la teoría en cuestión a través de los libros de texto correspondientes. Los libros de texto son los objetos concretos que mejor transmiten las entidades abstractas que llamamos teorías. Es por esta razón que una tarea importante del filósofo de la ciencia consiste en analizar críticamente y reconstruir lógicamente las obras estándar de las disciplinas científicas: ésta es la mejor manera de aproximarse a su meta de identificar la estructura de teorías científicas particulares.

Soy consciente de que esta evaluación positiva de la importancia de los libros de texto *no* es compartida por muchos filósofos de la ciencia, quienes o bien proceden normativamente presuponiendo alguna concepción a priori de la ciencia, o bien, en el caso de que estudien piezas concretas de la ciencia, tienden a sentirse más atraídos por lo que en otro lugar[3] he denominado «ciencia heroica»: los grandes cambios y las grandes innovaciones en la construcción de teorías. Sin duda, la ciencia heroica es muy interesante y hermosa (por mi parte he tratado de contribuir a su estudio en el caso de la física newtoniana)[4], pero también creo que la filosofía contemporánea de la ciencia ha exagerado su importancia en detrimento de los aspectos más cotidianos de la ciencia. Las representaciones de teorías científicas que se han sistematizado en los libros de texto son mucho más típicas de la estructura real de la ciencia.

7. Una objeción que se suele hacer al énfasis puesto en los libros de texto como objetos de reconstrucción es que éstos sólo contienen

[3] C. U. Moulines, *Reply to John North's «On Marking History»*.

[4] Cf. los capítulos 3.1 y 3.2 de este libro.

las formas «disecadas» de las teorías científicas, no las formas «vivas» de la ciencia, sobre todo las más recientes. Se dice que la ciencia es un «proceso en flujo» que no se puede «congelar» en los libros de texto.

En este tipo de rechazo de los libros de texto hay una gran dosis de retórica romántica, retórica conectada naturalmente con la imagen de la «ciencia heroica» y de los «grandes genios revolucionarios», tan favorecida por muchos filósofos. Nadie pone en duda (por lo menos yo no) que la ciencia es un «proceso en flujo» –o, mejor dicho, múltiples procesos en flujo–. Y el mejor testimonio de ello son precisamente los libros de texto. Nada más revelador para captar el sentido de la evolución de la mecánica newtoniana durante el primer siglo de su existencia que comparar sucesivamente cuatro libros de texto: los *Principia* de Newton, la *Mechanica* de Euler, la *Mécanique Analytique* de Lagrange y, finalmente, la *Mécanique Céleste* de Laplace. Cierto que estos textos apenas contienen alguna mención de anécdotas sobre la personalidad de sus autores, sobre las noches de insomnio que pasaron antes de llegar a sus resultados, sobre sus amistades o enemistades, sobre sus prejuicios políticos y religiosos. Pero, ¿es esto realmente tan esencial para entender la naturaleza y evolución de la mecánica clásica?

El menosprecio por las exposiciones estándar de las teorías suele ir acompañado del menosprecio por el estudio de las formas clásicas de una disciplina; por ejemplo, en el caso de la filosofía de la física es frecuente oír el reproche de que los filósofos han tomado demasiada inspiración de la física clásica en vez de concentrarse en aquellas áreas de la física «que verdaderamente importan hoy día»: la relatividad generalizada y la mecánica cuántica. En realidad, el hecho histórico es que ha ocurrido todo lo contrario: los filósofos de la ciencia del siglo XX se han visto mucho más impresionados por la física relativista y cuántica que por la física clásica, y han dedicado desproporcionadamente sus esfuerzos a los fundamentos de las primeras en detrimento de la segunda. (Basta observar cuánta inspiración sacaron de la física relativista para sus concepciones filosóficas generales autores tan influyentes como Russell, Whitehead, Reichenbach, Popper y Carnap.)

Por mi parte creo que lo que hace falta, aunque sólo sea para restablecer el equilibrio, es ocuparse más de las formas clásicas de la física. A veces se replica a esto que las teorías de la física clásica carecen de verdadero interés para una filosofía de la ciencia que quiera «estar al día» porque son teorías «ya superadas», «muertas». Lo cual es un doble absurdo. En primer lugar, es grotesco decir que la física clásica está «muerta» cuando sigue aprendiéndose, propagándose y usándose en todas las Universidades del mundo, por «avanzadas» que éstas sean. Teorías como las de la mecánica de Newton, la electrodinámica de Maxwell o la termodinámica de Gibbs siguen formando parte esencial del bagaje del físico moderno. Pero en un sentido más importante para nosotros, esa objeción es absurda

porque confunde la misión de la filosofía de la ciencia con la de una divulgación o popularización de los «progresos recientes» de la ciencia, con lo cual, naturalmente, nada tiene que ver. Valoraciones de mayor o menor actualidad deben ser ajenas a la filosofía de la ciencia. Para un filósofo desprejuiciado en principio da igual si una teoría científica fue diseñada en la Italia del Renacimiento hace quinientos años o bien es la última hipótesis propuesta por un grupo de jóvenes físicos en Berkeley. Lo importante es que pueda identificarse su estructura conceptual y reconstruirla. Decirle a un filósofo que las teorías de la ciencia clásica no le deben interesar porque «ya están muertas» es como decirle a un musicólogo que debe abandonar su análisis de composiciones tan anticuadas como los *Conciertos de Brandeburgo* o bien recomendarle a un crítico literario que en vez de dedicarse a una obra «tan superada» como el *Quijote* se dedique a la última novela de García Márquez. (Estas cosas a veces se dicen, por cierto, pero no creo que ningún crítico literario o musicólogo serio se deje impresionar por ellas.)

8. La analogía con la actividad del crítico literario, del musicólogo o, en general, de cualquier analista de objetos culturales del pasado o del presente nos puede servir para aclarar otro punto respecto a la metodología del filósofo de la ciencia.

La propuesta de concentrarse principalmente en el análisis de los libros de texto u otras representaciones estándar de las teorías científicas *no* debería confundirse con una recomendación en el sentido de limitarse a una *descripción* acrítica del contenido de las obras en cuestión. Ya he señalado que el objetivo del filósofo de la ciencia no es describir, sino interpretar o reconstruir productos científicos. En este contexto puede ayudarnos establecer una comparación entre el objetivo del filósofo de la ciencia y la tarea usual de un crítico literario moderno. Pienso en la clase de análisis literarios que usan categorías formales generales y que son típicos, por ejemplo, de la escuela estructuralista. (No es relevante ahora que uno concuerde o no con las conclusiones de esa escuela: lo único que interesa para la discusión presente es la metodología que usan.) El crítico literario moderno no concibe su propia tarea como una mera descripción del contenido de una obra literaria ni como una exégesis de la misma, sino más bien como un análisis interpretativo que le permite revelar las estructuras implícitas, subyacentes, de las cuales el autor mismo es inconsciente en la mayoría de los casos. Para lograr esto, el crítico literario tiene que subsumir esa obra como objeto cultural bajo ciertas categorías formales presupuestas. De una manera similar, la tarea del filósofo de la ciencia consiste en reconstruir, es decir, interpretar las obras científicas para poner al descubierto e identificar las estructuras abstractas subyacentes que llamamos teóricas.

9. Hasta aquí, los términos de nuestra discusión se han mantenido dentro de un marco muy general. Quisiera exponer ahora más concretamente cómo aparecería la implementación del programa de reconstrucción lógica de la ciencia que tengo en mente.

He subrayado ya que tomo las teorías científicas por el objeto propio de investigación para la filosofía de la ciencia. Sin embargo, esta determinación ciertamente tiene que ser matizada; de lo contrario conduciría a una visión supersimplificada de la tarea de la filosofía de la ciencia. Las cualificaciones necesarias discurren en dos direcciones. Primera, tenemos que ser más explícitos acerca de la estructura de las teorías individuales que tomemos como unidades de nuestra investigación. Segunda, debemos tener en cuenta el hecho de que las teorías, aunque unidades básicas, *no* son la única clase de entidades con que nos enfrentamos en un análisis detallado del *corpus científico*. Las teorías científicas no se dan, por así decir, como unidades aisladas: están esencialmente interrelacionadas entre sí. A fin de discutir este punto con algún detalle, podemos introducir una imagen gráfica, intuitiva. Podemos visualizar la totalidad del *corpus* de la ciencia como una enorme *red* de teorías. Los nudos de esta red son las teorías particulares que identificamos; sus cuerdas son las relaciones entre ellas, por ejemplo, la relación lógica de implicación deductiva.

Naturalmente, ésta es todavía una imagen muy burda de la estructura de la ciencia, y más adelante explicaré por qué creo que esta imagen es tan simplista. No obstante, como primera aproximación, creo que esta imagen de la red es útil para discutir algunos puntos. El concepto de red visualiza la aseveración de que la tarea de la filosofía de la ciencia en realidad es doble: detectar los nudos *a la vez que* las cuerdas de la red de la ciencia. No obtendríamos una visión completa de la ciencia si sólo consideráramos los nudos o las cuerdas. En primera aproximación, es recomendable metodológicamente dedicarse solamente a una de las dos tareas, ya que no podemos hacerlo todo simultáneamente. Pero debemos tener siempre bien presente que la identificación completa de una teoría no es posible a menos que sepamos qué clase de relaciones inter-teóricas ella tiene con otras teorías, y recíprocamente, no podemos elucidar tipos particulares de relaciones inter-teóricas de manera completamente adecuada si no sabemos qué estructura tienen de hecho las teorías mutuamente relacionadas.

10. De momento, sin embargo, voy a separar ambas tareas y decir algo sobre la primera tan sólo. Para identificar y reconstruir teorías científicas individuales necesitamos, ante todo, un *criterio de identidad*, y esto significa un concepto explícito general de teoría. Los diversos intentos de proporcionar un criterio satisfactorio de identidad para las teorías han influido mucho en el desarrollo de la filosofía

contemporánea de la ciencia. Consideremos brevemente la historia de este desarrollo.

El concepto de teoría de la filosofía «clásica» de la ciencia fue extraído de las matemáticas y aplicado a la totalidad de la ciencia. De acuerdo con esta elucidación clásica de la noción de teoría, ésta no es más que un *sistema axiomático*. Y un sistema axiomático es simplemente un conjunto de proposiciones básicas que se toman como a priori válidas, es decir, sin prueba, junto con sus consecuencias lógicas (sus «teoremas»). Dicho brevemente, de acuerdo con esta concepción, una teoría es simplemente un conjunto de proposiciones. El método que debería seguir el filósofo de la ciencia para identificar teorías particulares no es sino el método axiomático que ya se conoce en las matemáticas: tenemos que averiguar cuáles son los enunciados básicos de la teoría que sirvan como axiomas a partir de los cuales todo lo demás (dentro de la teoría) se puede derivar lógicamente.

La representación axiomática de una teoría estará controlada al máximo cuando la derivación de sus teoremas esté controlada al máximo. Esto ocurre cuando la teoría aparece expresada en un lenguaje formal. Los lenguajes formales de la lógica son sin duda mucho más precisos y mejor controlables que el lenguaje normal en que se expresan las teorías en los libros de texto usuales. Por lo tanto, es comprensible que muchos filósofos de la ciencia hayan considerado no sólo la axiomatización, sino también la formalización como un método esencial para su objetivo de reconstruir la ciencia. Esto también explica por qué, en los comienzos de la moderna filosofía de la ciencia, se dedicó tanto esfuerzo a la construcción de lenguajes formales adecuados. Rudolf Carnap es uno de los adalides más famosos de este tipo de enfoque. Dedicó toda su vida a la tarea de construir y perfeccionar lenguajes formales –no, como algunos críticos malévolos han sugerido, para jugar con esos formalismos, sino porque estaba firmemente convencido de que no se puede alcanzar un grado razonable de precisión en el análisis de la ciencia, a menos que se usen lenguajes formales. Para él y sus dicípulos, la formalización pasó a ser mandamiento ineludible.

Aunque la obra de Carnap es interesante y valiosa en tantos respectos, la historia reciente de la filosofía de la ciencia ha mostrado que su enfoque formalista y sintactista extremo frenó más que promovió un progreso rápido en la filosofía de la ciencia. Es un hecho que sólo muy pocas teorías fuera de la matemática pura se han podido formalizar en el sentido de Carnap (por la simple razón de que una formalización completa de cualquier teoría desarrollada es muy difícil en la práctica). Además, incluso en los casos en que se logró una formalización estricta, esto tuvo como consecuencia indeseable que muchas cuestiones interesantes, que no podían ser tratadas por medios puramente sintáctico-formales, fueron simplemente ignoradas por los filósofos.

Poco después de los primeros intentos de Carnap, Patrick Suppes y sus colaboradores iniciaron un método de reconstrucción de las teorías científicas que, aun siendo menos formalista, era igual de riguroso y más práctico. Dentro del estilo de Suppes, no es preciso construir sistemas formales previos para representar las teorías, sino que las axiomatizamos directamente dentro de la llamada teoría «ingenua» de conjuntos. La idea básica consiste en incluir todas las teorías científicas dentro del lenguaje de la teoría elemental de conjuntos –de un modo similar a la famosa reconstrucción de la matemática emprendida por el grupo Bourbaki, que sólo presupone los conceptos y procedimientos constructivos de la teoría de conjuntos. El lema del enfoque de Suppes podría ser: «Todo lo que puede decirse científicamente, puede decirse dentro de la teoría de conjuntos.» El grupo Bourbaki ha mostrado que el programa de reconstrucción conjuntista es factible y fructífero para las teorías de la matemática pura. El objetivo de Suppes y sus colaboradores es desarrollar un programa análogo para las teorías de la ciencia empírica. Este programa ya ha logrado algunos éxitos en los últimos veinticinco años. Ha sido, al menos, más fructífero que el programa carnapiano. Un buen número de teorías empíricas han sido ya reconstruidas, es decir, identificadas por medios conjuntistas a lo Suppes.

11. No obstante, aunque la labor realizada por Suppes y su grupo ha llevado a algunos resultados positivos importantes, el programa contiene aún ciertos defectos inherentes[5]. Los resultados obtenidos sólo se pueden considerar como un primer paso. La razón es que el concepto suppesiano de teoría todavía es demasiado simplista. Suppes cree que la estructura de una teoría empírica, como *teoría*, no es esencialmente diferente de la estructura de una teoría de la matemática pura. Pero eso es un error metodológico. Las teorías empíricas poseen una estructura más compleja, que una axiomatización normal no puede abarcar completamente. Este mayor grado de complejidad resulta bien claro cuando consideramos los aspectos semánticos y pragmáticos de las ciencias empíricas. Estos aspectos se hallan involucrados en lo que podemos llamar la «justificación *externa*» de las teorías empíricas, mientras que la matemática pura sólo se preocupa por la justificación *interna*, es decir, por la consistencia lógica. Creo que todos estos aspectos semánticos y pragmáticos adicionales de las teorías empíricas se pueden resumir alrededor de la noción clave de una *aplicación* de una teoría. El análisis del concepto de aplicación resulta, pues, central para la filosofía de la ciencia.

Algunos filósofos (entre ellos quizá Suppes) probablemente pien-

[5] Una exposición y crítica más detalladas del programa de Suppes se encontrarán en C. U. Moulines y J. D. Sneed, *Suppes' Philosophy of Physics*.

san que, aunque el concepto de aplicación va típicamente asociado a las teorías empíricas, no pertenece realmente a las últimas y que, por tanto, podemos separar la tarea de reconstruir teorías particulares de la tarea de aclarar la noción de aplicación. Pero no creo que esto sea metodológicamente aceptable. Las teorías empíricas están siempre construidas en vistas a algunas aplicaciones «externas» concretas, de tal modo que las aplicaciones están incluidas *también conceptualmente* en la teoría misma. No podemos realmente separar las dos cosas. Algunos autores, que de algún modo se han dado cuenta de que las teorías y sus aplicaciones están entrelazadas conceptualmente, han propuesto la famosa tesis de la «carga teórica» de todas las observaciones empíricas. Pero parece que han olvidado señalar la otra cara de la moneda, a saber, lo que podríamos llamar la «carga aplicativa» de todos los constructos teóricos.

Carnap y Hempel ya habían tratado de hacer justicia a este aspecto al introducir las llamadas «reglas de correspondencia» y los «postulados de significación». Sin embargo, su construcción de tales correspondencias es forzada y en parte ficticia. No puede decirse que proporcionaran una solución verdaderamente satisfactoria, aunque se encontraban en la dirección correcta.

En los últimos años cierto número de filósofos de la ciencia se han ocupado del problema de proporcionar un concepto adecuado de teoría que incluya la noción de aplicación de un modo realista. Los enfoques de autores tales como Sneed, Stegmüller, Ludwig, Van Fraassen y el grupo polaco, aunque proviniendo de tradiciones diferentes, todos tratan de proporcionar una solución adecuada al problema de la inclusión de las aplicaciones dentro de un concepto complejo de teoría. Aquí no puedo entrar en los detalles de estos enfoques recientes, pero quisiera señalar un rasgo básico del nuevo concepto de teoría, tal como ha surgido de los trabajos de autores tales como Sneed, Stegmüller y Ludwig (y quizá también los polacos), aunque los detalles puedan diferir grandemente entre estos autores. En ellos, una teoría ya no se concibe como un conjunto de enunciados o proposiciones, sino más bien como una estructura conceptual compleja, cuyas unidades, por así decir, son, a su vez, estructuras elementales a veces llamadas «modelos», a veces «aplicaciones» (Ludwig las llama «Bereich der Gegebenheiten» –«dominio de cosas dadas»–). Una teoría determinada no tiene un único modelo estándar en la realidad, como la concepción clásica había dado por supuesto implícita o explícitamente. Por el contrario, una teoría dada consiste en una multiplicidad abierta de modelos o aplicaciones que, por así decir, sistematizan diferentes pedazos de realidad en el marco conceptual propio de la teoría. Cada modelo o aplicación es una estructura a dos niveles en la que se distinguen dos clases de conceptos: aquellos que son específicos de la teoría en cuestión y que no tienen sentido fuera de ella, y aquellos que presuponen teorías previas y que

constituyen algo así como la base confirmatoria de la teoría en cuestión. Es importante observar, sin embargo, que esta distinción entre dos niveles conceptuales dentro de cada aplicación de la teoría dada no tiene nada que ver con la distinción clásica entre un lenguaje observacional y uno teórico, tal como aparece representada principalmente en la obra de Carnap y Hempel. La nueva distinción no es epistemológica, sino funcional, y no es absoluta, sino relativa a cada teoría.

La multiplicidad de aplicaciones que, según estas concepciones más recientes, constituyen una teoría es un conjunto abierto en el sentido de que no se puede determinar extensionalmente de una vez por todas. Depende de consideraciones pragmáticas, las cuales, lo mismo que cualquier otro aspecto pragmático de la ciencia, cambian en el transcurso del tiempo e incluso según el usuario. Dado que, por otro lado, la teoría misma está esencialmente constituida por esta multiplicidad (siendo la idea básica la de que no podemos separar completamente la teoría de sus aplicaciones), se desprende que la teoría misma es una especie de entidad *abierta*, cuya determinación conceptual no sólo debe tomar en cuenta los aspectos sintácticos y semánticos, sino también los cambiantes aspectos pragmáticos. Así, pues, una teoría resulta ser una entidad esencialmente determinada no sólo por su estructura formal y por su referencia, sino también por su *uso*. Esto hace que el criterio de identidad para teorías sea indudablemente más difícil de manejar, pero también más realista y adecuado.

12. Sin embargo, la cuestión de la identificación de las estructuras teóricas aparece ante un análisis cuidadoso aún más complicada que lo anterior. No sólo necesitamos un criterio de identidad para teorías individuales. También precisamos de criterios de identidad para *grupos* de teorías. Tenemos que ver ahora por qué la imagen original del *corpus* de las ciencias como red de teorías individuales es demasiado simple. La razón es ésta. Ni los nudos ni las cuerdas de esa red pertenecen todos al mismo nivel lógico. Existen diferentes *clases* de nudos y diferentes *clases* de cuerdas conectándolos. Por ejemplo, hay nudos menores y mayores. Hay nudos insertados en otros nudos de diversas maneras; hay «super-nudos» de nudos inter-conectados. En otras palabras, las teorías científicas tienen distintos niveles jerárquicos, y un análisis lógico cuidadoso detectaría no sólo teorías, sino estructuras más globales que las teorías propiamente dichas, que podemos llamar *familias* de teorías. Las teorías particulares no son los únicos componentes en los que podemos analizar la ciencia. La división de la ciencia en las llamadas disciplinas (física, biología, etc.), aunque metodológicamente ingenua, apunta ya al hecho de que podemos encontrar estructuras más generales que las teorías individuales. No creo que la división académica tradicional en disciplinas

sea de mucha relevancia para la filosofía de la ciencia. Pero el hecho de que estas divisiones académicas sean posibles e incluso muestren cierto grado de plausibilidad nos induce a considerar la posibilidad de que la ciencia consista en más cosas que simplemente teorías, no sólo desde un punto de vista sociológico o institucional, sino también desde el punto de vista metodológico. Las teorías individuales pueden reunirse de acuerdo a ciertos criterios lógicos, metodológicos y semánticos, constituyendo así familias de teorías. Las teorías individuales pueden agruparse por su forma y contenido según ciertas «semejanzas de familia». Trato de esta cuestión con cierto detalle en un capítulo posterior[6]. Introduzco allí el término «marco de teorías» para las estructuras globales que abarcan toda una familia de teorías y trato de argüir en favor de la posibilidad de un análisis lógico de tales estructuras tomando como ejemplo revelador el caso particular de la termodinámica. Creo que es posible identificar algo así como el «marco de teorías termodinámicas»: una estructura global común a todas las teorías de la termodinámica clásica. Aún no sabemos mucho sobre los marcos de teorías en general, pero ciertamente creo que éste es un nuevo campo prometedor de la filosofía de la ciencia.

13. La constitución de marcos de teorías a partir de teorías individuales está relacionada de forma natural con el problema de identificar las relaciones interteóricas, es decir, las cuerdas de la red científica. Estas también son más complejas y diversas de lo que la imagen primitiva de la red parecía sugerir. La filosofía estándar de la ciencia asumía implícita o explícitamente que la única relación interesante entre las teorías es la implicación lógica. Por ejemplo, el problema de la reducción de teorías, que ha jugado un papel importante en las discusiones metodológicas de las últimas décadas, se entendió generalmente como el problema de mostrar si, dado cualquier par de teorías, T_1, T_2, podemos probar que T_1 es lógicamente *deducible* a partir de T_2, o a la inversa. La reducción se identificó con la deducción. Esta es, naturalmente, una visión simplista de la reducción en las ciencia empíricas. La reducción es una relación más compleja que la deducción; se da a diferentes niveles, y es una relación entre estructuras conceptuales globales y no simplemente una relación de deducibilidad entre enunciados.

Pero además, la reducción no es la única clase de relación interteórica que podemos descubrir analizando el *corpus* científico. Quisiera mencionar, a modo de ejemplos, las relaciones de *aproximación* o de incrustación aproximativa, de *teoretización* entre teorías cuyos conceptos pertenecen a niveles metodológicos distintos y de *presuposi-*

---

[6] Cf. capítulo 2.10.

ción entre teorías con distintas prioridades epistemológicas dentro del *corpus* de la ciencia. Todas estas diferentes clases de relaciones interteóricas determinan que la red de las disciplinas científicas sea mucho más compleja de lo que se pensaba. De esto se hablará con mayor detalle en el capítulo 2.8.

14. Hay todavía otro sentido en el que la imagen anterior de la ciencia era parcial. No deberíamos visualizar la red de la ciencia como una red estática, inmóvil. Por el contrario, las disciplinas y teorías científicas, al igual que sus interconexiones, cambian continuamente; la ciencia es una red (o configuración de redes) que crece o decrece en diferentes direcciones: a veces se añaden nuevos nudos, a veces se abandonan otros; o bien se establecen nuevas conexiones, se cambian conexiones viejas, y así sucesivamente. Dicho brevemente, no podemos olvidar la *dinámica* de la ciencia si queremos obtener una imagen completa.

Además de los estudios *sincrónicos* en la filosofía de la ciencia, hay lugar para un enfoque *diacrónico* sistemático. En los últimos quince años aproximadamente ha crecido la importancia de la filosofía diacrónica de la ciencia. Muchos filósofos con intereses históricos y muchos historiadores con intereses filosóficos han hecho contribuciones valiosas a este campo.

La filosofía diacrónica de la ciencia y la filosofía sincrónica de la ciencia muestran frecuentemente diferencias substanciales en método e intereses. La primera tiende a hacer uso de métodos historiográficos solamente, la segunda a aplicar métodos formales como la axiomática. Esto puede ser la razón por la que han surgido tensiones y polémicas sin fin entre representantes de los dos enfoques. Sin embargo, creo que este conflicto ha sido en gran parte construido artificialmente. No veo ningún motivo real para una contradicción metodológica entre la filosofía sincrónica de la ciencia y la diacrónica: más bien creo que ambos enfoques necesitan de mutua complementación. Un uso combinado de análisis e interpretaciones históricas por un lado y de reconstrucciones formales por otro es más probable que nos lleve a una mejor comprensión de la complicada y cambiante red que es eso que llamamos ciencia.

125679

2. ANALISIS ESTRUCTURAL DE LA CIENCIA

2.1. HACIA UN NUEVO CONCEPTO DE TEORIA EMPIRICA

Según la concepción estándar en filosofía de la ciencia, una teoría científica es un conjunto de enunciados, donde por enunciado se entiende una entidad lingüística con significado, que puede ser verdadera o falsa (una entidad lingüística interpretada). Según esta concepción, las teorías científicas ideales son las que están perfectamente axiomatizadas y formalizadas de modo que usen un cálculo deductivo. La teoría ideal consta de unas fórmulas primitivas no deducibles de otras que, al ser interpretadas sobre un determinado universo, dan lugar a los axiomas específicos de la teoría. De estos axiomas se deducen formalmente otros enunciados. El conjunto de los axiomas y de sus consecuencias lógicas constituye la teoría. En las teorías «menos ideales», que aún no han sido formalizadas en un cálculo deductivo, sino a lo sumo semiformalizadas, las consecuencias lógicas no podrán obtenerse por vía de deducción formal, sino sólo por razonamientos más o menos informales. Pero en cualquier caso, la teoría no es más que un conjunto de enunciados.

A esta concepción general la ha llamado Stegmüller «concepción enunciativa de las teorías»[1] y, aunque la denominación sea quizá discutible, subraya con el calificativo el punto en el que se hace hincapié: las teorías son entidades lingüísticas, a saber, series de enunciados.

Esta es la concepción básica compartida explícita o implícitamente por la mayoría de los filósofos de la ciencia más notables del último medio siglo, por distintos que sean sus puntos de vista respecto de

[1] «Statement view»; cf. W. Stegmüller, *Theorienstrukturen und Theoriendynamik*.

otras cuestiones: Ramsey, Carnap, Reichenbach, Popper, Braithwaite, Nagel, Hempel. Es el punto de partida para los análisis de esos autores, la base desde la cual tratan cuestiones tan diversas como la explicación científica, la confirmación o corroboración de teorías, la distinción entre términos teóricos y observacionales, las relaciones interteóricas, etc.

La concepción lingüística de las teorías tiene dos grandes ventajas: primera, que es extraordinariamente simple, elegante y fácil de comprender; segunda, que es máximamente general: pretende abarcar todos los campos del saber científico (e incluso para-científico), tanto las ciencias formales, como las empíricas, y dentro de estas últimas, tanto las que usan un lenguaje cuantitativo, como las que sólo han llegado al estadio cualitativo. En una palabra, teorías, en el sentido de conjuntos axiomatizados (o axiomatizables) de enunciados, deben ser los constituyentes de todas las ramas del saber que pretendan ser científicas, desde el álgebra abstracta, hasta el psicoanálisis y la etnología.

Es obvio que históricamente esta concepción ha sido tomada de las ciencias formales. Lógicos y matemáticos no han considerado que tenían una idea clara de lo que eran la teoría de conjuntos o la geometría, por ejemplo, hasta que las han tenido perfectamente axiomatizadas y formalizadas. Y con esto han quedado por lo general satisfechos. Si alguien le pregunta a un matemático hoy día: «¿Qué es la aritmética (como teoría)?», éste puede dar una respuesta bien sencilla: mostrar los cinco axiomas de Peano y decir: «Eso es la aritmética (o bien: eso, junto con sus consecuencias lógicas).» Así queda perfectamente claro lo que es la teoría aritmética. Lo mismo puede hacer con la geometría euclídea, con la teoría de grupos o con las diversas teorías de conjuntos existentes.

Muchos filósofos, entre ellos los arriba mencionados, han supuesto que el mismo tipo de respuesta sencilla podía darse, mutatis mutandis, en las ciencias empíricas. Una de las aserciones más claras y explícitas que conozco de este punto de vista aplicado a las ciencias empíricas es la de Braithwaite, justo al comienzo del capítulo segundo de su *Explicación científica*: «Una teoría científica es un sistema deductivo en el que se siguen lógicamente consecuencias observables de la consideración conjunta de hechos observables y el conjunto de hipótesis fundamentales del sistema; por tanto, todo estudio de la naturaleza de una teoría científica es estudio de la del sistema deductivo que se utilice en ella.»

Esto es una expresión del famoso *covering-law model*, el modelo estándard del funcionamiento de las teorías. Según él, tener una teoría es tener un conjunto H de axiomas o hipótesis fundamentales. Si tenemos además un conjunto de hechos ya observados («condiciones iniciales») o, mejor dicho, un conjunto C de enunciados sobre hechos ya observados, podemos inferir de $C \cup H$ un conjunto de resultados

empíricos R. Con una también famosa metáfora de Hempel, las teorías científicas son como máquinas ideales de hacer salchichas: por un lado entran los cerdos, por el otro salen las salchichas. Podemos añadir: si las salchichas que salen responden a nuestras expectativas gastronómicas, consideramos que la máquina funciona bien; en caso contrario, abandonamos la máquina (=la teoría) y tratamos de construir una nueva, que realmente responda a nuestras expectativas.

El lector quizá ya habrá observado que en esta concepción de las teorías empíricas hay una asimetría respecto al caso de las teorías formales. Allí sólo hablábamos de un conjunto de axiomas y sus consecuencias; aquí hay que hablar de condiciones iniciales y de resultados empíricos. Surge la pregunta: ¿está constituida una teoría sólo por el conjunto H, o bien además por los conjuntos C y R? (Si está constituida por C, también estará constituida por R, y viceversa, puesto que, por regla general, habrá simetría entre C y R desde el punto de vista lógico.) A esa pregunta, la respuesta del lógico que conoce bien las teorías matemáticas será seguramente: la teoría sólo está constituida por H. La respuesta del empírico sería seguramente mucho más imprecisa e indecisa: en muchos casos le gustaría poder decir (y de hecho dice) que C y R forman parte de la teoría. El científico dirá, por ejemplo, que las condiciones iniciales C se expresan en el mismo lenguaje (con las mismas funciones métricas, por ejemplo) que el utilizado para H; incluso, si se sustituye H por H', también habrá que cambiar C por C' en la mayoría de los casos, puesto que los conceptos usados serán distintos[2]. Si las teorías son sólo conjuntos de enunciados H, no se ve claro por qué tendrían que afectar otros enunciados C que, según el lógico, no pertenecen a la teoría.

Pero existen aún algunas otras dificultades, de carácter más formal. Supongamos que un investigador pretende sistematizar mediante una ley super-primitiva del tipo

$$\forall x(Px \rightarrow Qx)$$

un dominio de individuos D finito y no-vacío, y sólo ese dominio. Supongamos, además, que el investigador no dispone de una descripción extensional de D, sino sólo intensional (éste es, con mucho, el caso más frecuente en las sistematizaciones de dominios finitos). El investigador formula entonces la siguiente teoría:

$$H: = \forall x(x \in D \rightarrow (Px \rightarrow Qx)).$$

---

[2] Este hecho sencillo corresponde en parte a lo que algunos autores recientes, como Feyerabend, han expresado con algo más de oratoria, diciendo que «todo es teoría» o algo por el estilo.

Ahora bien, independientemente de lo que el investigador *sepa* o no, el dominio D estará constituido por cierto número determinado de individuos: $a_1, ..., a_n$ (puesto que es finito).

$$D = \{a_1, ..., a_n\}.$$

Supongamos que el investigador descubre que a_1 concretamente cumple: $a_1 \in D$ y Pa_1. Su teoría le permitirá hacer el siguiente razonamiento:

$$\begin{array}{ll} H: = & \forall x(x \in D \rightarrow (Px \rightarrow Qx)) \\ C: = & a_1 \in D \wedge Pa_1 \\ \hline R: = & Qa_1 \end{array}$$

Por otra parte, admitamos que los i primeros individuos de D satisfacen el predicado P y los $n-1$ restantes no. Con respecto al dominio de interpretación D, el enunciado H es equivalente entonces a

$$H': = (Pa_1 \wedge Qa_1) \wedge ... \wedge (Pa_i \wedge Qa_i) \wedge \neg Pa_{i+1} \wedge ... \wedge \neg Pa_n.$$

Según la concepción que criticamos, Pa_1 y Qa_1 no son enunciados que pertenezcan a la teoría, puesto que forman parte de C y R en el razonamiento del investigador. Pero, en cambio, constituyen una parte de H', que es una teoría equivalente a H con respecto a su aplicación al dominio D[3], puesto que dice lo mismo sobre D (en efecto, los modelos empíricos que satisfacen H son exactamente los mismos que satisfacen H')[3]. Esta situación es totalmente insatisfactoria desde un punto de vista intuitivo.

Luego, por lo general, será razonable admitir los conjuntos de enunciados C y R en el concepto de teoría. Una teoría empírica ya no será tan sólo un conjunto de axiomas con sus consecuencias lógicas, sino que habrá que añadir condiciones iniciales y resultados[4]. El concepto de teoría empírica se ha hecho así más complejo que el de teoría formal. Pues si la teoría ha de ser un conjunto de enunciados, a la pregunta «¿qué es la teoría T?», el científico ya no puede responder

---

[3] *Pragmáticamente*, H' es distinto de H, puesto que, por lo general, el científico será capaz de enunciar H, pero no H'. No obstante, la concepción lingüística de las teorías prescinde de consideraciones pragmáticas en sus enunciaciones.

[4] En rigor, además de las condiciones iniciales en sentido propio, hay que añadir lo que se llama el *marco de condiciones* implícito en la formulación de la teoría (en inglés, *boundary conditions;* en alemán, *Rahmenbedingungen*) fuera del cual la aplicación de una ley general no tiene sentido.

«eso» y señalar simplemente una lista de unos cuantos axiomas. Ya no se puede dar una lista en absoluto.

Con todo, esto no representa todavía un grave *handicap* para la concepción enunciativa de las teorías. Pues se puede seguir sosteniendo que las teorías son conjuntos de enunciados, aun cuando no se pueda especificar exactamente cuáles son *todos* esos enunciados. Es más, se puede argüir con sentido que los enunciados que realmente «caracterizan» la teoría son los axiomas. Los axiomas constituyen el «núcleo» de la «teoría»; lo demás se da como por añadidura. Lo que «realmente afirma» la teoría sobre el mundo viene dado por los axiomas. Para determinar una teoría basta indicar los axiomas y señalar vagamente qué *tipo* de condiciones iniciales tiene sentido subsumir bajo la teoría y dentro de qué *tipo* de marco de condiciones puede aplicarse la teoría.

La concepción enunciativa, así entendida, contiene un núcleo importante de verdad, pero no «toda la verdad». Es aún en gran parte insatisfactoria precisamente por querer caracterizar las teorías *sólo* como conjuntos enunciados.

Para ver en qué sentido es insatisfactoria esta concepción, vamos a considerar primero brevemente el ejemplo de una teoría empírica bien conocida, la *mecánica clásica de partículas* de Newton, y trataremos de sacar de este ejemplo una caracterización general de las dificultades con que choca la concepción enunciativa.

Prescindiendo ya del problema de la clase de condiciones iniciales y de marcos de condiciones, la pregunta es: ¿Cuáles son los enunciados que definen «nuclearmente» la teoría?

Si alguien que presupone la concepción lingüística abre un libro de texto de mecánica por las primeras páginas creerá encontrar en muchos casos una respuesta enteramente satisfactoria a la pregunta anterior. Así, por ejemplo, si toma un excelente manual de mecánica clásica que se estudia en los primeros cursos universitarios, podrá leer algo como sigue: «Isaac Newton fue el primero en dar una *formulación completa* de las leyes de la mecánica:

1. Todo cuerpo persevera en su estado de reposo o de movimiento uniforme y rectilíneo, a menos que se vea obligado a cambiar ese estado por fuerzas exteriores a él.

2. La razón del cambio de movimiento es proporcional a la fuerza aplicada y se produce según la recta a lo largo de la cual actúa la fuerza.

3. A cada acción se le opone siempre una reacción igual» (subrayado nuestro)[5].

---

[5] Cf. K. R. Symon, *Mechanics*, p. 7

El lector ingenuo exclamará: «¡Ajá! Prescindiendo del problema de las condiciones de aplicación, ya sé lo que es la mecánica clásica: los axiomas de Newton junto con sus consecuencias lógicas.» Y verá confirmada su reacción por el propio Newton, quien puso a la cabeza de sus *Philosophiae Naturalis Principia Mathematica* sus tres «Axiomata sive leges motus», presentando esta base axiomática de modo muy similar, incluso estilísticamente, a los axiomas y postulados que Euclides propuso para su teoría geométrica. Nuestro lector ingenuo verá en ello una estupenda confirmación de la concepción lingüística: la teoría mecánica de las partículas se identifica con tres enunciados, a saber, los tres axiomas de Newton. La teoría de Newton será verdadera en la medida exacta en que lo sean los axiomas. Si uno de ellos resulta falseado por una experiencia, habrá que abandonar la teoría entera (esto es consecuencia del *covering-law model*).

Sin embargo, si tiene la paciencia de seguir hojeando las páginas de los *Principia Mathematica* de Newton, empezará a entrar en confusión. A medida que se adentre en el libro, encontrará cada vez más leyes más o menos generales, «hipótesis», «lemas» y los famosos «Scholia» («recetas» que daba Newton de modo asistemático para resolver problemas mecánicos particulares), cuya relación lógica con los «Axiomata» es poco menos que nula. Mayor confusión, si cabe le depararán manuales actuales avanzados. En el libro de Goldstein[6] utilizado en cursos universitarios superiores, el autor procede con mucha mayor cautela que Newton, y el lector no encontrará nada que se parezca remotamente a una construcción axiomática de la mecánica.

Claro que nuestro paciente lector podrá primero creer que esto es debido al tradicional desinterés que sienten los físicos por cualquier sistematización formal. Pensará que si se formalizara adecuadamente la teoría, se vería que todo se deduce de los axiomas. Pero no es necesario formalizar mucho para darse cuenta de que la ley de la gravitación universal, la ley de Hooke o las leyes de la hidrodinámica, pongamos por caso, todas las cuales se consideran pertenecientes a la mecánica clásica, no se desprenden de los tres axiomas de Newton. ¿Cómo vamos a deducir de estos principios que la fuerza de atracción entre dos cuerpos es inversamente proporcional al cuadrado de las distancias? Con los tres axiomas es consistente tanto una proporcionalidad cuadrática como una cúbica, como otra cualquiera.

Bien, replicará el lector paciente, esto sólo significa que en el sistema axiomático de la mecánica hay que añadir todas esas leyes. Hay que postular como axiomas no sólo los tres de Newton, sino todas las demás leyes más o menos especiales que se consideran pertenecientes a la mecánica clásica. Esto se puede hacer, ciertamente.

---

[6] Cf. H. Goldstein, *Mecánica clásica.*

Pero a partir de este momento surgen por lo menos tres dificultades muy enojosas para la concepción enunciativa de las teorías. Primera, cualquier físico dirá que hay una diferencia manifiesta entre el segundo principio de Newton y la ley de la gravitación, o aún más las leyes del rozamiento, pongamos por caso. El físico dirá que el segundo principio es «mucho más general y abstracto», que es «un principio-guía» que «*permite* la formulación de las demás leyes». Se puede ir más allá y sostener incluso que es prácticamente inimaginable una experiencia que refutara el segundo principio. En realidad, en una construcción axiomática se puede demostrar que, eligiendo convenientemente la función-fuerza, se puede «forzar» cualquier sistema físico que cumpla el segundo principio. ¿Pero cómo se expresa esta noción de «axioma más que general que sirve de principio-guía» en un sistema de enunciados (sin recurrir a expresiones metalingüísticas)? Un sistema axiomático es un sistema democrático. Todos los axiomas que lo componen son «considerados iguales».

Segunda dificultad. Supongamos que se admite que la mecánica clásica vendrá caracterizada unívocamente por *todos* los axiomas que se postulen en ella. Entonces, si queremos ser prudentes y exactos, deberemos admitir también que aún no sabemos qué es la mecánica clásica de partículas. Pues no es inverosímil suponer que dentro de unos años puedan postularse nuevas leyes especiales para sistemas clásicos de partículas, leyes no deducibles de las conocidas hasta ahora. Siendo muy optimistas e imprudentes, podríamos admitir que la mecánica clásica se completó hacia finales del siglo pasado. Desde luego no podríamos decir entonces que Newton fue el fundador de la mecánica clásica. Sin ir más lejos, la importante ley de Coulomb, que es una ley especial de la mecánica clásica de partículas, no fue formulada hasta 1785, cien años más tarde que los *Principia* de Newton. La situación es aún peor, si cabe, para las demás teorías físicas. Está claro que no podemos asegurar de ninguna axiomatización actual de una teoría empírica T que contenga *todos* los axiomas que en el futuro se considerarán intuitivamente como pertenecientes a T. El lógico puro puede replicar: aquí no hay ningún problema, lo que falla es la intuición, el modo de hablar de los físicos. En realidad, dirá, en la época de Newton se tenía una teoría T_1 sobre las partículas; a fines del siglo XVIII, una teoría T_2 tal que $T_1 \subset T_2$, y a fines del XIX, una teoría T_3 tal que $T_1 \subset T_2 \subset T_3$; la expresión «mecánica clásica de partículas» es sencillamente ambigua, pues con ella a veces se entiende T_1, a veces T_2, a veces T_3. Esta réplica es impecable formalmente. Pero no responde a las *condiciones de adecuación* que debemos establecer para una precisión adecuada y plausible del concepto de teoría física. No se trata de inventar un concepto formalmente impecable de teoría física que luego resulte inservible por anti-intuitivo, sino de precisar formalmente lo que se quiere significar realmente cuando se dice que Newton fundó la teoría T (la mecánica clásica de partículas) y que

luego, a lo largo de más de dos siglos, se han ido encontrando «nuevas aplicaciones», como dicen los físicos, es decir, *nuevas leyes especiales* de *esa misma teoría T.* Y tal precisión formal es posible, como luego veremos, si bien a costa de rechazar en buena parte la concepción enunciativa.

Tercera dificultad. Según la concepción que criticamos, una teoría formal se abandona cuando se deduce de ella alguna contradicción (interna); se piensa entonces que, correspondientemente, una teoría empírica se abandonará cuando o bien se deduzca de ella una contradicción interna, o bien una contradicción con la experiencia. Si admitimos la concepción enunciativa, esto no puede ser de otra manera. Dado el conjunto de enunciados H (los axiomas), si de H se deduce una contradicción, se abandona H; si de $C \cup H$ se deduce un $r \in R$ tal que r está en contradicción con nuestras observaciones, también se abandona H. Ahora bien, Thomas Kuhn ha mostrado en innumerables casos concretos que, históricamente, esto es precisamente lo que *no* ocurre. Los científicos predicen con sus teorías cosas que luego no se cumplen, y en vez de abandonar la teoría, olvidan los datos negativos y siguen aplicando la teoría; incluso hallan a veces contradicciones internas en sus teorías. Ello les causa ciertas molestias, pero en realidad no les preocupa mucho. Seguirán aplicando sus teorías «a trancas y barrancas», y sólo las abandonarán en casos muy especiales (las revoluciones científicas), cuando el barco hace agua por todas partes. Si uno se adhiere por una parte a la concepción enunciativa de las teorías, y por otra admite que las descripciones de Kuhn son históricamente correctas, tendrá la impresión de que lo que Kuhn llama la «ciencia normal» es un quehacer de gente sumamente terca, irrazonable y hasta fanática; todo lo contrario de la imagen del investigador objetivo y desapasionado en busca de la verdad, imagen que se ha formado en nuestras mentes después de trescientos años de éxitos científicos.

Vale la pena subrayar que esta impresión negativa acerca de la ciencia empírica que se saca de la lectura de Kuhn proviene (por lo menos en gran parte) de admitir como supuesto implícito o explícito que las teorías *no son más* que conjuntos de enunciados: está claro que si encontramos experiencias que refutan algunos de esos enunciados y, a pesar de ello, seguimos considerando que la teoría como conjunto es válida, seremos gente bastante insensata. Podemos ciertamente llegar a la conclusión de que realmente la inmensa mayoría de los investigadores empíricos son gente irrazonable. Pero, a la inversa, también puede sospecharse que hay algo que falla en la concepción enunciativa, y tratar de llegar a una nueva concepción de las teorías que arroje más luz sobre la aparente irracionalidad de la evolución científica.

Además de las tres expuestas, existen aún otras dificultades importantes con las que tropieza inevitablemente la concepción criticada. Señalemos sólo brevemente el siguiente punto. Para la

concepción enunciativa de las teorías es esencial poder indicar exactamente cuáles son los modelos que satisfacen el sistema formal correspondiente a la teoría. Que esto debe ser así está bien claro para las teorías matemáticas, si éstas se consideran interpretadas. En las ciencias empíricas la situación es completamente distinta. No sólo no está claro nunca cuáles son exactamente los modelos empíricos a los que se aplica la teoría, sino que si a alguien alguna vez se le ocurriera tratar de aclararlo con exactitud total, conseguiría una teoría que sería considerada por el resto de investigadores o bien como falsa o bien como ininteresante por trivial. En este sentido hay en las teorías empíricas un elemento de vaguedad que las hace fructíferas. Este hecho, en un lenguaje algo distinto, también ha sido descrito por Kuhn.

Alguien podría decir, por ejemplo, que los modelos de la mecánica clásica de partículas son precisamente partículas físicas en movimiento. Ahora bien, en rigor, partículas en sentido estricto no las hay por ninguna parte. Un modelo bien conocido de la mecánica clásica de partículas es el sistema Tierra-Luna. Se consideran, pues, la Tierra y la Luna como «partículas». El sentido común consideraría esto como una barbaridad. El físico replicaría que no lo es; que es una idealización permisible gracias, entre otras cosas, al concepto de centro de masa. Pero no siempre tiene sentido hacer tales idealizaciones. El sentido común le podría decir al físico: si la Tierra es una partícula en tu teoría, no veo por qué no ha de serlo también un árbol; a ver cómo me describes el crecimiento de un árbol en tu teoría, pues es también el movimiento de una partícula. El físico respondería seguramente: no tiene sentido tratar de subsumir el crecimiento de un árbol bajo la mecánica clásica de partículas; no es un modelo de esta teoría.

Una jugada de billar es uno de los modelos más obvios de la mecánica clásica de partículas; una tirada de dados es uno de los casos más obvios de algo que *no* es un modelo al que se aplica esa teoría. Al revés del caso del árbol, en principio *tendría* sentido considerar el dado como una partícula en movimiento. Pero el físico *sabe* de antemano que este sistema no «cabe» en su teoría, no debe «interesarle». ¿Cómo se sabe cuáles son los modelos de la mecánica clásica de partículas? ¿Por qué el árbol en crecimiento y el dado no son modelos de esta teoría? ¿Cuáles son las parcelas de la realidad de las que tendría por lo menos sentido considerarlas como partículas, y dentro de este conjunto, cuál es el subconjunto de parcelas que *efectivamente* son modelos de la teoría?

No hay una respuesta unívoca a estas cuestiones. El físico respondería que para saber eso hay que «tener olfato». Esto es lo mismo que decir que no se puede caracterizar formalmente de modo unívoco y definitivo el conjunto de modelos de la teoría. Pueden darse, no obstante, ciertas precisiones más o menos formales, de las

características de estos modelos y de sus relaciones con las teorías. En capítulos subsiguientes trataré con más detalle de esta cuestión.

En los últimos quince años se han levantado algunas voces más o menos confusas en contra de la idea de que las teorías empíricas son caracterizables sencillamente con una lista de axiomas; entre ellas, las de Feyerabend y Kuhn. Los análisis históricos concretos de Kuhn no sólo han desarticulado algunos prejuicios simplistas y «vulgoprogresistas» en cuanto al desarrollo de las ciencias, sino que contienen en germen una alternativa a la concepción enunciativa de las teorías. Los «paradigmas» de Kuhn son, sin duda, mucho más que meros conjuntos de axiomas, aunque, por desgracia, no se sabe exactamente qué más, ni parece saberlo el propio Kuhn a ciencia cierta[7]. Por lo demás, a las propuestas metodológicas de Kuhn y Feyerabend (en la medida en que en estos autores se pueda hablar todavía de metodología), que están basadas en casos históricos concretos, siempre se puede oponer la objeción de que una cosa es *historia* de la ciencia y otra muy distinta *teoría* de la ciencia.

El primero en ofrecer una verdadera y tangible alternativa a la concepción enunciativa de las teorías físicas ha sido J. D. Sneed, en su formidable obra *The Logical Structure of Mathematical Physics*, después de años de arduas investigaciones. Por lo demás, digamos ya en este punto que la metateoría de Sneed no puede considerarse una refutación en sentido estricto de la concepción enunciativa de las teorías, sino un análisis mucho más fino de la estructura de las teorías físicas que el que permitía dicha concepción. Después de estudiar la obra de Sneed, uno se da cuenta de que las teorías físicas son cosas de estructura mucho más compleja de lo que Popper, Carnap o Hempel han imaginado, y piensa que a fin de cuentas no podía ser de otro modo. Ello significa, no obstante, que la premisa necesaria para las investigaciones de Sneed eran ciertos resultados positivos de la concepción enunciativa, en particular, los resultados de la axiomática. Sin las axiomatizaciones de algunas teorías físicas emprendidas por H. Simon. P. Suppes y E. W. Adams, entre otros, Sneed no habría podido llegar a los resultados obtenidos. Por esto es injustificado en este campo el menosprecio de Feyerabend por las axiomatizaciones de teorías empíricas[8]. La axiomatización no es una condición suficiente, pero sí *necesaria* para hacer metateoría de las teorías. Por otro lado, el análisis más fino de Sneed permite reinterpretar de un modo mucho más preciso ciertas tesis expuestas (con métodos casi exclusivamente literarios) por algunos autores recientes, sobre todo Kuhn; entre ellas, la de que las teorías no suelen ser abandonadas aun cuando se encuentren hechos que las «refutan». Esto es debido a que las teorías

---

[7] Cf. los intentos de precisión de Kuhn en su *Postscript*-1969.

[8] Cf., por ejemplo, sus manifestaciones peyorativas acerca del método axiomático en la nota 82 de su *Explanation, Reduction, and Empiricism*, pp. 67-68.

tienen una estructura mucho más compleja que la de una simple lista de enunciados.

Sneed trata en su libro un gran número de cuestiones, algunas totalmente nuevas, y con una perspectiva muy distinta de la habitual en la filosofía de la ciencia actual, tanto la «formalista» como la «antiformalista». En el capítulo siguiente se explica con mayor detalle el sentido de los resultados de Sneed. El objetivo del presente capítulo es sólo describir las dificultades con que choca una concepción sostenida por muchos filósofos y que, por lo demás, hasta ahora era la única concepción básica bien articulada de que se disponía para emprender el análisis de las teorías científicas. Y con tal descripción pretendíamos hacer ver en qué sentido es crucial la obra de Sneed para la filosofía de la ciencia, en qué sentido representa un punto de inflexión. Los fallos de la concepción enunciativa apenas si están insinuados en algunos pasajes aislados del libro de Sneed. De ahí que una primera lectura del mismo haga difícil una apreciación cabal del papel renovador de las ideas expuestas.

El libro de Sneed abre nuevas perspectivas en la filosofía de la ciencia. No es una obra conclusa. En realidad, a pesar de su volumen y densidad, es sólo un esbozo de nuevos métodos metateóricos. Sneed pone a prueba su concepción «sólo» en unas cuantas teorías físicas: la mecánica clásica de partículas, la formulación lagrangiana de la mecánica, la formulación hamiltoniana y la mecánica del sólido rígido. Y aun de una manera detallada solamente en la primera teoría, si bien ésta es una de las más centrales y acabadas de la física. Falta por ver hasta qué punto otras teorías caben en los mismos esquemas. Por otro lado, los puntos oscuros, las formulaciones defectuosas, los errores de detalle son numerosos en los análisis de Sneed. Incluso puntos bastante centrales de la metateoría de Sneed están necesitados, sin duda, de revisión radical. Pero, a pesar de todo ello, por mi parte creo que su concepción es, en líneas generales, correcta, y mucho más fructífera que los análisis metateóricos empíricos que hasta ahora se habían llevado a cabo.

125679

2.2. UN PROGRAMA DE RECONSTRUCCION ESTRUCTURAL DE LAS TEORIAS FISICAS

0. Introducción

El siglo XX ha presenciado, desde la época de los trabajos pioneros de Frege y Russell, enormes avances en el campo de la filosofía de las matemáticas. En cambio, en el campo de la filosofía de las ciencias empíricas hemos sido menos afortunados. Especialmente por lo que se refiere al análisis de la estructura de las teorías *físicas*, que son indudablemente las de estructura más compleja dentro de dicho campo, los resultados logrados han permanecido hasta hace poco a un nivel a todas luces insuficientes.

Dos han sido, a mi entender, los defectos principales responsables de esta situación poco brillante en la filosofía de las ciencias empíricas: (1) los métodos aplicados en los análisis lógicos han sido demasiado simples; (2) la filosofía de la ciencia se ha dedicado casi exclusivamente a construir y discutir esquemas conceptuales de supuesta validez universal para todo el conocimiento empírico, desde las formas más burdas de conocimiento cotidiano hasta las leyes más complicadas de la mecánica cuántica, e incluyendo disciplinas de metodología tan diversa como la física, la psicología y la economía. La consecuencia de este exceso de generalización, con métodos demasiado simples, ha sido la obtención de resultados de escaso contenido informativo y poco relevantes para las teorías científicas más complejas, particularmente, repito, las de la física.

Con estas preliminares observaciones críticas no quiero decir que todo lo hecho en los últimos cuarenta años en filosofía de la ciencia carezca de valor. Muy al contrario, el trabajo realizado es seguramen-

te la premisa necesaria para desarrollos ulteriores y, además, ha permitido ya análisis relativamente convincentes de las formas más simples del conocimiento humano, como son las sistematizaciones empíricas que hacemos en nuestra vida cotidiana o en ciencias poco desarrolladas. Y ésta es, sin duda, una zona muy significativa del conocimiento de la mayoría de los hombres. Pero está claro que no debemos quedarnos ahí, sino que hay que pasar a niveles más teóricos. Este capítulo presenta un programa de factura bastante reciente para acometer tal empresa. (Aquí me referiré sobre todo a la reconstrucción de teorías físicas, aunque mucho de lo que se expone podría aplicarse también al caso de teorías matematizadas y «teoretizadas» de otras disciplinas. En realidad, el término «teoría física» se usa aquí como abreviación para «teoría física o teoría que tiene una estructura similar a las teorías físicas». Esto incluye algunas teorías de la economía y de la psicología, por ejemplo.)

El primer paso que hay que dar para remediar la situación de indigencia señalada es no pretender generalizar demasiado y analizar y reconstruir teorías físicas *concretas*. Para ello es condición necesaria *axiomatizarlas* correctamente. Los primeros pasos en esa dirección se han dado ya en los últimos treinta años por autores como Simon, Suppes, Adams y otros. Los esfuerzos por conseguir axiomatizaciones formalmente correctas se han concentrado en la mecánica clásica y en la teoría de la relatividad.

Sin embargo, contrariamente a lo que ocurre en matemáticas, no basta con axiomatizar una teoría física para clarificar su estructura lógica. Las teorías físicas son estructuras complejas cuyos componentes no quedan suficientemente analizados dando sólo una lista de axiomas. Si consideramos las teorías físicas meramente como sistemas de axiomas, aparecen en seguida graves dificultades tanto de carácter formal como de contenido, que se han expuesto en el capítulo precedente.

Pero, además, no sólo ocurre que la estructura estática de las teorías físicas es más compleja que la de las teorías matemáticas; un aspecto adicional de diferenciación es la importancia que tienen en las teorías físicas los momentos evolutivos, históricos, o lo que puede llamarse la dinámica interna de las teorías físicas.

Finalmente, un tercer elemento de diferenciación con respecto a las teorías matemáticas es el papel que juega la llamada «intuición física» en la determinación del campo de aplicaciones de la teoría, o sea, de su conexión con la «realidad». Este es un elemento fuertemente pragmático que no tiene parangón en las teorías matemáticas.

Los filósofos de la ciencia de la primera etapa, deslumbrados por los brillantes resultados en metamatemáticas, olvidaron o prescindieron de los aspectos diferenciales mencionados. Esta negligencia provocó en la década de los años sesenta una reacción en sentido contrario, y autores de diversa procedencia criticaron el análisis

axiomático-formal para propugnar una aproximación historicista e intuitiva a las teorías empíricas. La culminación de esta corriente puede contemplarse en el nihilismo oscurantista del último Feyerabend, según el cual, cuanto más oscura sea una teoría, tanto mejor.

La obra más conspicua y, me atrevo a añadir, más seria de dicha reacción historicista fue sin duda *La estructura de las revoluciones científicas*, de Thomas S. Kuhn. A diferencia de Feyerabend y otros, Kuhn no partió de ningún *a priori* filosófico para criticar la filosofía de la ciencia tradicional, ni tampoco rechazó dogmáticamente cualquier tipo de análisis lógico de las teorías empíricas. Sólo expuso un análisis histórico detallado de las formas de existencia de las teorías científicas hasta la fecha y llegó a la conclusión, enteramente justificada, de que el resultado de su análisis histórico no concuerda con la imagen de la ciencia ofrecida por los filósofos de la ciencia anteriores, tanto inductivistas como deductivistas, tanto empiristas como racionalistas. La obra de Kuhn puede interpretarse no tanto como una filosofía de la ciencia alternativa, sino como el reto filosófico de un historiador a los filósofos de la ciencia «profesionales». El reto consiste en exigir de los filósofos un aparato conceptual preciso que dé cuenta de los aspectos de las ciencias empíricas que Kuhn ha revelado en sus análisis históricos. Los aspectos, a mi entender, realmente cruciales revelados por Kuhn pueden resumirse en los siguientes puntos:

(1) Una teoría empírica es algo más que un conjunto de hipótesis o axiomas.

(2) Las aplicaciones empíricas de una teoría forman parte esencial del concepto mismo de la teoría. La estructura teórica y sus aplicaciones están conceptualmente ligadas de manera estrecha.

(3) Una teoría empírica es una entidad que evoluciona en el tiempo; esta evolución es lo que Kuhn llama «ciencia normal».

(4) En la determinación de una teoría interviene un elemento oscuro que Kuhn denomina «paradigma». El propio Kuhn ha reconocido en el *Postcript* a la segunda edición de su libro que su uso original de este término fue confundente y que él mismo no tenía las ideas muy claras al respecto, pero es obvio que ese término denota algo esencial a las teorías científicas, que había escapado a los análisis anteriores.

(5) El rechazo de hipótesis particulares de una teoría durante un período de «ciencia normal» no significa el rechazo de la teoría misma. Científicos que usan la misma teoría pueden divergir en cuanto a las hipótesis empíricas que postulan mediante la teoría.

(6) Una teoría se acepta o se rechaza en su totalidad, no por partes.

(7) Las revoluciones científicas son eventos en los que una teoría sustituye completamente a otra, con profundos cambios conceptuales y sin que medie una relación lógica directa entre ambas teorías.

Las tesis de Kuhn provocaron una década de fuertes polémicas en el mundo filosófico. Ellas parecían a mucha gente incompatibles no sólo con las concepciones metodológicas usuales, sino también con cualquier programa de reconstrucción racional de la empresa científica.

Sin embargo, lo único que los análisis kuhnianos realmente contradecían era, como lo ha mostrado Stegmüller detalladamente, una concepción simplista de la ciencia, según la cual las teorías científicas son meros conjuntos de axiomas con sus consecuencias lógicas, que pueden ser confirmadas o refutadas directamente por hechos aislados. Kuhn había mostrado la necesidad de una concepción más compleja y profunda de las teorías científicas, concepción que él mismo, por falta de herramientas lógico-matemáticas, no había podido formular con claridad.

En este contexto surgió una nueva metateoría de las ciencias, a la que, siguiendo a Yehoshua Bar-Hillel, podemos bautizar, a falta de nombre mejor, como la concepción estructural de las teorías empíricas. (El epíteto «estructural» no indica ninguna conexión con las escuelas llamadas «estructuralistas» en lingüística y etnología, aparte de subrayar la importancia del análisis de estructuras globales en las ciencias.)

La concepción estructural, sin este nombre, fue inaugurada por Sneed en 1971. Aunque el libro de Sneed es de difícil lectura, sus ideas han ido penetrando paulatinamente en el mundo filosófico gracias, sobre todo, a los esfuerzos de divulgación de Stegmüller en su *Theorienstrukturen und Theoriendynamik* y en otros ensayos posteriores (cf. en especial «Estructuras y dinámica de teorías»). Finalmente, el propio Kuhn se ha acercado a dicha concepción y ha admitido su adecuación para reproducir en forma lógica, exacta, los puntos antes reseñados de sus análisis, con la posible salvedad del último, es decir, de las revoluciones científicas (cf. «A Letter to W. Stegmüller Concerning *Theorie und Erfahrung*» y «A Formalism for Scientific Change»).

En realidad, Sneed no escribió su libro para tratar fundamentalmente de la problemática kuhniana. Su objetivo principal era otro: desarrollar un nuevo aparato conceptual con el cual reconstruir formalmente la estructura estática de las teorías físicas, así como su contenido empírico. Sin embargo, como efecto secundario, ese aparato permite dar un sentido preciso a las tesis de Kuhn. El último capítulo

del libro de Sneed contiene ya un esbozo de la aplicación a la dinámica de las teorías según Kuhn. Pero ha sido Stegmüller quien más ha profundizado en esta reinterpretación estructural de Kuhn. Mi intención en lo que sigue es explicar los aspectos fundamentales de la concepción estructural y señalar su conexión con las tesis de Kuhn.

1. Axiomatización por definición de un predicado conjuntista

Los elementos mínimos del análisis estructural de las teorías físicas son sus *modelos* y no sus enunciados. Los modelos de una teoría son los correlatos formales de los trozos de realidad que la teoría explica. El concepto de modelo es pues fundamental.

Un modo posible de explicar el concepto de modelo de una teoría empírica sería construir primero un lenguaje formal en el cual se expresase la teoría en cuestión, dar luego una interpretación semántica de ese lenguaje sobre un universo empírico y definir finalmente la noción de satisfacción de una fórmula del lenguaje en ese universo. Se dice que la interpretación es un modelo de un determinado conjunto de fórmulas que se toman como axiomas si esas fórmulas son satisfechas en ese universo bajo esa interpretación.

Esta precisión del concepto de modelo de una teoría, que es la que se suele emplear en los estudios metamatemáticos, es formalmente intachable, pero tiene algunos inconvenientes prácticos cuando se trata de teorías físicas. El mayor seguramente es que supone la construcción previa de un lenguaje formal y de una interpretación sobre un universo dado, lo cual resultaría considerablemente trabajoso y complicado en el caso de las teorías físicas.

Afortunadamente disponemos de otro procedimiento para definir el concepto de modelo, mucho más práctico, y que, aunque no es completamente formal, es suficientemente exacto. Con este procedimiento se puede axiomatizar perfectamente la teoría e indicar inmediatamente qué cosas son los modelos de la misma sin necesidad de presuponer un lenguaje formal y una interpretación. Se trata de lo que se llama «axiomatización por introducción de un predicado conjuntista». (Una exposición sistemática de este procedimiento se hallará en los últimos capítulos de la *Introducción a la Lógica* de P. Suppes.) Es el método empleado por Sneed en sus axiomatizaciones.

Consideremos un ejemplo concreto de axiomatización de una teoría física: la mecánica clásica de partículas (en versión algo simplificada), que nos servirá también para todas las discusiones subsiguientes.

La estructura básica de esta teoría se axiomatiza definiendo, mediante el instrumental de la lógica y la teoría de conjuntos, el predicado «MCP» (léase «mecánica clásica de partículas»). Las condiciones de definición de este predicado son lo que normalmente se considerarían como axiomas de la teoría:

Definición: MCP (x) si y sólo si existen P, T, s, m, f, tales que:

(1) $x = < P, T, s, m, f >$.
(2) P es un conjunto finito no-vacío (que representa un conjunto de *partículas* físicas).
(3) T es un intervalo cerrado de números reales (que representa el *intervalo temporal* durante el cual se consideran las partículas).
(4) s es una función del producto cartesiano $P \times T$ en el espacio vectorial R^3, y s es dos veces diferenciable en T (s representa la función que determina la *posición* en el espacio de cada partícula en cada instante).
(5) m es una función de P en los números reales positivos (que representa la *masa* de cada partícula).
(6) f es una función de $P \times T$ en R^3 (que representa la *fuerza* resultante que actúa sobre cada partícula en cada instante).
(7) Para todo p en P y para todo t en T se cumple:

$$m(p) \cdot D_t^2 s(p, t) = f(p, t)$$

(es decir, la masa de una partícula multiplicada por la derivada segunda de su posición respecto al tiempo es igual a la fuerza resultante).

La condición (7) es naturalmente la condición esencial o ley fundamental de esta axiomatización: se trata del famoso «segundo principio de Newton», que suele enunciarse en forma breve diciendo que «fuerza es igual a masa por aceleración».

Ahora podemos decir inmediatamente qué es un *modelo de la mecánica clásica de partículas:* es simplemente cualquier entidad que satisfaga el predicado MCP, o sea, cualquier entidad constituida a su vez por otras cinco entidades (un conjunto de partículas, un intervalo temporal, una función de posición, una función masa y una función fuerza) que cumplan las condiciones estipuladas (los axiomas de la teoría) y en particular que estén entre sí en la relación especificada en la condición (7) (segundo principio de Newton).

En general, si la teoría no es trivialmente vacía, habrá innumerables modelos de la misma en este sentido que acabamos de precisar, y todos ellos serán igualmente interesantes. (Esta es otra diferencia importante con respecto a las teorías matemáticas, donde por lo general se prefiere que haya *un solo* modelo estándar)[9].

Esta multitud de modelos de la misma teoría, todos determinados por la misma ley fundamental (en el presente caso el segundo

[9] Para los matemáticos, las teorías *categóricas*, es decir, aquellas cuyos modelos son todos isomorfos entre sí (que es como decir que tienen un solo «tipo» de modelo), son particularmente atractivas. En física, en cambio, se trata de que las teorías se alejen, por así decir, lo más posible de la categoricidad.

principio de Newton), corresponden a diversas aplicaciones de la teoría a la realidad. Un modelo de MCP puede corresponder, por ejemplo, al sistema planetario en su totalidad, otro modelo al sistema Tierra-Luna-Sol, otro modelo a la trayectoria de un proyectil sobre la superficie de la Tierra, otro a una bolita suspendida de un alambre oscilando, etc. No tiene objeto postular la existencia de un único modelo universal de la teoría, pues esto no respondería a la praxis del científico; éste aplica su teoría solamente a pequeños trozos de la realidad.

Las entidades que intervienen en la teoría se especifican en cada modelo. En principio no tendrán nada que ver con las entidades correspondientes que aparecen en otro modelo. En particular, cada modelo tendrá sus propias funciones. En el caso concreto que estamos examinando, supongamos que tenemos una serie de modelos distintos de la teoría, todos los cuales satisfacen el predicado MCP,

$$a_1, a_2, \ldots, a_i, \ldots .$$

Las funciones posición, masa y fuerza serán distintas en cada uno de ellos (pues su dominio de definición y su recorrido de valores son también distintos). Tendremos, pues, las funciones

$$s_1, s_2, \ldots, s_i, \ldots$$
$$m_1, m_2, \ldots, m_i, \ldots$$
$$f_1, f_2, \ldots, f_i, \ldots .$$

A cada una de estas funciones se la llama una «función concreta» (o «función en sentido propio»): por ejemplo, f_i es la función concreta fuerza correspondiente al i-ésimo modelo de la mecánica clásica de partículas. Pero como, en general, al hablar de una función física, no se especifica el modelo particular en el que está definida, conviene introducir la noción de «función abstracta» (o «función en sentido impropio»): es simplemente la clase de todas las funciones que ocupan el mismo lugar en los tuplos que caracterizan los modelos de la teoría. Por ejemplo, la fuerza «abstracta» es la clase de funciones-fuerza $\{f_1, f_2, \ldots, f_i, \ldots\}$. En las discusiones filosóficas y metodológicas en que se habla de la fuerza, por ejemplo, se entiende por tal generalmente la noción abstracta, no las funciones concretas.

2. Condiciones de ligadura

Ahora bien, todos esos modelos de la teoría no aparecen aislados entre sí, sino que están interconectados formalmente tal como corresponde a la realidad intuitiva, formando una estructura global.

Frecuentemente, aunque no siempre, estas interconexiones se deben a que el mismo individuo aparece en modelos diversos: por ejemplo, el planeta Tierra aparece tanto en un modelo planetario como en uno que estudia la trayectoria de un proyectil. Estas relaciones entre modelos distintos (nuevamente a diferencia del caso de las matemáticas) son esenciales a la teoría misma, pues ellas permiten la mayor parte de las formas interesantes de predicción y explicación por medio de la teoría. A estas relaciones «inter-modélicas» las llama Sneed «constraints»; este término está tomado de los manuales corrientes de física, donde posee un sentido distinto, aunque relacionado. La traducción castellana de ese término en los manuales suele ser «ligadura»; podemos pues llamar a estas relaciones entre modelos distintos «condiciones de ligadura», pues ligan de determinadas maneras fijas los valores que pueden tomar las funciones correspondientes de los diversos modelos.

Mostremos ejemplos concretos de tales condiciones en el caso de la mecánica clásica de partículas. Sean x, y, dos modelos distintos que satisfacen el predicado MCP y que tienen algunos individuos en común. (Un ejemplo sería el modelo constituido por el sistema planetario y el modelo constituido por el sistema Tierra-Luna: son dos sistemas distintos que satisfacen el predicado MCP, pero tienen un individuo en común, a saber: el planeta Tierra.) Las funciones concretas que intervienen en uno y otro modelo, como ya hemos dicho, son en principio completamente distintas (desde un punto de vista formal); pero, empíricamente, deberá haber cierta relación entre ellas si tienen argumentos en común.

Supongamos que el individuo a (la Tierra en el ejemplo considerado) aparece tanto en el modelo x como en el modelo y. Entonces en la teoría de la mecánica clásica de partículas se exigirá sin duda que la masa de a en x sea idéntica a la masa de a en y (la masa de las partículas es invariante con respecto a las distintas aplicaciones de la teoría clásica):

$$m_x(a) = m_y(a).$$

Esta condición es intuitivamente trivial, pero formalmente no puede postularse en las condiciones del predicado MCP porque, como ya hemos observado, este predicado se refiere sólo a cada sistema tomado aisladamente y no a pares, o en general, conjuntos de sistemas. Esa condición de ligadura de las funciones masa de los modelos x e y debe postularse adicionalmente, de lo contrario no representaríamos adecuadamente el contenido empírico de la teoría. Lo mismo puede postularse para las demás funciones.

Otra condición de ligadura que hay que postular en la mecánica clásica de partículas es la siguiente. En la física existen con frecuencia

ciertas magnitudes, llamadas «extensivas», para las cuales se introduce una operación de composición o combinación de individuos (a la que denominaremos «o»), que debe cumplir ciertas condiciones. La masa, como función abstracta, es una de estas magnitudes extensivas.

Esto significa lo siguiente. Supongamos que tenemos el individuo a en el modelo x (por ejemplo un péndulo simple) y el individuo b en el modelo y (otro péndulo simple), y combinamos a con b: a o b (construimos un péndulo compuesto). Si queremos que este individuo compuesto forme parte de un modelo z de la mecánica clásica de partículas, que además esté con los modelos de partida x e y en conexión empíricamente adecuada, deberemos exigir que se cumpla:

$$m_z(a \text{ o } b) = m_x(a) + m_y(b).$$

Esto se expresa a veces diciendo que «la masa de un sistema compuesto es la suma de las masas de los sistemas componentes». Esta es también una condición de ligadura que conecta *tres* modelos distintos (mejor dicho: las tres funciones masa de tres modelos distintos) de la mecánica clásica de partículas.

Otra condición de ligadura sería la invariancia de las funciones mecánicas respecto a las transformaciones de Galileo.

Todas estas condiciones de ligadura pueden considerarse como axiomas de la teoría, pero tienen un *status* distinto de los axiomas que aparecen como condiciones de definición de los modelos de la teoría, pues las condiciones de ligadura se aplican a varios modelos a la vez. Por otro lado son una parte esencial e imprescindible de la teoría, sobre todo para hacer predicciones, ya que ellas nos permiten deducir mediante puro cálculo los valores de las funciones en un modelo partiendo de los valores de las funciones correspondientes en otro u otros modelos distintos.

Aquí no entraré en la definición formal y general del concepto de condición de ligadura, pues es algo complicada. Creo que con los ejemplos expuestos está bastante clara, por lo menos intuitivamente, su naturaleza. Sólo quisiera añadir a este respecto que las condiciones de ligadura constituyen un elemento muy importante en la identificación de cada teoría física, quizás incluso más importante de lo que el propio Sneed sugiere. No sólo en la mecánica clásica hay seguramente más condiciones de ligadura de las postuladas por Sneed, sino que en otras teorías, como la termodinámica clásica, juegan un papel aún más importante que en la mecánica clásica de partículas (cf. cap. 2.6).

Podemos entrever ya un primer sentido en el que una teoría es más que el conjunto de axiomas que determinan sus modelos, pues existen importantes condiciones adicionales que ligan los distintos modelos entre sí.

3. Dos niveles

A semejanza de las concepciones más tradicionales, la metateoría de Sneed también considera esencial la distinción entre dos niveles conceptuales dentro de una teoría física: un campo de aplicaciones empíricas y una superestructura teórica que sistematiza dicho campo. El contenido empírico de una teoría viene determinado por la relación postulada entre ambos niveles.

Pero a diferencia de sus predecesores, Sneed no identifica esa distinción entre dos niveles con la dicotomía absoluta y general entre lenguaje observacional y lenguaje teórico. La distinción propuesta por Sneed es *relativa* a cada teoría concreta analizada, según el funcionamiento de los conceptos dentro de la teoría. Dada la teoría T, algunos de sus conceptos funcionan de modo que han de considerarse como parte del campo de aplicaciones empíricas de T, mientras que otros funcionan como conceptos de la superestructura teórica de T. Y puede ocurrir, y ciertamente ocurre, que un mismo concepto que forma parte de la superestructura teórica de la teoría T, funcione en cambio como parte de la base empírica en otra teoría T'. (Es lo que ocurre, por ejemplo, con los conceptos dinámicos, que son teóricos en la mecánica y no-teóricos en la termodinámica.)

Se plantea entonces el problema de hallar un criterio funcional que permita clasificar, para cada teoría física, los conceptos que intervienen en la misma en dos grupos: los pertenecientes a la base empírica (o campo de aplicaciones) y los pertenecientes a la superestructura.

El criterio de clasificación buscado deberá atender a los rasgos esenciales de los conceptos que aparecen en una teoría. En las teorías físicas tenemos por lo general conceptos *métricos o magnitudes*, es decir, determinados tipos de *funciones* matemáticas (reales, complejas, vectoriales). La clasificación deberá referirse, pues, a ciertos rasgos distintivos de estas funciones. Según sean estos rasgos, unas funciones pertenecerán a la base empírica y otras a la superestructura teórica.

Ahora bien, la base empírica de una teoría dada es la base sobre la que la teoría se comprueba (es decir, se confirma o se refuta). Luego, las funciones de la base empírica deberán ser aquellas cuyos valores puedan determinarse (por lo menos en algunos casos) sin presuponer la validez de la teoría en cuestión. Si la medición de los valores de todas las funciones supusiera siempre la validez de la teoría, ésta sería *a priori* válida, por tanto, no sería una teoría empírica.

A las funciones cuya medición no presuponga la validez de la teoría T las podemos llamar T-no-teóricas. Pero puede haber, y en realidad hay, funciones cuya determinación presupone la validez de la teoría. A éstas las llamaremos funciones *teóricas*. Para poder expresar formalmente el contenido empírico de la teoría teniendo en cuenta la existencia de funciones teóricas, Sneed utiliza una forma altamente modificada del conocido método de Ramsey para tratar conceptos

teóricos. No tiene objeto entrar aquí en los detalles técnicos de esa modificación; pueden analizarse en los libros de Sneed o Stegmüller mencionados.

La distinción estructural entre funciones teóricas y no-teóricas dentro de una teoría dada es independiente de la dicotomía tradicional entre conceptos observacionales y teóricos. Se trata de una distinción relativa a cada teoría, basada en el «funcionamiento» de los conceptos en cuestión, y no en su «significado». En suma, la dicotomía tradicional, propuesta por Carnap y otros, es *semántica*, mientras que la distinción propuesta por Sneed es *funcional*, y por tanto de cierto carácter pragmático.

4. El contenido empírico de la teoría

El campo de aplicaciones empíricas (llamémosle *A*) viene precisado formalmente como un conjunto de estructuras, cada una de las cuales es un modelo *parcial* de la teoría, parcial en el sentido de que en él no aparecen funciones teóricas, sino sólo descripciones no-teóricas. Por ejemplo, en el caso de la mecánica clásica se desprende del criterio funcional de teoricidad que masa y fuerza son conceptos teóricos, mientras que posición y tiempo son no-teóricos –relativamente a la mecánica clásica–. Los modelos parciales serán entonces descripciones puramente cinemáticas de los sistemas físicos, es decir, descripciones en las que sólo intervienen parámetros espaciotemporales. El campo de aplicaciones empíricas de la mecánica clásica es un conjunto de tales descripciones cinemáticas.

Si tenemos en cuenta que tanto los modelos teóricos como los modelos parciales están interconectados por las condiciones de ligadura antes mencionadas, se comprende que la relación entre la base empírica y la superestructura teórica, que expresa el contenido empírico de la teoría, no podrá establecerse tomando los modelos parciales y teóricos individualmente, sino que el conjunto de los modelos parciales como un sólo cuerpo deberá ser subsumible bajo los modelos teóricos tomados como un todo. Esto corresponde también al aspecto holista señalado por Kuhn. Esta relación es bastante complicada, pero perfectamente expresable dentro del formalismo estructural.

5. El núcleo

Hasta aquí hemos reseñado los siguientes elementos de una teoría: el predicado conjuntista, que incluye la ley fundamental de la teoría y que fija sus modelos; las condiciones de ligadura que conectan los distintos modelos entre sí; la distinción entre modelos parciales y

modelos teóricos, que permite dar un sentido preciso al contenido empírico de la teoría. Todos estos elementos se hallan interrelacionados entre sí formando una estructura compleja y muy general que Sneed denomina el *núcleo* de la teoría. El núcleo corresponde a la parte formal de lo que Kuhn llama «el paradigma», dado que esta estructura no es cuestionada por los científicos mientras no se produzca una revolución. El núcleo es el presupuesto mismo del uso de la teoría en sus aplicaciones.

6. El conjunto de aplicaciones paradigmáticas

No obstante, la intuición kuhniana de paradigma no queda agotada con el concepto estructural de núcleo. Cuando Kuhn habla de paradigma, muchas veces se refiere a ciertas aplicaciones especiales de la teoría que sirven de ejemplos paradigmáticos para el uso de la teoría en otros casos distintos. De hecho, Kuhn ha admitido recientemente que ésta fue su intuición primaria al introducir la discutida noción de paradigma (cf. «A Letter to W. Stegmüller...»). En nuestra terminología se trata de una parte del campo de aplicaciones empíricas, o sea, de un subconjunto (generalmente pequeño) del conjunto total de los modelos parciales de la teoría: A_0 contenido en A.

En el ejemplo de la mecánica de partículas, el conjunto paradigmático de aplicaciones estaría constituido por los primeros modelos parciales que Newton consiguió explicar con su teoría, a saber, el sistema planetario, con sus subsistemas, así como los cuerpos en caída libre sobre la Tierra, y algún otro ejemplo. Estas aplicaciones primarias sirvieron de guía heurística para el posterior desarrollo de la teoría newtoniana y nunca más fueron abandonadas. Son tan esenciales para el aprendizaje del uso de la teoría que pueden considerarse como parte misma de la teoría. Es decir, en la definición del concepto de teoría física interviene no sólo el núcleo, sino también el conjunto de aplicaciones paradigmáticas. Para identificar una teoría física no basta con captar su estructura formal especificada en el núcleo, sino que además hay que conocer sus aplicaciones paradigmáticas. Esto responde a la idea kuhniana de que parte de las aplicaciones de una teoría son indisolubles de la noción misma de la teoría.

Para averiguar cuáles son exactamente las aplicaciones paradigmáticas no puede recurrirse a un criterio formal. Hay que acudir a la «biografía» de la teoría, es decir, hay que tener en cuenta aspectos histórico-pragmáticos. Hay que ver cómo nació la teoría, cómo evolucionó, qué aplicaciones sirvieron de ejemplos distinguidos para hallar nuevas aplicaciones. Este es uno de los sentidos en que la historia de la ciencia es relevante para la filosofía de la ciencia.

7. Leyes especiales

El núcleo y las aplicaciones paradigmáticas constituyen la porción intangible, inatacable, de una teoría –a menos que se produzca una revolución. Núcleo y aplicaciones paradigmáticas se constituyeron conjuntamente en el inicio de la teoría y son el supuesto indiscutido con el que se trabaja durante un período de ciencia normal.

Pero hay otros elementos en una teoría que son cambiantes y no substanciales para la identidad de la teoría. Se trata, por una parte, de las aplicaciones no-paradigmáticas que se proponen, se analizan y eventualmente se aceptan o se rechazan en base al núcleo y a los ejemplos paradigmáticos previos; parte del trabajo cotidiano del científico normal que dispone ya de un paradigma consiste en ir buscando nuevas aplicaciones. Pero otra parte de ese trabajo consiste en añadir a la teoría *leyes especiales*, más estrictas que la ley fundamental, pues ésta suele ser de carácter tan general (como en el caso del segundo principio de Newton) que con ella sola no puede hacerse gran cosa. Las leyes especiales son válidas sólo para ciertos subconjuntos de A, y se introducen con el fin de obtener medidas o predicciones cada vez más exactas.

Las leyes especiales no forman parte de la teoría en sentido estricto. Son «ampliaciones» del núcleo, que varían con el desarrollo de la teoría y que no se pretende que sean universalmente válidas. Los resultados obtenidos con ellas son sólo de validez limitada a ciertos *tipos* de aplicaciones propuestas. En algunos casos nos daremos cuenta de que estos resultados obtenidos mediante leyes especiales son falsos; ello no afecta ni al núcleo ni al conjunto de aplicaciones paradigmáticas A_o. En tales casos simplemente cambiamos de ley especial.

Utilizando una metáfora zoológica, podemos comparar una teoría con un pulpo: la cabeza del pulpo sería el núcleo; el fondo del mar, del que el pulpo obtiene su alimento, sería el campo de aplicaciones y los tentáculos representarían las leyes especiales. Para el pulpo (la teoría) sólo es vital e imprescindible su cabeza (el núcleo) y una cierta porción segura del ambiente en que vive (las aplicaciones paradigmáticas). En cambio, al pulpo (a la teoría) se le pueden cortar algunos tentáculos (leyes especiales) sin que deje de existir como pulpo (como teoría). Incluso a veces puede regenerar nuevos tentáculos (construir nuevas leyes especiales).

Hablando más formalmente, las leyes especiales resultan de restringir la extensión del predicado fundamental (MCP, por ejemplo), añadiendo nuevas condiciones de definición, o sea, nuevos axiomas. Este proceso puede continuar indefinidamente, clasificando los modelos de la teoría de manera cada vez más fina, según sean las leyes especiales que cumplan. Por ejemplo, al segundo principio de Newton, expresado en la condición (7) de la definición antes expuesta,

podríamos añadir una condición (8) que expresase formalmente la ley de acción y reacción (que, como es sabido, no es estrictamente válida para todas las aplicaciones propuestas de la mecánica clásica de partículas, pero sí para un enorme conjunto de ellas). Así introduciríamos un predicado conjuntista más restringido que MCP, al cual podríamos llamar MCPN («mecánica clásica de partículas newtoniana»). Luego, para obtener predicciones aún más exactas, y válidas sólo para determinados tipos de sistemas, podríamos introducir condiciones aún más restrictivas, como la ley de la gravitación (para sistemas que se mueven libremente en el espacio) o la ley de Hooke (para deformaciones elásticas), y así sucesivamente. La postulación sucesiva de leyes especiales constituye el otro aspecto de la «dinámica de las teorías».

Los avances y retrocesos en el campo de las aplicaciones propuestas y los avances y retrocesos en la postulación de nuevas leyes especiales constituyen los dos planos, el «empírico» y el «teórico», respectivamente, en que se mueve la ciencia normal. Ninguna de esas variaciones puede considerarse «revolucionaria». Las revoluciones científicas son únicamente los cambios profundos que atañen a la parte esencial de la teoría: el núcleo. Por eso son cambios eminentemente *conceptuales*, y en ese sentido puede interpretarse la aseveración de Kuhn de que las revoluciones científicas implican cambios en «la visión del mundo», y no meramente en los hechos conocidos.

Aunque esta exposición de los elementos principales de la concepción estructural ha tenido que ser aquí necesariamente en términos informales, quisiera subrayar que todos esos elementos son precisables por medio de las nociones de la lógica, de la teoría de conjuntos y la teoría de modelos. La metateoría estructural no pretende hacer ninguna concesión a oscuridades supuestamente inherentes a la ciencia.

2.3. FORMA Y FUNCION DE LOS PRINCIPIOS-GUIA EN LAS TEORIAS FISICAS

0. Introducción

En el *Postscrip* de 1969 a la segunda edición de *La estructura de las revoluciones científicas*, Thomas Kuhn se dispuso a replicar a una serie de objeciones que se le habían hecho a raíz de la publicación de la primera edición de su libro en 1962. Entre esas objeciones, una de las que Kuhn tomó más en serio fue la de que su famosa noción de «paradigma», como estructura subyacente a toda una disciplina durante un período de «ciencia normal», era irremediablemente vaga y ambigua. Kuhn trató de precisar dicha noción, y para ello incluso cambió de nomenclatura: el «paradigma» de la primera edición se convirtió en la «matriz disciplinaria» del *Postscript*. No parece que el nuevo término haya tenido buena fortuna; todo el mundo sigue hablando de «paradigmas», y, lo que es peor, sigue hablando de ellos con la misma vaguedad con que Kuhn utilizó el término en la primera edición de su obra. No se trata, naturalmente, de una mera cuestión de nombre (y, a efectos de la presente discusión, podemos seguir utilizando el popular término «paradigma», en vez de «matriz disciplinaria»). Lo que importa es que Kuhn intentó precisar la estructura de su noción básica y para ello identificó lo que, en su opinión, son los cuatro componentes esenciales de un paradigma: generalizaciones simbólicas, modelos ontológicos o heurísticos, valores metodológicos y ejemplares «modélicos» de aplicaciones. Poco después, Stegmüller, en su *Theorienstrukturen und Theoriendynamik*, utilizó el aparato formal de la metateoria de Sneed en *The Logical Structure of Mathematical Physics* para precisar formalmente las

nociones involucradas en los componentes primero y último identificados por Kuhn en un paradigma. En particular, según el análisis de Stegmüller, las «generalizaciones simbólicas» de Kuhn no son sino las *leyes fundamentales* del llamado «núcleo estructural» de una teoría.

El presente capítulo trata de ese primer componente de los paradigmas identificado por Kuhn y precisado por Sneed-Stegmüller. Para empezar, por razones que luego se verán más claras, propongo un nuevo nombre para lo que Kuhn llama «generalizaciones simbólicas» y Stegmüller «leyes fundamentales», aun a costa de la consiguiente inflación terminológica. Propongo llamar al primer componente de los paradigmas de Kuhn «*principios-guía*», una denominación que, por lo demás, ya existe en la literatura metodológica. A mi entender, ese rótulo refleja mejor la función de las entidades en cuestión que las anteriores denominaciones. Pero, naturalmente, el nombre es lo de menos. De lo que se trata aquí es de ofrecer un análisis más completo y detallado (y también más «comprometido») que el ofrecido hasta ahora. La concepción de los principios-guía aquí propuesta pretende ser válida al menos para dos casos importantes de paradigmas científicos: la mecánica y la termodinámica fenomenológica. Que el mismo tipo de análisis se aplique o no a otros casos es una cuestión que quisiera dejar abierta de momento. Admito la posibilidad de que los principios-guía tengan otra forma y función en otros paradigmas científicos. Lo que me interesa subrayar es lo siguiente: si los principios-guía de un paradigma tienen la forma aquí propuesta, eso explicaría algunas de las características peculiares de los paradigmas que Kuhn ha hecho notar y que han intrigado a muchos de sus críticos; fundamentalmente:

a) que el contenido del paradigma sea más una *promesa* de futuros éxitos científicos que una realización palpable;

b) que los paradigmas den lugar a la extraña clase de actividad que Kuhn llama «*resolución de rompecabezas*» («puzzle-solving»);

c) que los paradigmas sean esencialmente *irrefutables* por la experiencia.

Estas características de los paradigmas han sido en parte explicadas por Stegmüller en la obra mencionada, así como en ensayos más recientes (cf., por ejemplo, su artículo «Accidental ("Non-substantial") Theory Change and Theory Dislodgment»). Sin embargo, el análisis de Stegmüller es todavía demasiado crudo, a mi entender, y no muestra claramente el papel que los principios-guía juegan (y *por qué* lo juegan) en los paradigmas. Aquí intentaré complementar el análisis de Stegmüller atendiendo a la *forma* lógica peculiar que poseen los

principios-guía (al menos en los casos de la mecánica clásica y la termodinámica fenomenológica)[10].

La estrategia que seguiré aquí será la de concentrarme en el principio-guía de la mecánica clásica, para llegar, a través del estudio de este ejemplo particular, a las características generales de ese tipo de principios. La razón de esta preferencia es en parte histórica (el principio de Newton ha sido probablemente el más discutido en la historia de la metodología) y en parte expositiva (la mecánica clásica es la teoría física más ampliamente conocida). Al final trataré de hacer plausible que las mismas consideraciones se aplican al caso de la termodinámica.

1. La discusión en torno al Segundo Principio de Newton

El principio-guía de la mecánica clásica es el llamado «Segundo Principio de Newton», o sea, la ecuación dinámica fundamental «$F = m \cdot a$». En las exposiciones «ingenuas» de la mecánica clásica esta ecuación aparece siempre como ley fundamental; es más básica que la «Tercera Ley de Newton», que no es válida para todas las aplicaciones de la teoría (por ejemplo en balística o en el caso de cargas en movimiento), y es lógicamente más básica que la «Primera Ley de Newton», o Principio de Inercia, que se deriva trivialmente del Segundo Principio. Por supuesto, también es más fundamental con respecto a leyes de alcance restringido o más especiales, como la ley de la gravitación, las leyes de Coulomb para fuerzas electrostáticas y magnetostáticas, o las ecuaciones dinámicas para fuerzas elásticas, de fricción, etc. En esta caracterización del Segundo Principio coinciden no sólo los libros de texto «ingenuos», sino también axiomatizaciones concienzudas como las proporcionadas por Hamel, McKinsey-Sugar-Suppes y Sneed, y asimismo las discusiones más filosóficas de Kuhn y Stegmüller. Hay, sin embargo, autores que han negado, o al menos han puesto en duda, el carácter fundamental del Segundo Principio. Sus razones para ello han sido esencialmente lógico-filosóficas. Las discutiremos en seguida.

La historia de la controversia metodológica alrededor del Segundo Principio es larga y densa. Ella es asimismo la historia de la discusión sobre el concepto newtoniano de *fuerza*. En pocos casos un concepto y un enunciado han estado tan indisolublemente ligados como aquí. Para comprender el sentido de esta historia hay que tener bien presente una distinción metodológica que ha jugado un gran papel en

[10] La axiomatización de la mecánica clásica de Sneed-Stegmüller, basada a su vez en la axiomatización «ejemplar» de McKinsey et al., no da cuenta de las características formales del principio-guía correspondiente que aquí nos interesen, y en este sentido es incompleta. Lo mismo vale para mi propia axiomatización de la termodinámica en «A Logical Reconstruction of Simple Equilibrium Thermodynamics».

la moderna filosofía de la ciencia: me refiero a la distinción entre *enunciados de hecho y definiciones* (o convenciones terminológicas). Esta distinción puede considerarse una particularización en el caso de las teorías científicas de la dicotomía más general «*descriptivo-prescriptivo*». Según ella, los enunciados de hecho serían las descripciones de ciertos estados de cosas postulados por la teoría, mientras que las definiciones serían prescripciones acerca del uso de ciertos términos de la teoría.

Asumiendo la dicotomía anterior, podemos preguntarnos si «$F = m \cdot a$» es un enunciado de hecho, es decir, una descripción de situaciones empíricas encontradas en la naturaleza, o bien una definición; si se adopta la segunda alternativa, lo plausible es sostener que se trata de una definición del concepto de *fuerza*, puesto que se admitirá generalmente que los conceptos de espacio, tiempo y masa tienen un significado independiente del de fuerza[11].

Esas dos alternativas de interpretación han constituido el núcleo de la controversia alrededor del Segundo Principio. Unos sostienen que es un enunciado empírico, otros que se trata de una «mera convención», una definición del término «fuerza», y puede decirse sin exagerar que la controversia ha durado siglos. Cada uno de los puntos de vista rivales adquiere cierta plausibilidad por el hecho de que puede mostrarse que la concepción opuesta implica serias dificultades. Así, por ejemplo, alguien que quiera sostener que «$F = m \cdot a$» es un enunciado empírico puede hacer notar que si fuera simplemente una definición, entonces la totalidad de la mecánica clásica estaría basada en una mera tautología y sería, por tanto, una ciencia a priori, por lo menos en su parte fundamental. Quien, por el contrario, quiera persuadirnos de que «$F = m \cdot a$» es una definición de «F» podrá argüir, con razón, que es difícil admitir que dicha ecuación describa un hecho concreto (o siquiera una conjunción de hechos), puesto que cualquier situación empírica imaginable es compatible con ella: siempre pueden construirse vectores-fuerza para cualquier caso concreto, de modo que su resultante sea el producto de la masa por la aceleración de los cuerpos involucrados. El principio, por tanto, es irrefutable. Es más, las fuerzas mecánicas de cualquier sistema no pueden determinarse cuantitativamente como no sea presuponiendo lógicamente la validez del Segundo Principio, como ha mostrado Sneed con detalle en la obra arriba mencionada.

---

[11] En algunas exposiciones de la mecánica clásica que se presentan en cursos introductorios de física se dice, a veces, que la masa «viene definida como cociente de la fuerza por la aceleración». El motivo de tal caracterización es, aparentemente, que en muchos casos se miden masas a través de sus correspondientes pesos en el campo gravitatorio terrestre. Pero esta «definición» de masa no sólo es un exabrupto histórico y metodológico, sino que, además, suele estar en flagrante contradicción con una de las «reglas» más básicas que aprende un estudiante de mecánica en esas mismas exposiciones: «no confundir masa con peso.»

Algunos autores han intentado hallar una solución al dilema del estilo de una «vía media», es decir, ecléctica. Según ella, el Segundo Principio «a veces funcionaría» como un enunciado empírico, «a veces como una definición»; o bien, «desde cierto punto de vista» sería una cosa, «desde otro punto de vista», la otra. De estas posiciones eclécticas o conciliadoras, la mejor articulada que conozco es la de Nagel en *La estructura de la ciencia.* Vale la pena citar a Nagel extensamente. De hecho, su análisis del Segundo Principio de Newton lo lleva a una posición muy cercana a la que es mi *punto de partida:*

... frecuentemente se considera el segundo axioma no como una afirmación acerca de las condiciones en las cuales se producen aceleraciones, sino como una formulación compacta de una *guía* especial para la investigación, como una *regla metodológica* que orienta al físico con respecto a lo que tiene que buscar cuando está analizando los movimientos de los cuerpos. Pues en lo relativo a su característica de no admitir una refutación concluyente, el segundo axioma es muy semejante a una regla. Cualquier número de fracasos del físico para hallar lo que el axioma lo induce a buscar, es insuficiente para concluir de ellos la necesidad de abandonar la búsqueda y descartar la regla. Pues, no obstante esto, la regla puede ser buena porque la investigación conducida de acuerdo con ella puede haber sido recompensada frecuentemente con el éxito y porque aun una regla que sólo es útil a veces puede ser mejor que no tener ninguna regla. De hecho, el segundo axioma, considerado como principio regulador, ha sido sumamente fecundo para guiar la construcción de un cuerpo sistemático de conocimiento bien fundado, y si se continúa aceptándolo como regla de procedimiento, evidentemente no es porque sea una regla arbitraria e infundada para investigar los movimientos de los cuerpos (*op. cit.*, pp. 184-185).

Esta descripción del modo como *se usa* el Segundo Principio nos parece, en términos generales, correcta, por lo que esperaríamos que de este análisis particular se derivara una visión sistemática y general de esta clase de principios físicos que escapara realmente al dilema antes apuntado. Sin embargo, cuando Nagel resume sus conclusiones sobre el análisis de la mecánica newtoniana general, vuelve a poner los «tres axiomas» de Newton en el mismo saco (lo cual, a mi entender, es un grave error) y toma una postura no comprometida y, en último término, decepcionante:

... es evidente que no puede darse ninguna respuesta breve y simple a la pregunta: ¿cuál es el *status* lógico de los axiomas newtonianos del movimiento? ... una respuesta razonablemente satisfactoria a la pregunta indicada exige una referencia al lugar que los axiomas ocupan en alguna codificación particular de la teoría de la mecánica y a los usos que se dan a los axiomas en diversos contextos especiales. Quizá lo que puede afirmarse con toda generalidad es, por una parte, que los axiomas newtonianos a menudo pueden desempeñar el papel de esquemas para analizar los movimientos de los cuerpos o de estipulaciones para definir ciertas nociones experimentales, y, por otra parte, cuando se agregan a los axiomas supuestos adicionales (entre otros, supuestos concernientes a funciones-fuerza) se les puede considerar enunciados que poseen un contenido empírico definido (*op. cit.*, p. 193).

Dos son los defectos principales de esta interpretación, que me parecen totalmente irremediables: primero, que pone al mismo nivel lógico los tres axiomas de Newton; segundo, y más importante, que la apelación a «distintos contextos», «distintos usos», etc., conduce a una ambigüedad y eclecticismo insatisfactorios, que no corresponden de modo alguno a lo que se espera de una verdadera reconstrucción lógica de un aspecto fundamental de la ciencia empírica. En último término, la ambigüedad inherente al análisis de Nagel, a pesar de su prometedor comienzo, proviene de que este autor no ha podido liberarse plenamente de la tajante dicotomía «descripción-prescripción», y por ello, ante un caso como el del Segundo Principio, tiene que concluir que «en algunos contextos» actúa como un enunciado de hecho, mientras que «en otras ocasiones» actúa como una definición. Resumiendo, su punto de partida en el análisis lógico del Segundo Principio es correcto, pero no así su conclusión general.

De estas últimas observaciones puede colegirse ya mi poco interés en mantener en pie la famosa dicotomía. Mi posición general es que hay que desconfiar de las dicotomías sospechosamente simples: las formas lógicas del discurso en general, y *también* las del discurso científico, son más variadas de lo que tienden a suponer los filósofos en sus ansias simplificadoras. En la investigación filosófica, como en la detectivesca, hay que ser a la vez cauto e imaginativo. Aplicando esta actitud general al caso que nos ocupa, y teniendo en cuenta las deficiencias obvias de las dos posiciones controvertidas, así como la superficialidad de la posición ecléctica, creo que es hora de buscar un nuevo enfoque al problema, y, si es posible, un enfoque tal que no sólo sea aplicable al caso particular de «$F = m \cdot a$», sino a otros casos también, con lo que quizá lleguemos a una concepción más general de ciertas estructuras científicas. Esto es precisamente lo que ha faltado a la mayoría de los análisis lógicos del Segundo Principio: su inserción en un esquema metateórico general.

Como se ha hecho notar antes, la discusión en torno al Segundo Principio ha ido normalmente asociada a la discusión del concepto mismo de fuerza. Este concepto es la cruz del problema. Clarificar su naturaleza es condición necesaria para llegar a una concepción adecuada del Segundo Principio.

2. La discusión en torno al concepto de fuerza

Desde el momento mismo de la cristalización de la mecánica newtoniana, el concepto de fuerza fue visto con suspicacia e incluso repugnancia por algunas de las mentes más clarividentes de la historia de la ciencia y la filosofía, en particular (aunque no sólo) debido a su asociación con la idea de «acción a distancia». Entre los contemporáneos de Newton, baste mencionar a Huygens, Leibniz y Berkeley

como críticos agudos del concepto newtoniano de fuerza. Más tarde, y a pesar de los frutos obvios e impresionantes de la mecánica newtoniana, el concepto siguió siendo considerado sospechoso por la mayoría de los autores ocupados en los fundamentos de la mecánica, que querían ver a esta ciencia «libre de oscuridades metafísicas» (según una conocida frase de D'Alembert, en su *Traité de Dynamique*). La principal «oscuridad metafísica» era la fuerza newtoniana (y con ella el Segundo Principio). Así se entienden históricamente los esfuerzos por eliminar el concepto, o al menos someterlo a rígido control, por parte de una serie de autores positivistas o «parapositivistas», desde D'Alembert a mediados del siglo XVIII hasta Hermes y Simon hace treinta años, pasando por Kirchhoff, Mach y Hertz en la segunda mitad del XIX[12]. Todos estos esfuerzos, que en general pueden interpretarse como el intento de reducir la dinámica a la estática o a la cinemática, son ciertamente profundos e interesantes desde un punto de vista metodológico; pero no cambiaron en nada el hecho de que el concepto newtoniano de fuerza siguió campando a sus anchas en todos los libros de texto de mecánica clásica y en la práctica normal de los físicos. Por mucho que famosos físicos como Kirchhoff y agudos lógicos como Herbert Simon hayan sostenido que «fuerza» no es sino una abreviación de «masa por aceleración», y que por consiguiente el Segundo Principio no es sino una tautología, no parece que se les haya hecho mucho caso en la «praxis mecánica». La «comunidad científica» sigue usando el concepto de fuerza como una entidad independiente de masa, espacio y tiempo.

En años relativamente recientes algunas voces filosóficamente ilustradas se han levantado contra el programa de eliminación de fuerzas. Primero fueron Suppes y sus colaboradores quienes, en su ya clásica axiomatización de la mecánica de Newton, mostraron con un argumento formal que el Segundo Principio *no* puede tomarse como definición de fuerza (ni tampoco de masa, claro), a menos que se modifique sustancialmente la estructura de la teoría. Más recientemente un autor versado por igual en filosofía e historia de la ciencia, Clifford Truesdell, ha subrayado expresamente[13] que, a su entender, el concepto newtoniano de fuerza, lejos de ser una superfluidad metafísica, es la contribución más revolucionaria y sustancial de Newton al desarrollo de la mecánica. Según Truesdell, este concepto fue el máximo responsable de todas las grandes consecuciones de la mecánica durante dos siglos; y tal valoración se puede aplicar a la par a su correlato enunciativo, el Segundo Principio. Ambos, concepto y principio, fueron, para hablar en términos kantianos, algo así como «las condiciones de posibilidad de toda mecánica». Tanto Suppes en

[12] Esta historia se relata con detalle en el capítulo 3.5.

[13] Por ejemplo, en su artículo «Rückwirkungen der Geschichte der Mechanik auf die moderne Forschung».

el aspecto formal como Truesdell en el histórico calibran correctamente la importancia del concepto de fuerza y muestran su irreducibilidad en el sentido de que no puede ser definido a partir de otros conceptos supuestamente más básicos ni asociados directamente a medidas y observaciones empíricas. (Por eso Truesdell lo llama un concepto «a priori».) Sin embargo, ni Suppes ni Truesdell insertan su caracterización en un marco conceptual que *explique* (y no sólo describa) la peculiaridad del concepto. Es necesario un análisis más profundo que dé cuenta del papel de «fuerza» en el desarrollo de la mecánica.

Un paso importante en ese camino lo dio Sneed con su reconstrucción de la mecánica clásica en la obra mencionada. Al aplicar su criterio general de teoricidad de conceptos al caso particular de la mecánica newtoniana, Sneed llega a la conclusión de que fuerza y masa son conceptos mecánico-teóricos, mientras que los conceptos cinemáticos serían mecánico-no-teóricos. Las magnitudes *T*-teóricas en general poseen una serie de características peculiares respecto al modo como son usadas dentro de una teoría dada *T* para hacer determinaciones «indirectas» de los valores de las magnitudes *T*-no-teóricas y para establecer correlaciones entre modelos distintos de *T*. Estas características explican, entre otras cosas, la fecundidad predictiva de los conceptos teóricos y su papel en dar un carácter «holista» o global a las teorías empíricas. Así quedarían justificadas algunas de las peculiaridades más notables del concepto de fuerza en la mecánica. Este es ciertamente un resultado relevante para nuestra discusión. Pero no es completo todavía. *No todas* las peculiaridades del concepto de fuerza quedan suficientemente analizadas al subsumirlo bajo la categoría sneediana de magnitud *T*-teórica. Para darse cuenta de ello basta observar que, en el análisis sneediano, fuerza y masa quedan al mismo nivel de conceptos mecánico-teóricos; no habría, pues, una diferencia metodológica de principio entre fuerza y masa. Ahora bien, a pesar de la impresionante fecundidad de la metateoría de Sneed en tantos otros respectos, este punto de su análisis me pareció, desde el principio, insatisfactorio. Mi intuición (y creo que la de muchos otros) sobre la estructura de la mecánica clásica es la de que fuerza y masa no pueden ponerse, por así decir, al mismo nivel. Si la masa es un concepto *T*-teórico, como parece serlo, entonces la fuerza, además de *T*-teórica, es otra cosa; es «más abstracta todavía». Y es un hecho histórico que la mayoría de autores que han examinado la mecánica bajo la lupa crítica se han sentido mucho más preocupados y perplejos por el concepto de fuerza que por el de masa.

Otro punto es digno de mención: la discusión del concepto de fuerza ha ido invariablemente ligada, repetimos una vez más, a la del Segundo Principio, y en especial a su carácter aparentemente tautológico o «aprioristico». Ahora bien, en la reconstrucción sneediana se da una precisión formal del Segundo Principio en la que su aparente carácter tautológico o no-empírico queda reflejado en los siguientes

términos: a pesar de que el Segundo Principio *no* puede tomarse como una definición, su estructura es tal que, en la terminología de Sneed, *cualquier* modelo parcial (no-teórico) puede ser «extendido» o «completado» trivialmente hasta transformarse en un modelo completo (teórico) de la mecánica, satisfaciendo por tanto el Segundo Principio. En este sentido, el Segundo Principio no contiene ninguna restricción de contenido empírico, y por tanto es empíricamente vacuo. Sin embargo, esta última caracterización, por razones que veremos más adelante, es confusa y no refleja el verdadero contenido del principio; anticipamos que éste no es «empíricamente vacuo» en el sentido en que lo son enunciados analíticos como «Ningún soltero está casado» o «Si Juan se come el pollo, entonces el pollo es comido por Juan». En vez de la expresión «empíricamente vacuo» propongo la menos comprometida de «empíricamente irrestricto» para aplicarla al Segundo Principio. Veremos luego por qué la segunda expresión parece más adecuada en este contexto que la primera.

El punto que me interesa hacer resaltar en este momento es el siguiente. En la reconstrucción de Sneed, el hecho de que el concepto de fuerza sólo sea discutible en términos del Segundo Principio y el hecho de que este principio sea empíricamente irrestricto son dos características, por así decir, «contingentes» de la mecánica clásica, casualidades históricas de la forma en que se desarrolló la teoría y, por ende, dos casualidades lógicamente desconectadas entre sí.

Esto último es, a mi entender, el punto débil del análisis sneediano. No es una «casualidad» que el principio fundamental de la mecánica (en cualquiera de sus múltiples formulaciones equivalentes) sea empíricamente irrestricto (sin ser una mera definición), ni es una contingencia histórica que la discusión del concepto de fuerza se haya asociado a la del Segundo Principio. Mi tesis es que: (*a*) es esencial a esta teoría (como a otras seguramente) que su principio-guía sea empíricamente irrestricto; y (*b*) que la peculiaridad de «$F = m \cdot a$» proviene de la estructura implícita de «F» y que es precisamente la estructura de «F» lo que hace al principio tan fecundo, a pesar de ser empíricamente irrestricto y aunque esto suene a paradoja metodológica. En lo que sigue se darán las razones para sostener esta tesis y más adelante se mostrará que una tesis análoga es válida para la termodinámica.

3. Reconstrucción lógica del Segundo Principio de Newton

Empecemos por analizar con más detalle la forma lógica del Segundo Principio. La formalización estándar del mismo es la debida a McKinsey *et al.* en el artículo antes citado[14]. En ella se basan todos

[14] No es fácil conseguir este artículo en su versión original. Existe una traducción castellana hecha por Adolfo G. de la Sienra bajo el título «Fundamentos axiomáticos

los análisis posteriores de la mecánica clásica (excepto los de los «reduccionistas» Simon y Hermes), incluyendo los de Sneed, Stegmüller y los míos propios en trabajos anteriores al presente. En la formalización de McKinsey *et al.*, los conceptos primitivos son los de partícula, posición, instante, masa y fuerza, junto con un índice numérico para distinguir las diferentes fuerzas que pueden actuar sobre una partícula. El sentido de introducir esos índices numéricos es que con este «truco» no es necesario cuantificar sobre las distintas funciones fuerza y por tanto no es necesario introducir variables para funciones. Más adelante veremos, sin embargo, que dicho truco no cumple su propósito.

Si P es un conjunto de partículas (=un sistema mecánico), T es el intervalo temporal durante el que se considera el sistema, i un índice numérico para tipos de fuerzas, s es la función «posición», m la función «masa» y f la función «fuerza», entonces el Segundo Principio en la formalización considerada afirma que, para toda p de P y todo t de T:

$$\sum_i f(p, t, i) = m(p) \cdot D_t^2 s(p, t),$$

donde «$D_t^2 s(p, t)$» es una abreviación para la «derivada segunda de la posición respecto al tiempo», es decir, la aceleración.

Lo que nos interesa subrayar de esta formalización es el hecho de que fuerza y masa aparecen aquí «al mismo nivel» en el siguiente sentido: ambas son funciones de las partículas, es decir, de las variables primitivas del sistema. La única diferencia formal entre masa y fuerza estriba en que la primera es una función solamente monádica y la segunda es triádica, pues tiene dos argumentos más además de las partículas. Pero, en cualquier caso, m y f se hallan al mismo nivel ontológico de funciones de partículas.

Notemos otra característica de la anterior formalización: las únicas variables cuantificables son las variables individuales p, t e i. Por tanto, el Segundo Principio podría ser formalizado en una lógica de primer orden. Lo cual puede parecer una ventaja. Pero en este caso no lo es, porque distorsiona su verdadera forma lógica. Digámoslo de una vez: para explicar la forma lógica del Segundo Principio hay que cuantificar sobre variables *funcionales* y no sólo individuales; es decir, si queremos formalizar necesitamos al menos una lógica de segundo orden. La razón no es meramente que debemos introducir distintas funciones fuerza (pues para ello bastaría el truco del índice numérico),

---

para la mecánica de partículas clásica». Un resumen del mismo (aunque algo confuso para nuestros propósitos) puede encontrarse en la *Introducción a la lógica* de Suppes. También puede acudirse a la reseña del artículo original hecha por C. Lorena García Aguilar para *Crítica* 28, vol. X.

sino que tales funciones fuerza son *funciones de funciones* (o, más correctamente, funciones de tuplos de funciones), y no simplemente funciones de partículas e instantes. En la literatura lógico-matemática[15] se reserva a veces el nombre de «*funcional*» para tales funciones de tuplos de funciones. Adoptando esta denominación, podemos enunciar nuestra tesis así: cada sistema mecánico se describe mediante cierto número de fuerzas y cada una de estas fuerzas es un funcional. (También algunos físicos usan la expresión «funcional» con un sentido idéntico o similar al aquí empleado. Así, por ejemplo, Mittelstaedt, en su texto sobre mecánica clásica, califica a la función fuerza explícitamente como «funcional» de la velocidad, posición, etc.)[16].

¿Por qué decimos que las fuerzas son funcionales y no meramente funciones de partículas? Porque sólo así podemos dar cuenta de una característica esencial de la noción de fuerza, que se revela en su uso en las aplicaciones concretas de la mecánica, pero en cambio se borra en la formulación normal del Segundo Principio: que para cada sistema físico, las fuerzas consideradas (frecuentemente más de una) dependen de una serie de «parámetros adicionales», que sólo se especifican en el momento de la aplicación concreta. Estos parámetros, naturalmente, son a su vez funciones de las variables individuales (básicamente partículas e instantes). El Segundo Principio no especifica ni su número ni su naturaleza, ni tampoco la forma en que la fuerza es función de ellos. Por todo ello, están «implícitos» en la formulación usual del Segundo Principio, lo cual quiere decir, en realidad, que se presupone un número indeterminado de cuantificadores existenciales sobre variables funcionales que representan dichos parámetros.

Esto, por lo demás, corresponde formalmente al modo de hablar intuitivo de los físicos; en efecto, éstos nunca dirían que «la fuerza es una función de partículas e instantes», sino que «distintos tipos de fuerzas son función de distintos parámetros», según las aplicaciones de la teoría. Estos parámetros pueden ser, por ejemplo, coordenadas espaciales, instantes, velocidades, masas, cargas eléctricas, polos magnéticos, coeficientes elásticos, coeficientes de fricción, etc. Todos ellos son, naturalmente, funciones de partículas y/o instantes. Y, en cada caso, la forma específica de las funciones será distinta. Podrá ser, por ejemplo, una función lineal de la distancia o del tiempo, o bien una función cuadrática inversa de la distancia, o bien una función exponencial de la velocidad, etc. Al especificar todas esas diferentes posibilidades en aplicaciones concretas de la mecánica, obtenemos diferentes leyes dinámicas, que pueden ser comprobadas empíricamente. El Segundo Principio contiene implícitamente la información de que *existen* todas esas posibilidades. Pero no dice nada específico acerca de ellas.

---

[15] Cf. por ejemplo, M. Davis, *Computability and Unsolvability*, p. 124.
[16] Cf. P. Mittelstaedt, *Klassische Mechanik*, p. 59.

Si explicitamos esta información contenida en el Segundo Principio, la forma que éste adquiere resulta entonces la siguiente, en términos informales:

«Dado un sistema P durante un intervalo T, puede encontrarse un conjunto de funcionales vectoriales que son funciones de funciones de partículas de P e instantes de T, tales que la suma vectorial de los funcionales es igual al producto de la masa por la aceleración de cada partícula considerada en cada instante considerado.»

Enunciado un poco más formalmente:

(*SPN*) «Dados P y T: Existen n funcionales vectoriales $f_1, \ldots, f_n$ (en $\mathbb{R}^3$) y m funciones (escalares o vectoriales) de $P \times T$, $g_1, \ldots, g_m$, tales que: para cada p en p y t en T se cumple:

$$\sum_{i=1}^{n} f_i(g_1(p, t), \ldots, g_m(p, t)) = m(p) \cdot D_t^2 s(p, t).»$$

Esta enunciación aún no representa enteramente la complejidad de la forma lógica implícita del Segundo Principio; en el capítulo II.5 daremos una versión más exacta y matizada de *SPN*, teniendo en cuenta todas las posibles complicaciones. De todos modos, la formulación presente basta para darse cuenta de que la forma lógica del Segundo Principio es mucho más compleja de lo que se supone en la formulación estándar de McKinsey-Sugar-Suppes. Tal complejidad formal responde al verdadero contenido del principio, como he tratado de hacer ver.

Para ser totalmente exactos y honestos, sin embargo, hay que hacer notar que el contenido implícito del Segundo Principio todavía no queda enteramente cubierto con la formulación que se da aquí (SPN). Esta sólo atiende a la parte sintáctico-semántica del principio, no a lo que podríamos denominar su aspecto pragmático. En efecto, en el uso del SPN en la praxis física está implícito el requerimiento de que los funcionales f_i sean a la vez «no-triviales» y «relativamente simples». Debe excluirse la trivialidad del funcional en el sentido de que éste no sea meramente una forma subrepticia y camuflada de repetir el término derecho de la ecuación, es decir, «masa por aceleración»; en tal caso, SPN no nos guiaría en busca de funciones interesantes, aparte de ser una ingeniosa tautología. Para dar una formulación exacta de este requerimiento sería conveniente hacer uso de la noción formal de «ocurrencia esencial de una variable en una fórmula». Pero aparte de que no veo claro cómo se aplicaría estrictamente esta noción al presente caso (incluso si estuviera clara en otros contextos), probablemente ella por sí sola no inmunizaría el principio frente a todas las posibilidades de «trivialidad» en que piensan los físicos intuitivamente. «Trivialidad en física» seguramente

cubre más casos que «trivialidad en lógica». Por otro lado, la forma matemática de las f_i tampoco debe ser demasiado complicada, hasta el punto de que hiciera el cálculo imposible, en cuyo caso la «guía» ofrecida por SPN sería inútil, no por trivial, sino por desesperante. Ahora bien, mientras no tengamos un criterio formal ampliamente aceptado de *simplicidad* (y no parece que lo tengamos, a pesar de los arduos esfuerzos de Goodman y otros autores), el requerimiento de *no-demasiada-complicación* quedará relegado a la intuición de los físicos (lo cual se suele expresar, a modo de consuelo pedante, diciendo que el concepto de simplicidad es «pragmático»).

Con respecto a los dos problemas que acabamos de discutir (no-trivialidad y no-demasiada-complicación), la presente reconstrucción formal expresada en SPN es todavía incompleta. Sin embargo, no creo que tal incompletud sea fatal para los presentes propósitos. A pesar de ella, creo que la forma lógica presentada es mucho más rica y adecuada que cualquiera de las ofrecidas hasta la fecha para el Segundo Principio. Y sean cuales sean las formalizaciones definitivas de los criterios de no-trivialidad y simplicidad que se adopten en física, presumiblemente serán compatibles con la presente formulación de SPN; en realidad, no puedo imaginar ninguna razón por la que no lo fueran.

Consideremos ahora la forma lógica de SPN con algo más de detenimiento. El aspecto más significativo desde un punto de vista metodológico es la gran cantidad de cuantificadores existenciales involucrados en este enunciado. No sólo se cuantifica sobre las funciones de primer orden g_{ij}, sino también sobre su número i y sobre los funcionales de segundo orden f_i: número y naturaleza de funciones y funcionales se cuantifican existencialmente, es decir, se dejan indeterminados. Por ende, aquí no cabe aplicar el truco de introducir un índice numérico para «condensar» todas las f_i en una sola f, puesto que cada funcional f_i posiblemente dependerá de un conjunto de g_{ij} distinto en cardinalidad y naturaleza.

Ahora bien, es un hecho metodológico bien conocido que cuantos más cuantificadores existenciales se introduzcan en un enunciado empírico, más débil deviene su contenido empírico y más difícil resulta refutarlo. Está claro que es menos comprometido afirmar «ha habido o habrá alguna vez en algún lugar alguna persona con tres piernas» que «Juan Pérez González, nacido el 16 de abril de 1978 en la ciudad de México, tenía tres piernas en el momento de nacer». Si el dominio sobre el que discurren los cuantificadores existenciales es potencialmente infinito, entonces, como bien se sabe, el enunciado en cuestión es irrefutable empíricamente. Llamemos a este tipo de irrefutabilidad «irrefutabilidad de primer orden», puesto que es debida a la cuantificación sobre variables individuales de un dominio. Para la irrefutabilidad de primer orden podemos diseñar estrategias matemáticas y/o empíricas por las que, aun cuando el enunciado es, en principio,

irrefutable, puede hacerse cada vez menos plausible o menos probable hasta resultar «prácticamente refutado». Algunas de estas estrategias pertenecen a la teoría de la decisión. No obstante, la irrefutabilidad de SPN no es de primer orden, pues el problema no estriba en la cuantificación sobre partículas e instantes (estas entidades se hallan cuantificadas universalmente). Se trata de una «irrefutabilidad de segundo orden» (para máximo escándalo a la vez de falsacionistas popperianos y de inductivistas carnapianos), dado que se cuantifica existencialmente sobre variables de segundo orden. Por tanto, SPN no sólo es un principio irrefutable, sino que es incluso difícil imaginar qué clase de estrategia empírica o cálculo inductivo nos podría determinar el grado de inverosimilitud o improbabilidad de SPN. En realidad, la característica indeterminación del principio no proviene de su cuantificación existencial sobre realidades empíricas (cuerpos, lugares, instantes), sino sobre funciones teóricas y sobre funcionales abstractos; es decir, se trata aquí de una cuantificación sobre las potencialidades matemáticas de nuestra mente al tratar con problemas empíricos. Me es difícil imaginar que en tal contexto puedan decirse cosas con sentido, no ya sobre verificación o falsación, sino siquiera sobre grados de corroboración, confirmación, *likelihood*, etc. Todo este aparato metodológico tan popular en la filosofía clásica de la ciencia no parece aplicable aquí.

Podría objetarse que es posible rechazar SPN por su incompatibilidad lógica, no ya con la base empírica, sino con algún otro principio de igual o mayor generalidad que haya sido previamente aceptado por «buenas razones empíricas». Pero también es difícil imaginar, en el caso de SPN, qué otro principio mecánico nos podría hacer dudar de su validez, dado que SPN no afirma prácticamente nada o «casi» nada. Y es notable que la dinámica relativista, a pesar de haber transformado radicalmente los conceptos de espacio, tiempo y masa, haya conservado SPN como ecuación dinámica fundamental *en la misma forma*[17].

De la reconstrucción lógica anterior se desprende también por qué el concepto mismo de fuerza aparece tan indisolublemente ligado a la discusión del Segundo Principio, y por qué es un concepto de interpretación tan elusiva. Para empezar, las fuerzas son funcionales que sólo adquieren sentido dentro del SPN. Además, por la formulación anterior vemos que la idea general de fuerza resume un número indeterminado de funcionales o, mejor dicho, de variables de *tercer* orden para funcionales de distinta forma y contenido. No es de extrañar, pues, que la fuerza newtoniana aparezca en las discusiones metodológicas como un concepto sumamente «abstracto» y difícil de interpretar. Es difícil de interpretar precisamente porque, por su propia naturaleza, no tiene interpretación fija (es el nombre de una

[17] Cf., por ejemplo, S. J. Prokhovnik, *The Logic of Special Relativity*, p. 91.

serie indeterminada de variables funcionales); y parece muy abstracto (y no simplemente *T*-teórico) obviamente porque es un concepto de conceptos (un predicado de predicados).

El Segundo Principio de Newton es un paradigma de paradigmas en un doble sentido: en primer lugar, juega el papel más esencial (el más «paradigmático», si se quiere) dentro del paradigma que llamamos «mecánica clásica»; en segundo lugar, el análisis de la forma y función del Segundo Principio puede resultar metodológicamente paradigmático en el sentido de que en otras teorías pueden aparecer principios-guía de naturaleza semejante y, por tanto, el estudio de las características esenciales del Segundo Principio puede darnos la clave para reconstruir lógicamente esos otros principios-guía y comprender su función dentro de los correspondientes paradigmas. No podemos esperar un paralelismo estricto, pues cada teoría tiene, por así decir, su propia «personalidad». Pero sí es plausible que hallemos ciertos rasgos comunes que sean significativos para un esquema general de la ciencia (por ejemplo, para el esquema de Kuhn). Para sustentar esta tesis propongo el caso de la termodinámica.

4. El principio-guía de la termodinámica reversible

Consideremos primero una sub-teoría de la termodinámica reversible: la termodinámica de los sistemas simples. Los conceptos primitivos de esta teoría son: entropía, energía interna, volumen, presión y moles. En el capítulo 2.6 se presentan las razones de esa elección, así como de que haya que tomar entropía y energía interna como magnitudes termodinámico-teóricas.

Pero aquí lo que nos interesa fundamentalmente es lo siguiente. Siguiendo a Callen en su *Thermodynamics*, podemos identificar el principio básico de esta teoría como la correlación postulada entre la entropía y el resto de funciones extensivas (=aditivas) o, alternativamente, entre la energía y el resto. Callen escribe esas dos «ecuaciones fundamentales» (equivalentes) así:

$$S = S(U, V, N_1 \ldots, N_r);$$

o bien,

$$U = U(S, V, N_1, \ldots, N_r),$$

donde S es la entropía, U la energía interna, V el volumen y N_i son los moles de cada sustancia química.

Esta forma de escribir las ecuaciones fundamentales es naturalmente incorrecta (aunque uno entiende de qué se trata). Porque todas las magnitudes consideradas son funciones *de estado*, es decir, funciones de las variables individuales de la teoría, que representaremos por z. S

no puede ser a la vez función de primer orden (función de estado), tal como aparece a la izquierda de la ecuación, y función de segundo orden (funcional de funciones de estado), tal como aparece a la derecha de la ecuación. Lo mismo vale para U en la segunda ecuación.

Está claro que Callen, como tantos otros físicos despreocupados que escriben este género de ecuaciones, lo que quiere expresar es que los valores numéricos que tome la entropía como función de estado *dependen* de los valores tomados por las otras funciones de estado (las que aparecen a la derecha de la ecuación). Lo análogo vale para la energía. Ahora bien, la manera formalmente correcta de expresar esto es decir que *hay un funcional* (llamémosle f^s) de las funciones extensivas U, V, N_i, tal que el valor tomado por la entropía es igual al valor de dicho funcional (para cada estado z dado). Es decir, existe f^s tal que:

$$S(z)=f^s(U(z),\ V(z),\ N_1(z),\ ...,\ N_r(z))\,.$$

Y análogamente para el principio alternativo existe f^u tal que:

$$U(z)=f^u(S(z),\ V(z),\ N_1(z),\ ...,\ N_r(z))\,.$$

Nótese que f^s, o respectivamente f^u, son funcionales en el mismo sentido en que lo son las funciones fuerza consideradas en el caso de la mecánica. Y, análogamente, este funcional es una variable de segundo orden cuantificada existencialmente, una variable que discurre sobre funcionales de las funciones de estado: energía, volumen y moles. De este funcional la «ecuación fundamental» no nos da más información sino la de que *existe* y toma valores reales. Su forma específica, por lo demás, se deja completamente indeterminada. Al aplicar el principio-guía a casos particulares, se tratará de hallar una forma apropiada de f^s que dé lugar a alguna ley más especial de los sistemas simples (por ejemplo, la ley de Gay-Lussac, o la ley de Van der Waals, o la ley de Stefan). Es sólo a este nivel donde es posible la comprobación empírica en la termodinámica de los sistemas simples. El principio-guía mismo,

$$S(z)=f^s(U(z),\ V(z),\ N_1(z),\ ...,\ N_r(z))\,,$$

es irrefutable experimentalmente.

Nótese finalmente que también en este caso cuantificamos existencialmente no sólo sobre f^s, sino también sobre el número r de funciones moles y, por tanto, en definitiva, también sobre el número de «parámetros» de estado a la derecha de la ecuación. Nuevamente tenemos una doble indeterminación del principio-guía: una parte es debida a la indeterminación en que se deja al funcional y la otra parte

se debe al número de parámetros considerados. No es de extrañar, pues, la «vacuidad empírica» del principio –ni tampoco su fecundidad, por cierto.

De todos modos, puede admitirse que el principio-guía de la termodinámica de los sistemas simples es «menos» irrestricto o vacuo, si se puede hablar así, que el Segundo Principio de Newton. En efecto, aquí sólo se cuantifica existencialmente sobre *un* funcional (f^s), y no sobre un número indeterminado de ellos, y por otro lado se especifican, hasta cierto punto, las funciones de primer orden cubiertas por el funcional.

No obstante, el principio-guía de la termodinámica de los sistemas simples es, en realidad, sólo un «sub-principio-guía», un caso particular de un principio-guía mucho más general que define el paradigma de la termodinámica reversible y que es totalmente comparable en forma y función con el principio-guía de la mecánica clásica[18].

La formulación del principio-guía de toda la termodinámica reversible que propongo a continuación se basa en la dada por Tisza en su *Generalized Thermodynamics*, p. 113, aunque difiere en ciertos detalles[19]. Para comprender bien la formalización que sigue, es necesario recordar que un sistema termodinámico Z en general puede considerarse compuesto de una serie de subsistemas parciales Z_a, Z_b, ..., Z_n (relativamente), simples, y que por tanto un estado z de Z puede considerarse formado por el tuplo de estados parciales $\langle z_a, z_b, ..., z_n \rangle$. Entonces el principio-guía dice lo siguiente:

(EFT) «Para todo sistema termodinámico Z existe una descomposición del mismo en subsistemas Z_a, Z_b, ..., Z_n, tal que para cada subsistema Z_k puede encontrarse un funcional f_k^s y ciertas funciones extensivas de estado G_1^k, G_2^k, ..., G_r^k, tales que, para todo z en Z:

$$S(z) = \sum_k f_k^s(U_k(z), G_1^k(z), ..., G_r^k(z)).»$$

El grado de indeterminación de este principio-guía debido al uso de tantos cuantificadores existenciales es perfectamente comparable al caso del Segundo Principio de Newton. No sólo se cuantifica sobre el funcional f_k^s, sino sobre los parámetros de estado, sobre su número y sobre el número de subsistemas en que descomponemos el sistema original Z. El nivel de «abstracción» de este principio es incluso superior al de SPN, pues mientras que en este último aparecía

[18] Cuando emprendí la reconstrucción lógica de la termodinámica en mi trabajo de 1975, «A Logical Reconstruction of Simple Equilibrium Thermodynamics», no era consciente de la existencia de ese principio más general, y por ello el alcance de mi análisis quedó restringido a los sistemas simples.

[19] La formulación de Tisza es «formalmente despreocupada» (y algo confusa), como suelen ser las de los físicos en tales contextos.

explícitamente al menos la aceleración como magnitud *T*-no-teórica, en el presente caso las únicas funciones que aparecen explicitadas son entropía y energía interna, ambas *T*-teóricas. No es sorprendente que la termodinámica suela considerarse como una teoría más «abstracta y general» que la mecánica clásica.

5. A modo de conclusión: qué promete un paradigma

Partiendo de los análisis anteriores de SPN y EFT podemos retomar el hilo de la discusión general planteada al comienzo de este artículo. Para empezar, ahora está claro por qué la simple dicotomía «descriptivo-prescriptivo» no se puede aplicar en buena ley a los principios-guía considerados. El uso generalizado de cuantificadores existenciales, y en particular la cuantificación de un funcional de segundo orden, nos explica por qué dichos principios no pueden considerarse como descripciones de hechos particulares ni siquiera como una descripción de un conjunto de hechos particulares. Afirmar no sólo que existen ciertas funciones de los objetos básicos de la teoría, sino además que existe un funcional indeterminado de esas funciones, no es describir nada concreto acerca del mundo, a menos que se use el término «descripción» en un sentido abstruso y forzado. A lo sumo podríamos decir que se trata de descripciones acerca de las posibilidades de expresión de nuestro aparato conceptual. Pero éste también sería un sentido extraño de «descripción»; pues, evidentemente, al introducir SPN o EFT no pretendemos meramente hacer una descripción lingüística o psicológica.

Por otro lado tampoco se trata, evidentemente, de definiciones. (Es relativamente fácil mostrar en ambos casos, por el método de Padoa, que ninguno de los conceptos involucrados en ambos principios puede considerarse como un *definiendum* con respecto al resto.) En ninguno de los dos casos se están estableciendo reglas lingüísticas sobre el uso de términos. Por el contrario, ambos principios son considerados por los físicos como matrices conceptuales a partir de las cuales se pueden derivar[20] importantes leyes empíricas sobre la naturaleza.

La posición ecléctica, naturalmente, tampoco funciona. En efecto, no se trata de que SPN o EFT se usen *a veces* como descripciones de hecho y *a veces* como definiciones; ambos principios se usan siempre del mismo modo, precisamente del modo que hemos tratado de explicitar aquí, y que no se puede caracterizar ni como descripción ni como definición, a menos que se deforme totalmente el sentido genuino de esas categorías semánticas.

Finalmente, del análisis propuesto se desprende también en qué sentido los principios-guía analizados son «paradigmáticos» o, dicho

[20] «Derivar» *no* debe entenderse aquí en el sentido de una deducción lógica.

más exactamente, en qué sentido ciertas características de los paradigmas, reveladas en lenguaje metafórico-intuitivo por Kuhn, provienen realmente de la forma lógica de los principios-guía.

En primer lugar, ya hemos visto en qué sentido los principios-guía son empíricamente irrestrictos y, por tanto, irrefutables por la experiencia: introducir un gran número de cuantificadores existenciales en un enunciado empírico es ya una forma de reducir sustancialmente su contenido empírico. Pero si esos cuantificadores se aplican además a variables de *segundo orden*, es patente lo absurdo de «buscar contraejemplos» que refuten el principio. Las posibilidades de dar cuenta de cualquier situación empírica concreta mediante la elección de funciones y funcionales apropiados son, por razones estrictamente lógico-matemáticas, *infinitas supernumerables*.

¿Cuál es el valor de tales principios-guía aparentemente tan vacuos? Su valor es, usando el término de Kuhn, el de una *promesa*, el de un poderoso principio motriz para hacer ciencia. La promesa implícita en el principio-guía consiste en que se nos asegura que si adoptamos el esquema conceptual general propuesto por el principio-guía, a la larga, y con la suficiente paciencia y habilidad, obtendremos los resultados empíricos apetecidos. La promesa es, limitada al principio-guía, muy vaga e indeterminada; pero define la estructura general dentro de la que hay que buscar los parámetros y funcionales requeridos para tratar *cualquier caso particular*. La búsqueda de esos parámetros y funcionales (es decir, desde el punto de vista lógico, la búsqueda de constantes que substituyan adecuadamente las variables funcionales de segundo orden), es la actividad que Kuhn llama «resolución de rompecabezas». Es una actividad ciertamente no-trivial y difícil, a veces exasperante. Pero el principio-guía nos promete que, si perseveramos, a la larga tendremos éxito en nuestra búsqueda y lograremos dar con la substitución adecuada de variables por constantes.

Esa es, efectivamente, la función del Segundo Principio de Newton dentro de la mecánica clásica. Es la promesa de que si buscamos los parámetros dinámicos adecuados y una apropiada relación funcional entre todos ellos, al final obtendremos una ecuación dinámica para cada situación mecánica específica. Esta promesa puede parecer vaga. Pero el hecho histórico es que fue muy estimulante para físicos e ingenieros, generación tras generación, por lo menos durante dos siglos (e incluso hasta nuestros días: la mecánica clásica no está «muerta»). Adoptando el principio-guía expresado en SPN, el paradigma newtoniano permitió hacer lo que el cartesiano había sido incapaz de promover: mostrar el camino para resolver toda clase de problemas mecánicos incluyendo los detalles más complicados. Esa fue la tarea que se propusieron «los newtonianos» que tuvieron fe en la promesa inicial: encontrar el número y naturaleza adecuados de funciones fuerza no sólo para los problemas de la mecánica celeste,

sino para muchos otros tipos de problemas: los de la elastomecánica, la hidrodinámica, la acústica, la electrostática, la magnetostática, e incluso (aunque con menos éxito) problemas ópticos y químicos. Esas fueron las innumerables piezas del rompecabezas newtoniano, cuya resolución condujo a un éxito sin precedentes en la historia de la física.

2.4. REDES TEORICAS

1. Elementos teóricos

En el capítulo 2.2 se ha dado una caracterización intuitiva de la estructura de una teoría empírica según la concepción estructural de Sneed. En particular, se caracterizó allí el llamado «núcleo» de una teoría como una estructura que consta de los modelos determinados por un predicado conjuntista, las condiciones de ligadura entre modelos y la distinción de dos niveles conceptuales. Conviene ahora ser más precisos respecto a estas estructuras, de modo que podamos introducir una noción de teoría empírica lo más exacta posible. Para denotar esta noción más exacta no usaremos el término corriente «teoría», que aparece en la literatura enmarañado en una multiplicidad de usos, sino el de «elemento teórico». La conveniencia de este nuevo término se echará de ver más adelante. Baste decir de momento que, según la concepción presente, «elemento teórico» explicita una gran parte (aunque no la totalidad) de nuestras intuiciones presistemáticas acerca de lo que es una teoría empírica.

Un elemento teórico es una entidad que consta de un núcleo más un conjunto de aplicaciones, a las que una comunidad científica aplica el núcleo durante una etapa histórica determinada. Un elemento teórico es, pues, una entidad que posee no sólo un aspecto formal-semántico (el núcleo), sino también uno pragmático (las aplicaciones), uno sociológico (la comunidad científica) y uno histórico (un intervalo temporal). De los dos últimos componentes apenas trataremos en este capítulo. En el capítulo 3.3 se verá más claramente su relevancia.

Vamos a elucidar cada uno de estos componentes. Para nuestros

propósitos actuales el más importante y complejo es el núcleo. El núcleo es una entidad que consta a su vez de cuatro entidades modeloteóricas: un conjunto de modelos potenciales, un conjunto de modelos (actuales), un conjunto de modelos parciales y un conjunto de condiciones de ligadura.

El conjunto de modelos potenciales (al cual, siguiendo la nomenclatura estándar de Stegmüller, designaremos por M_p) viene determinado por las condiciones estructurales o «axiomas impropios» que intervienen en la caracterización de un predicado conjuntista. Para fijar ideas, en el ejemplo de la mecánica clásica de partículas considerado en 2.2.1 las condiciones de definición (1)-(6) del predicado MCP son los axiomas impropios para ese predicado; a cualquier entidad que satisfaga dichas condiciones la llamaremos «un modelo potencial de la mecánica clásica de partículas» y al conjunto de tales entidades lo podemos designar por $M_p[MCP]$ (=«el conjunto de los modelos potenciales de la mecánica clásica de partículas»). Intuitivamente, un modelo potencial de una teoría dada T es cualquier sistema del que sabemos que tiene la estructura conceptual requerida para ser un modelo de T, aunque posiblemente no sepamos si cumple realmente las leyes empíricas (los «verdaderos axiomas») de T, y por lo tanto no sepamos si es efectivamente un modelo de T. Con las estructuras contenidas en M_p «conceptuamos la realidad» de determinada manera (de la manera correspondiente al lenguaje de T). Pero con ello, naturalmente, aún no queda garantizado que esa conceptuación sea empíricamente válida y fructífera.

Si añadimos uno (o más) axioma(s) propio(s) a las condiciones que determinan los modelos potenciales (y suponemos que este o estos axiomas tienen «contenido empírico» o «fáctico»), entonces definimos un predicado conjuntista con contenido empírico y las entidades que lo satisfacen son los modelos (actuales) de la teoría correspondiente. En el ejemplo considerado, al añadir el segundo principio de Newton obtenemos el predicado MCP, y las entidades que lo satisfagan serán modelos de la mecánica clásica de partículas. Al conjunto de los modelos de una teoría dada lo designamos por M.

Los modelos parciales de una teoría T son las estructuras que describen, mediante conceptos no-teóricos o «empíricos» relativamente a T, los sistemas posibles a los que es concebible aplicar T. Constituyen, por así decir, la «base empírica» de T (en sentido relativo). Su construcción no presupone ni los conceptos específicos ni los principios de T. Generalmente son otras teorías, previas a T, las que proveen los modelos parciales de T. Al conjunto de todos los modelos parciales de una teoría dada lo llamamos M_{pp}. Este conjunto queda determinado por un predicado conjuntista definido por los mismos axiomas estructurales que determinan M_p, exceptuando aquellos que se refieren a los conceptos T-teóricos. En el caso de la mecánica de partículas, esos axiomas serían las condiciones (1)-(4). En

este ejemplo los modelos parciales son, pues, cualesquiera descripciones cinemáticas de sistemas físicos. Los conceptos dinámicos (masa y fuerza) quedan excluidos, puesto que suponemos que son *MCP*-teóricos[21]. Podemos decir, por tanto, que la cinemática es la «base empírica» (relativa) de la mecánica clásica.

El último componente del núcleo son las condiciones de ligadura. En 2.2.2 se han dado ejemplos de ellas para el caso de la mecánica clásica de partículas. Desde un punto de vista formal pertenecen a un nivel distinto, «superior» al de las estructuras que acabamos de mencionar. En efecto, las entidades destinadas a satisfacer las condiciones de ligadura no son modelos individuales, sino conjuntos de modelos. Las ligaduras son, por tanto, no conjuntos de modelos, sino conjuntos de conjuntos de modelos. Dado que, intuitivamente, las ligaduras son relaciones constrictivas entre modelos *distintos*, todos los conjuntos unitarios de modelos estarán contenidos en cualquier ligadura; y también es plausible admitir que si un conjunto de modelos está en una ligadura, también lo estará cualquier subconjunto del mismo, excepto el vacío, del cual no tiene sentido decir que satisface una condición de ligadura.

Podemos dar ahora una elucidación más formal de lo que es un núcleo K de una teoría. Es una estructura $K=\langle M_p, M, M_{pp}, C\rangle$, donde

(1) M_p es un conjunto de modelos potenciales, es decir, de estructuras que satisfacen ciertas condiciones estructurales para ciertos conceptos.

(2) M es un subconjunto de M_p cuyas estructuras, llamadas modelos (actuales) satisfacen además una o más leyes empíricas.

(3) M_{pp} es un conjunto de modelos parciales, es decir, de estructuras que coinciden con las de M_p si en estas últimas suprimimos o «recortamos» los conceptos considerados teóricos (relativamente al caso reconstruido). A cada elemento de M_p le

[21] El problema de determinar, para una teoría T dada, cuáles de sus conceptos son T-teóricos y cuáles T-no-teóricos, es decir, el problema de distinguir las descripciones que son modelos potenciales de las que son modelos parciales, así como la cuestión más general de hallar un criterio lo más eficaz y exacto posible de T-teoricidad, constituyen problemas arduos y complicados que han sido discutidos con detalle por Sneed, Stegmüller, Kamlah, Balzer y por mí mismo en otros escritos. En este libro no se incluye una discusión de estos problemas, en especial del de dar un criterio general de T-teoricidad. Simplemente asumiremos que la distinción entre modelos potenciales y parciales se puede hacer en cada caso concreto de teoría avanzada dentro de las ciencias empíricas y, en consecuencia, que se puede hacer una distinción entre la «base empírica» y la «superestructura teórica» relativas a cada teoría.

corresponde uno y sólo uno de M_{pp}, pero a uno de M_{pp} le pueden corresponder varios de M_p.

(4) C es la intersección conjuntista de conjuntos C_i tales que cumplen:

(a) $C_i \subseteq \mathcal{P}(M_p)$.
(b) $\forall x(x \in M_p \rightarrow \{x\} \in C_i)$.
(c) $\forall X,\ Y \in \mathcal{P}(M_p)$: $X \subseteq Y \wedge Y \in C_i \rightarrow X \in C_i$.
(d) $\varnothing \notin C_i$.

El núcleo es la estructura conceptual de una teoría que se aplica a ciertas porciones de la realidad que el científico se propone sistematizar, explicar y predecir. Para que esto sea posible, es necesario que previamente se hayan conceptualizado dichas porciones de la realidad en términos compatibles con el «lenguaje» del núcleo de la teoría, si bien no en términos que de buenas a primeras presupongan los conceptos y principios específicos de la teoría. Esto significa que para que la aplicación del núcleo tenga lugar, los sistemas empíricos que pretenden explicarse deberán ser conceptuados como estructuras del tipo de los modelos parciales de la teoría, o sea, estructuras que son compatibles con la teoría, pero no la presuponen. Al conjunto de las porciones de la realidad así conceptuadas lo llamamos «dominio de aplicaciones (propuestas)» de la teoría, y lo designamos por A.

En su obra original, Sneed caracterizó el dominio A de la manera más simple y formal posible: como un subconjunto de M_{pp}, que se admite *dado* de una vez por todas (aunque los usuarios de la teoría puedan no conocer su extensión exacta). Sin embargo, esta caracterización de A es insatisfactoria por simplista. Es preciso modificarla en al menos dos sentidos.

En primer lugar no hay que interpretar A simplemente como un conjunto amorfo de aplicaciones heterogéneas, sino como un conjunto de diversas *clases* de aplicaciones, es decir, como un conjunto de conjuntos de aplicaciones. Formalmente, esto significa que A será un subconjunto del conjunto potencia de M_{pp} y no un subconjunto de M_{pp}. Los elementos de A no son aplicaciones aisladas (modelos parciales singulares), sino clases de aplicaciones. Este cambio formal hace más realista la reconstrucción lógica del modo como las teorías se aplican. En efecto, cuando alguien dice que el dominio de aplicaciones exitosas de, pongamos por caso, la teoría de la gravitación de Newton en su formulación originaria fue el conjunto de las órbitas planetarias, casos de caída libre y péndulos, no está presuponiendo que podemos dar una lista de todos los modelos parciales *singulares* que pudo tratar la teoría de Newton; más bien, lo que se quiere decir es que este dominio de aplicaciones consistía en la *clase* de todas las órbitas planetarias conocidas, más la *clase* de todos los casos conocidos de caída libre, más la *clase* de todos los péndulos

construidos, etc. Un caso de aplicación de la teoría es toda una clase de sistemas semejantes en algún sentido pertinente. No importa mucho cuántos elementos tengan de hecho esas clases; normalmente, su número es irrelevante. Por otro lado, el número de *clases diferentes* de aplicaciones contenidas en un dominio dado de aplicaciones ciertamente no es irrelevante. Tener un número mayor o menor de péndulos construidos dentro del campo de aplicaciones de la teoría carece de verdadera importancia (desde el punto de vista teórico); pero, en cambio, es esencial que la clase de los péndulos y la clase de los cuerpos en caída libre aparezcan dentro del mismo dominio. De hecho, una lectura no demasiado literal de los escritos de Sneed y Stegmüller muestra que, cuando hablan de los miembros del dominio de aplicaciones de una teoría, en lo que realmente están pensando es en clases o tipos de aplicaciones, y no simplemente en casos singulares.

Un segundo aspecto problemático de la caracterización de Sneed es que encubre el origen pragmático de la determinación de A. La relación (pragmática, y por tanto fluctuante) entre el usuario de la teoría y las aplicaciones de ésta es inherente a la noción misma de «dominio de aplicaciones de una teoría», por lo que no podemos suponer que A está «dado» de una vez por todas como lo está el núcleo. La concepción sneediana del dominio de aplicaciones tiene indeseables, e innecesarias, connotaciones platónicas: A sería una entidad existente de manera absoluta, con independencia de la determinación pragmática de sus límites. En cambio, parece más realista tomar A como una entidad genidéntica (y en este sentido más parecida a una sociedad que a un sistema formal). El concepto de aplicación propuesta es, si se lo toma en serio, relativo a seres humanos y tiempos históricos. Las aplicaciones propuestas las propone *alguien* (la comunidad científica o algún subgrupo dentro de ella) en un momento dado. En consecuencia, considero que, en una reconstrucción más adecuada del dominio de aplicaciones (propuestas), éste debe ir indisolublemente ligado a dos tipos de entidades que no aparecen en el aparato sneediano original: comunidades científicas (abreviadamente, CC_i) e intervalos históricos (designados por h_i). En la noción de *elemento teórico* aquí propuesta intervienen, por tanto, además de un núcleo K y un dominio de A de aplicaciones, una comunidad científica CC y un intervalo histórico h.

Digamos ahora dos palabras acerca de estas nuevas nociones (las cuales, por su propia naturaleza, contendrán necesariamente un elemento importante de vaguedad).

Respecto a los intervalos históricos, es plausible admitir que la evolución de una teoría puede dividirse en períodos históricos. A cada elemento teórico irá asociado uno de estos períodos: por así decir, su «fecha». Estos períodos podrían determinarse, aunque no es necesario ni siquiera conveniente hacerlo, como intervalos de números reales.

También es plausible admitir que los períodos en cuestión están bien ordenados mediante una relación $\precsim$, que podemos denominar «...es históricamente anterior o simultáneo a...». Se trata de una relación diádica, antisimétrica, transitiva y conexa sobre el conjunto de los $h_i s$. Si H es un conjunto de intervalos históricos considerados h_i, entonces podríamos definir una «*historia*» como una ordenación $\langle H, \precsim \rangle$. Si se determinaran los intervalos históricos mediante intervalos de números reales, podríamos definir «$\precsim$» en términos de la relación «$\leqslant$» sobre los números reales (los cuales representarían instantes); pero ya se ha dicho que tal elucidación no es necesaria ni, probablemente, conveniente. De hecho, parece mucho más adecuado para el tratamiento del tema tomar los intervalos históricos como objetos de naturaleza cualitativa, «borrosa», no identificable con intervalos numéricos.

El uso de esta noción de períodos históricos no es esencial al presente contexto de discusión, dado que aquí sólo nos proponemos analizar la estática de las teorías. No obstante, esta noción devendrá pertinente cuando tratemos de los aspectos diacrónicos de las teorías, especialmente en el capítulo 3.3. por lo que conviene tener ya el aparato formal a nuestra disposición.

El concepto de comunidades científicas, en el sentido presente, se retrotrae a Kuhn. Puede describirse por medio de las siguientes características:

(a) Es un grupo de personas; el grupo como tal normalmente dura más de una generación. (Una comunidad científica es una entidad genidéntica.)

(b) Los miembros del grupo se comunican entre sí en un «lenguaje científico» específico.

(c) Todos ellos participan de técnicas de medición específicas y de procedimientos observacionales y de cálculo para contrastar sus hipótesis.

Puede admitirse que estas tres características están implicadas analíticamente por el concepto de comunidad científica. Puede que, después de las investigaciones pertinentes, resulte que las *CCs* están íntimamente conectadas con otras clases de entidades, tales como «leyes fundamentales» o «ejemplos paradigmáticos»; pero esta conexión debería tomarse como un hecho empírico, no como una cuestión meramente conceptual.

No hay criterios de demarcación exactos para las *CCs*. Son objetos borrosos. Naturalmente, lo mismo vale para otras clases de grupos sociales, y de hecho para la mayoría de nociones sociohistóricas. En la ciencia empírica de la ciencia tenemos que aprender a trabajar con el máximo de precisión posible con conceptos inexactos.

Hemos introducido así los cuatro componentes básicos de un *elemento teórico*. Su definición formal es como sigue.

*D*1: T es un *elemento teórico* si y sólo si existen K, A, CC y h, tales que:

(1) $T=\langle K, A, CC, h\rangle$.
(2) K es un núcleo.
(3) $A\subseteq\mathscr{P}(M_{pp})$.
(4) CC es una comunidad científica.
(5) h es un intervalo histórico.
(6) CC se propone aplicar K a A durante h.

2. Redes teóricas

Los elementos teóricos son, por así decir, las unidades autónomas mínimas de cualquier descripción teórica de la realidad; podríamos decir que son las «células» de una teoría. Un elemento teórico aislado, si se quiere, ya se podría considerar como *una* teoría (a diferencia de un núcleo o de un dominio de aplicaciones, que por sí solos sería impropio considerar como teorías). Por supuesto, las teorías más avanzadas e interesantes no serán «organismos unicelulares» de este tipo, sino que, por lo general, constarán de un gran número de elementos teóricos. Y lo mismo que en el caso de los organismos de la biología, todas esas «células» constituyentes no aparecerán en un conglomerado amorfo, sino firmemente atadas entre sí en una complicada madeja, interconectadas. Las teorías avanzadas se constituyen como redes de elementos teóricos. Las llamaremos simplemente «redes teóricas». Para introducirlas formalmente necesitamos aún un par de nociones auxiliares.

*D*2: Si K y K' son dos núcleos tales que $K=\langle M_p, M_{pp}, M, C\rangle$ y $K'=\langle M'_p, M'_{pp}, M', C'\rangle$, entonces K' es una *especialización nuclear* de K si y sólo si:

(1) $M'_p=M_p$.
(2) $M'_{pp}=M_{pp}$.
(3) $M'\subseteq M$.
(4) $C'\subseteq C$.

*D*3: Si T y T' son elementos teóricos tales que $T=\langle K, A, CC, h\rangle$ y $T'=\langle K', A', CC', h'\rangle$, entonces T' es una *especialización teórica* de T si y sólo si:

La noción de especialización nuclear no contiene ningún componente pragmático. En cambio, la de especialización teórica hace uso de los componentes pragmático considerados anteriormente.

(1) K' es una especialización nuclear de K.
(2) $A' \subseteq A$.
(3) $CC = CC'$.
(4) $h \lesssim h'$.

Debería estar claro después de esta definición que la relación «... es una especialización teórica...» es una relación de orden parcial.

Podemos pasar ahora a definir la noción clave de red teórica.

D4: R es una *red teórica* si y sólo si:

(1) R es un conjunto de elementos teóricos.
(2) R está parcialmente ordenado por la relación de especialización teórica.
(3) Para cada par T_i, T_j, tal que T_i, $T_j \in R$, $CC_i = CC_j$.
(4) Para cada par T_i, T_j, tal que T_i, $T_j \in R$, $h_i = h_j$.

La condición (4) significa que tomamos una red teórica como una estructura «estática»: todos sus elementos teóricos existen durante el mismo intervalo histórico. Esto implica que podemos asociar unívocamente un intervalo histórico h_R con cada red teórica R (es decir, hay una función de las redes teóricas a los intervalos históricos). Asimismo, a cada red teórica R va asociada una comunidad científica CC_R.

Llamemos $Q(R) = \{A / \exists T (T \in R \wedge T = \langle K, A, CC, h \rangle)\}$. A la gran unión $\cup Q(R)$ la llamaremos $I(R)$. Esto es, $I(R)$ cubre la totalidad de dominios de aplicaciones propuestas de los elementos teóricos de una red dada. Podemos denominarla «el dominio de las aplicaciones (propuestas) de la red teórica R».

Normalmente, las redes teóricas de las ciencias empíricas no serán órdenes lineales, es decir, la relación de especialización entre elementos teóricos no será conexa. No parece que los órdenes lineales sean estructuras interesantes para formalizar la estructura de las teorías empíricas. Sin embargo, un tipo de redes teóricas que parece ser particularmente significativo para representar teorías empíricas son las redes teóricas con «primer» elemento teórico. Gráficamente, estas redes teóricas son arbóreas: poseen un tronco común, a partir del cual derivan todas las demás especializaciones como ramas.

D5: R es una *red teórica arbórea* si y sólo si:

(1) R es una red teórica.
(2) Existe un $T_o \in R$, tal que, para todo $T_i \in R$, T_i es una especialización teórica de T_o.

En lo que sigue vamos a llamar a un T_o semejante de una red teórica arbórea el elemento teórico «*básico*» de la red y a su núcleo correspondiente, «núcleo básico».

Es plausible suponer que un gran número (si no la totalidad) de las redes de las teorías empíricas avanzadas son arbóreas. Presentan un «punto de partida» (a saber, las leyes fundamentales y las aplicaciones básicas), a partir del cual se derivan por especialización todos los demás elementos teóricos. Quizá se puedan encontrar contraejemplos a esta hipótesis. No obstante, los ejemplos considerados en este libro son claramente de este tipo, como veremos en seguida.

En 2.2 ya hablamos en términos intuitivos de lo que es el contenido empírico de una teoría según la concepción estructural: es básicamente la aserción que afirma que el dominio de aplicaciones propuestas es subsumible bajo el núcleo. Con el actual concepto de red teórica arbórea podemos desarrollar y precisar más esta idea.

Sea R una red teórica arbórea, $I(R)$ su dominio de aplicaciones propuestas, K_o su núcleo básico y K_i los núcleos que son especializaciones de K_o en R. Sea cualquier $Q \subseteq \mathcal{P}(M_{pp})$, donde M_{pp} corresponde a K_o. Diremos que Q es *subsumible* bajo R si y sólo si Q está constituido por subclases $q_1, \ldots, q_n$, de tal modo que *hay* extensiones teóricas de los elementos de cada q_i tales que, tomadas en su conjunto, satisfacen los axiomas propios y las ligaduras de cierto K_i (para cada q_i).

Sea $\bar{Q}(R) =: \{Q/Q \subseteq \mathcal{P}(M_{pp})$ y Q es subsumible bajo $R\}$. Entonces la aserción empírica de la teoría es que $I(R)$ es alguno de esos $Q \in \bar{Q}(R)$. O sea,

D6: Dada la red teórica R, su *aserción empírica* es: $I(R) \in \bar{Q}(R)$.

Dadas las definiciones y estipulaciones anteriores, debería estar claro ahora por qué se puede considerar a una aserción de tal tipo como un enunciado «holista», sobre todo teniendo en cuenta el supuesto de que R es arbórea.

El lector puede comprobar fácilmente que también se podría definir de manera plausible una aserción empírica para redes teóricas no-arbóreas. Pero dado que éstas no parecen ser interesantes para la reconstrucción de teorías empíricas, prescindimos de tal definición aquí.

2.5. LA RED TEORICA DE LA MECANICA CLASICA DE PARTICULAS

0. Introducción

En este capítulo se emprende una representación de la estructura de la mecánica clásica de partículas como red teórica en el sentido de 2.4. La representación es incompleta en dos sentidos. En primer lugar, dado que de momento sólo nos interesamos por los aspectos sincrónicos de la ciencia, reconstruiremos sólo lo que es la mecánica clásica de partículas *en la actualidad*, tal como aparece en los libros de texto del presente, ignorando por tanto su desarrollo diacrónico. Este último aspecto lo consideraremos en 3.3. En segundo lugar, no se presentarán todos los elementos teóricos que en la actualidad pueden tomarse como especializaciones de la mecánica clásica de partículas y que, por tanto, estrictamente, deberían aparecer en la reconstrucción de su red teórica (por ejemplo, especializaciones que ni siquiera aparecen en los libros de texto usuales en física, sino sólo en los de ingeniería). Ofrecer una lista completa de *todas* las especializaciones de la teoría, además de sumamente tedioso, sería prácticamente imposible si tenemos en cuenta el desarrollo diacrónico de las redes teóricas, el cual determina que, incluso en una teoría aparentemente tan «disecada» como ésta, se estén añadiendo constantemente nuevos elementos. Me limitaré, pues, a la reconstrucción de las líneas de especialización y los elementos teóricos más significativos y corrientemente utilizados en mecánica. Para apoyar la tesis de que las teorías pueden reconstruirse como redes basta una reconstrucción general de este tipo.

También hay que hacer notar que la reconstrucción propuesta se

concentrará en la parte más decisiva (y tratable) de los elementos teóricos, a saber, sus núcleos; a los dominios de aplicación correspondientes sólo me referiré someramente.

La presente reconstrucción de la mecánica clásica de partículas se basa en intentos interiores, sobre todo los de la tradición suppesiana de axiomatización por definición de predicados conjuntistas. Los trabajos básicos en esta tradición son: el de McKinsey *et al.* de 1953, el de Suppes contenido en su *Introducción a la lógica* de 1957 y el de Sneed en su libro de 1971. El primer trabajo, que no es fácil de encontrar en su versión original, ha sido traducido al castellano por Adolfo G. de la Sienra (cf. bibliografía) y analizado detenidamente por Lorena García Aguilar en *Crítica*, núm. 28. Los puntos básicos del enfoque de Suppes han sido expuestos y criticados por Moulines y Sneed en «Suppes' Philosophy of Physics». La reconstrucción sneediana original aparece analizada en diversos escritos de y sobre la concepción estructural de las teorías; quizá el estudio más extenso y asequible es el de Stegmüller en la primera parte de *Theorienstrukturen und Theoriendynamik*.

La presente reconstrucción no pretende ser una alternativa totalmente nueva, sino más bien una complementación y revisión de la tradición mencionada. Sus dos aspectos más novedosos son: el uso explícito y detallado de la noción de red teórica y la nueva concepción de la forma lógica de la ley fundamental de la teoría, de acuerdo con el esquema presentado en 2.3.

Aparte de las axiomatizaciones mencionadas, me he basado, naturalmente, en exposiciones estándar de la mecánica clásica. Sería ocioso dar una lista de títulos pertinentes. No obstante, quisiera mencionar tres libros de texto que han sido particularmente útiles para la tarea propuesta: la *Mechanics* de Symon, la *Klassische Mechanik* de Mittelstaedt y el tomo I de la *Einführung in die Grundlagen der theoretischen Physik* de G. Ludwig.

1. El núcleo básico

Admitiendo los resultados del análisis funcional que ofrece Sneed de los conceptos básicos de la M.C.P., la porción *MCP*-no-teórica de cada modelo estará constituida por la cinemática (fundamentalmente, los conceptos de partícula, instante y posición, más los conceptos derivados de velocidad y aceleración). Los modelos parciales vendrán, pues, sistematizados por el siguiente predicado conjuntista.

*D*1: x es una *cinemática de partículas* (abreviado por: $x \in \widehat{CP}$) si y sólo si existen P, T, s, tales que:

(1) $x = \langle P, T, s \rangle$.
(2) P es un conjunto finito no-vacío.

(3) T es un intervalo de números reales.
(4) s es una función de $P \times T$ en $\mathbb{R}^3$, y s es dos veces diferenciable con respecto al segundo argumento.

La extensión del predicado CP es el conjunto de modelos parciales de esta teoría; esto es: $M_{pp}[MCP] = \widehat{CP}$.

A diferencia de las reconstrucciones anteriores de la tradición Suppes-Sneed-Stegmüller, y de acuerdo con la propuesta hecha en 2.3, la completación MCP-teórica de las cinemáticas de partículas no consistirá simplemente en la adición de dos funciones sencillas, masa y fuerza, sino en la adición de la masa, de una clase $\bar{G}$ de parámetros indeterminados y de una clase $\bar{F}$ de funcionales vectoriales (a los que podemos denominar «funcionales fuerza»), cuyos argumentos son los valores de los parámetros de $\bar{G}$. Nótese que, si bien muchos de los parámetros en $\bar{G}$ pueden considerarse MCP-*no*-teóricos (por ejemplo: distancia, tiempo, velocidad, etc.), la *clase* $\bar{G}$ en sí misma es un concepto MCP-*teórico*, puesto que su introducción es un aspecto específico de la teoría que estamos reconstruyendo: la determinación de esta clase sólo tiene sentido si se presupone que la M.C.P. es aplicable. Naturalmente, aún con más razón, deberemos considerar la clase $\bar{F}$ de funcionales como MCP-teórica. A la completación teórica de CP mediante estos nuevos conceptos la llamaremos una «dinámica de partículas».

D2: x es una *dinámica de partículas* (abreviado por: $x \in \widehat{DP}$) si y sólo si existen P, T, s, m, g_{11}, ..., g_{mn}, f_1, ..., f_m, $\bar{G}$, $\bar{F}$, tales que:

(1) $x = \langle P, T, s, m, \bar{G}, \bar{F} \rangle$.
(2) $\langle P, T, s \rangle \in \widehat{CP}$.
(3) m es una función de P en $\mathbb{R}^+$.
(4) Cada g_{ij}, para $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, es una función de $P \times T \times \mathbb{N}$[22] en $\mathbb{R}$ o $\mathbb{R}^3$.
(5) $\bar{G} = \{g_{11}, ..., g_{mn}\}$.
(6) Para cada f_i $(1 \leqslant i \leqslant m)$ existen $p \in \mathbb{N}$ y $G_i \subseteq \bar{G}$ tales que f_i es una función de $\mathbb{R}^p \times T$ en $\mathbb{R}^3$, y si $r_j \hat{\in} D_I f_i$, entonces existe $g_j \in \bar{G}_i$, tal que $r_j \in D_{II} g_j$.
(7) $\bar{F} = \{f_1, ..., f_m\}$.

Así, pues, $M_p[MCP] = \widehat{DP}$.

Anticipamos aquí que todas las llamadas «constantes dinámicas», incluyendo la constante de gravitación G, que aparecen en leyes dinámicas especiales, se considerarán formalmente como funciones de

[22] En el caso en que una g_{ij} sólo dependa de P, o sólo de T, o sólo de $\mathbb{N}$, consideraremos que es función constante de los demás argumentos. También podríamos caracterizar las g_{ij} como funciones cuyo dominio es $P \times T \times \mathbb{N}$ *o bien* un subproducto cartesiano de éste.

$P \times T \times \mathbb{N}$. Con la excepción de G, todas ellas son «verdaderas» funciones, al menos funciones de $\mathbb{N}$; el caso especial de G puede tomarse como una función constante de $P \times T \times \mathbb{N}$.

El predicado fundamental de M.C.P., que fija sus modelos, vendrá determinado por los axiomas estructurales anteriores más la ley fundamental de M.C.P., a saber, el Segundo Principio de Newton. De una manera informal, ya se ha reconstruido este principio en 2.3. Con el aparato introducido en este capítulo podemos darle una forma más precisa.

*D*3: x es una *dinámica newtoniana de partículas* ($x \in \widehat{DNP}$) si y sólo si existen P, T, s, m, $\bar{G}$, $\bar{F}$, tales que:

(1) $x = \langle P, T, s, m, \bar{G}, \bar{F} \rangle$.
(2) $x \in DP$.
(3) Para cada $p \in P$ y cada $t \in T$ existen $f_{i_1}, \ldots, f_{i_n} \in \bar{F}$ y $r_1^{i_1}, \ldots, r_{p1}^{i_1}, \ldots, r_{pn}^{i_n} \in \mathbb{R}^{(3)}$, tales que:

$$\sum_{h=1}^{n} f_{i_h}(r_1^{i_h}, \ldots, r_{ph}^{i_h}, t) = m(p) \cdot D_t^2 s(p, t).$$

Obtenemos $M[MCP] = \widehat{DNP}$.

La expresión contenida en el miembro izquierdo de la ecuación en (3) suele denominarse la «fuerza resultante» que actúa sobre p en el instante t. El número n que aparece en el sumatorio, es decir, el número de funcionales (de «fuerzas») que intervienen en dicha fuerza resultante, es una característica propia de cada partícula p en cada instante t; cuando queramos subrayar que n es el número fijo de fuerzas correspondientes a p en t, escribiremos $n[p, t]$. Esta expresión se referirá siempre, a partir de ahora, al número de funcionales-fuerza característico del par $\langle p, t \rangle$.

Pasemos ahora a la formulación de las condiciones generales de ligadura del núcleo de M.C.P. Admitimos, siguiendo la estrategia de Sneed-Stegmüller, que las condiciones de ligadura referidas *exclusivamente* a conceptos *MCP*-no-teóricos pertenecen a teorías previas a M.C.P., por lo que no serán incluidas en las ligaduras de M.C.P. Esto es, por supuesto, una hipótesis metateórica que podría ser refutada en algún caso. Por el momento, sin embargo, no tenemos razón alguna para dudar de ella.

Las condiciones de ligadura aquí expuestas constreñirán, pues, m, $\bar{G}$ y $\bar{F}$. Son, esencialmente, las siguientes: las ligaduras de identidad y aditividad para la masa; la de identidad para los elementos de $\bar{F}$; una ligadura peculiar que llamaremos de «combinación» para los conjun-

tos $\bar{G}$ y $\bar{F}$; y la invariancia respecto a las transformaciones de Galileo para m y los elementos de $\bar{F}$[23].

D4: El conjunto $X \subseteq DP$ satisface la ligadura $C_m^=$ (abreviado por $X \in C_m^=$) si y sólo si:

$$\forall x,\ y \in X\ \forall\, p \in P_x \cap P_y\colon\ m_x(p) = m_y(p)^{24}.$$

D5: $X \subseteq \widehat{DP}$ satisface la ligadura $C_m^+ (X \in C_m^+)$ si y sólo si:

$$\forall x,\ y,\ z \in X\ \forall\, p_x \in P_x,\ p_y \in P_y,\ p_z \in P_z\colon$$
$$:p_z = p_x o p_y \rightarrow m_z(p_z) = m_x(p_x) + m_y(p_y)$$

(donde «$p_x o p_y$» denota la concatenación, física o conceptual, de las partículas p_x y p_y).

D6: $X \subseteq \widehat{DP}$ satisface la ligadura $C_{\bar{F}} (X \in C_{\bar{F}})$ si y sólo si:

$$\forall x,\ y \in X,\ \forall f_x \in \bar{F}_x,\ \forall f_y \in \bar{F}_y,\ \forall \langle r_1, \ldots, r_p \rangle \in D_l(f_x) \cap D_l(f_y),$$
$$\forall t \in T_x \cap T_y\colon f_x(r_1, \ldots, r_p, t) = f_y(r_1, \ldots, r_p, t).$$

(Nótese que no es necesario postular una ligadura de identidad análoga para los elementos de $\bar{G}$, puesto que éstos son o bien m, en cuyo caso se aplica $C_m^=$, o bien funciones específicas de teorías distintas de M.C.P., en cuyo caso suponemos que las ligaduras correspondientes ya han sido postuladas.)

D7: Dados $x,\ y \in \widehat{DP}$, designemos por $x \cup y$ el modelo potencial que resulta de la unión lógica de los dominios de definición correspondientes de x, y, es decir, el modelo potencial que tiene por conjunto de partículas $P_x \cup P_y$ y por intervalo temporal $T_x \cup T_y$. Entonces diremos que el conjunto $X \subseteq \widehat{DP}$ satisface la ligadura $C_{\bar{F}}^{\cup} (X \in C_{\bar{F}}^{\cup})$ si y sólo si: $\forall x,\ y \in X\colon \bar{F}_{x \cup y} = \bar{F}_x \cup \bar{F}_y$, y diremos que satisface la ligadura

$C_{\bar{G}}^{\cup} (X \in C_{\bar{G}}^{\cup})$ si y sólo si: $\forall x,\ y \in X\colon \bar{G}_{x \cup y} = \bar{G}_x \cup \bar{G}_y$.

---

[23] Es una cuestión aún controvertida la de si lo más adecuado es incluir las transformaciones de Galileo como ligaduras de M.C.P. o no. Pero como es conveniente dejar abierta esta posibilidad, las incluimos aquí.

[24] Aquí, y en lo que sigue, usamos la convención de que si x es un modelo (parcial, potencial, actual), entonces a_x es la entidad (conjunto, relación, función) del «tipo» a que aparece en x. Así, m_x es la función-masa del modelo x mientras que m_y es la función-masa del modelo y.

Esto significa que la clase de parámetros y la de funcionales dinámicos de un modelo compuesto serán las uniones conjuntistas de las clases respectivas de parámetros y funcionales de los modelos componentes.

A la intersección conjuntista de estas cinco ligaduras $C_m^=$, C_m^+, $C_F^=$, $C_G^\cup$, $C_F^\cup$, la llamaremos la «condición de ligadura básica de M.C.P.» y la designaremos por $C_B[MCP]$. Así, pues, el núcleo básico de M.C.P. es la siguiente estructura: $K_B[MCP]=\langle DP, DNP, CP, C_B[MCP]\rangle$.

2. Especializaciones

Las especializaciones del núcleo básico de esta teoría se construyen especificando sucesiva y paulatinamente las clases $\bar{G}$ y $\bar{F}$, o sea, los parámetros dinámicos g_j y las correlaciones funcionales f_i que los determinan. Aquí se desarrollarán cuatro «líneas» (o «ramas», o «cuerdas») de especialización, cuyas combinaciones pueden dar lugar a especializaciones adicionales. (Las cuatro líneas no son, en principio, incompatibles entre sí.) Cada línea principal de especialización se ramifica, a su vez, en «sublíneas» secundarias, y así sucesivamente. El conjunto interrelacionado de todas estas líneas y sublíneas de especialización constituye la red teórica de la mecánica clásica de partículas. Por supuesto, aquí no se pretende construir un grafo exhaustivo de todas las ramificaciones posibles de la red de M.C.P., sino sólo de las más importantes y frecuentemente utilizadas.

Para dar cuenta claramente de la ordenación jerárquica de las sucesivas especializaciones utilizaremos índices apropiados para los conjuntos de modelos: así, por ejemplo, M_1 es una primera especialización del conjunto básico de modelos M (en este caso: de DNP); M_2 será otra especialización principal de M; M_{21} será una primera subespecialización de M_2; M_{22}, una segunda del mismo conjunto, y así sucesivamente. Dado que en este contexto sólo tratamos de la mecánica clásica de partículas, economizaremos notación, y en vez de escribir $M_2[MCP]$, por ejemplo, escribiremos sencillamente M_2.

Las cuatro grandes líneas de especialización aquí reconstruidas pueden caracterizarse intuitivamente de la siguiente forma:

1) El principio de acción y reacción, primero en su forma más débil y luego en la más fuerte.
2) Fuerzas dependientes de la distancia (o de la posición) con todas sus importantes subespecializaciones hasta llegar a las leyes más concretas, como la ley de la gravitación o la ley de Hooke para materiales particulares.
3) Fuerzas dependientes de la velocidad (como las de fricción o rozamiento).
4) Fuerzas dependientes del tiempo.

Estas cuatro líneas representan las especializaciones dinámicas más corrientes en mecánica; la segunda línea (fuerzas dependientes de la distancia) se trata con especial detalle, puesto que contiene algunas de las leyes más «populares» de la teoría.

Primera línea de especialización: fuerzas de simetría

*D*8: x es una *dinámica de acción y reacción generalizada* ($x \in \widehat{DARG}$) si y sólo si:

(1) $x \in \widehat{DNP}$.

(2) $\forall t \in T \forall p \in P \forall i \leqslant n[P, t]\ \exists! q \in P \exists! j \leqslant n[q, t]$
$\exists g_{i_1}, ..., g_{i_p}, g_{j_1}, ..., g_{j_q} \in \bar{G}$: $q \neq p \wedge$
$f_i(g^i_{i_1}(p, t, i_1), ..., g^i_{i_p}(p, t, i_p), t) \neq 0 \rightarrow$
$f_i(g^i_{i_1}(p, t, i_1), ..., g^i_{i_p}(p, t, i_p), t) =$
$-f_j(g^j_{j_1}(q, t, j_1), ..., g^j_{j_q}(q, t, j_q), t)$.

$M_1 = \widehat{DARG}$.

El requisito de la cuantificación unívoca $\exists! q$ con respecto a *cada* p implica que hay una correspondencia biyectiva entre las partículas del sistema que interactúan de *una forma determinada* (según las componentes de fuerza i, j, respectivamente). A cada partícula p sobre la que se ejerce una fuerza especial le corresponde biunívocamente otra partícula q (necesariamente *distinta*) sobre la que se ejerce el mismo tipo de fuerza, con la misma magnitud, aunque de sentido opuesto. Esta es la explicitación de la versión popular de este axioma: «A cada acción se le opone una reacción igual y opuesta.»

Dada la univocidad de la partícula q y de la fuerza componente j que corresponden a la fuerza i actuante sobre p, es fácil ver que en un sistema $DARG$ se cumple la siguiente propiedad:

*T*1: Si $x \in DARG$, entonces existe una biyección φ tal que $D_I(\varphi) = D_{II}(\varphi) = P_x \times \mathbb{N}$, con la condición $\langle q, j \rangle = \varphi(\langle p, i \rangle) \rightarrow p \neq q$.

La biyección φ expresa la correspondencia entre las «parejas» de partículas interactuantes.

*D*9: x es una *dinámica de acción y reacción estricta* ($x \in \widehat{DARE}$) si y sólo si:

(1) $x \in \widehat{DARG}$.

(2) $\forall t \in T \forall p \in P \forall i \leqslant n[p, t] \exists! q \in P \exists! j \leqslant n[q, t]$
$\exists g_{i_1}, ..., g_{i_p}, g_{j_1}, ..., g_{j_q} \in g$: $q = p \wedge$
$f_i(g^i_{i_1}(p, t, i_1), ..., g^i_{i_p}(p, t, i_p), t) \neq 0 \rightarrow$
$f_i(g^i_{i_1}(p, t, i_1), ..., g^i_{i_p}(p, t, i_p), t) \times s(p, t) =$
$-f_j(g^j_{j_1}(q, t, j_1), ..., g^j_{j_q}(q, t, j_q), t) \times s(q, t)$.

$M_{11} = D\widehat{ARE}$.

Una dinámica de acción y reacción estricta es aquella en que la acción y la reacción no sólo son de magnitud igual y opuesta, sino que además ambas se producen en la misma dirección de la línea que une las partículas interactuantes. Originariamente, Newton creyó que en todo proceso de acción y reacción se cumplía también esta última condición. Pero esto no es cierto: existen sistemas, por ejemplo de cargas eléctricas en movimiento, que son subsumibles bajo *DARG*, pero no bajo *DARE*. El segundo predicado es más restrictivo que el primero. Como condición de ligadura específica de esta línea de especialización es apropiado postular la siguiente:

*D*10: El conjunto $X \subseteq D\widehat{ARG}$ satisface la ligadura $C_{\bar{\varphi}}$ si y sólo si:

$\forall x, x' \in X \forall p, q \in P_x \cap P_{x'} \forall i, j$: $\langle q, j \rangle = \varphi_x(p, i) \rightarrow \langle q, j \rangle = \varphi_{x'}(p, i)$.

Esto significa que la interacción de dos partículas según el principio de acción y reacción es invariante respecto a los modelos en que se consideran dichas partículas.

Segunda línea de especialización:
fuerzas dependientes de la distancia

La manera más sistemática de construir esta importante rama de M.C.P. es a través de la noción de potencial y del uso del operador matemático conocido como «gradiente» (de un potencial). Dado que el concepto de gradiente no está muy divulgado fuera de la física teórica, quizá convenga en este punto dar una breve explicación del mismo. Naturalmente, para una comprensión más detallada del uso y significado de los gradientes, hay que acudir a un libro de texto de mecánica teórica.

Sea f: $\mathbb{R}^3 \rightarrow \mathbb{R}$ una función escalar diferenciable de vectores definidos en un sistema rectangular de coordenadas espaciales, que denotaremos por x, y, z, según es habitual. Sean i, j, k, los tres vectores-unidad que determinan el sistema de coordenadas. Podemos tomar derivadas parciales de f con respecto a cada una de las coordenadas: $D_x f(x, y, z)$, $D_y f(x, y, z)$, $D_z f(x, y, z)$. Si a cada uno de

estos valores (escalares) lo multiplicamos por el vector-unidad correspondiente obtenemos los vectores $i \cdot D_x f(x, y, z)$, $j \cdot D_y f(x, y, z)$, $k \cdot D_z f(x, y, z)$, en cada una de las direcciones espaciales correspondientes. La suma vectorial $i \cdot D_x f(x, y, z) + j \cdot D_y f(x, y, z) + k \cdot D_z f(x, y, z)$ es un vector, cuya dirección, según se puede mostrar, es la misma que aquella en que f varía más. A este vector se le puede considerar como el resultado de aplicar un operador diferencial general $i \cdot D_x + j \cdot D_y + k \cdot D_z$ a la función escalar f; el resultado es un vector que indica en qué dirección varía más f y cuál es la magnitud de dicha variación. A este operador se le llama «gradiente» y lo abreviamos por «grad»:

$$\text{grad } f(x, y, z) = (i \cdot D_x + j \cdot D_y + k \cdot D_z) f(x, y, z) = \\ = i \cdot D_x f(x, y, z) + j \cdot D_y f(x, y, z) + k \cdot D_z f(x, y, z).$$

Nótese que grad es formalmente una expresión abierta (incompleta) del mismo tipo que una derivada «normal» D.

El núcleo de partida para la línea de especializaciones que consideramos aquí consiste simplemente en asumir que la función-posición s forma parte de los parámetros g_i que determinan el funcional fuerza.

*D*11: x es una *dinámica dependiente de la distancia* ($x \in \widehat{DDD}$) si y sólo si:

(1) $x \in \widehat{DNP}$.
(2) $s \in \bar{G}$.

$M_2 = \widehat{DDD}$.

Un tipo importante de fuerzas dependientes de la distancia entre partículas (o, equivalentemente, dependientes de la posición de cada partícula) son las llamadas *fuerzas conservativas.*

*D*12: x es una *dinámica conservativa de partículas* ($x \in \widehat{DCP}$) si y sólo si:

(1) $x \in \widehat{DDD}$.
(2) $t, D_t s \notin \bar{G}$.
(3) $\exists f_i, V \in \bar{F}\ \forall p \in P\ \forall t \in T$: $f_i(r_1^i, \ldots, r_p^i, t) = -\text{grad } V(s(p, t), \ldots)$.

$M_{21} = \widehat{DCP}$.

La función V es la llamada «energía potencial». La condición (3) expresa, pues, que en una dinámica conservativa, al menos una de las «fuerzas» que intervienen es el gradiente negativo de una energía potencial. A tal tipo de fuerza se le suele llamar «fuerza conservativa».

Por la condición (2) se excluye que esa fuerza pueda ser dependiente del tiempo y de la velocidad. En algunos libros de texto (como el de Ludwig, por ejemplo) se excluye de los parámetros determinantes de las fuerzas conservativas no sólo el tiempo y la velocidad, sino también cualquier otro parámetro que no sea la posición (o distancia). En otros, en cambio (como en Symon), se adopta una noción más amplia de fuerza conservativa, según la cual se admite cualquier parámetro que no sea tiempo o velocidad. Esta convención más liberal parece más adecuada, porque permite agrupar bajo la misma categoría los importantes casos de la caída libre, la gravitación y las fuerzas electrostáticas, entre otros.

La siguiente subespecialización que introducimos es una de las leyes más simples de la mecánica: la ley de Hooke. Su campo de aplicación son casos «puros» de oscilación de sistemas elásticos.

*D*13: x es un *oscilador armónico simple* ($x \in \widehat{OAS}$) si y sólo si:

(1) $x \in \widehat{DCP}$.
(2) Existen $a \in \bar{G}$, con $D_I(a) = \mathbb{N}$ y $D_{II}(a) = \mathbb{R}$, y $f_i \in \bar{F}$ tales que:

$$\forall p \in P\, \exists k \in \mathbb{N}\, \forall t \in T\colon f_i(r_1^i, \ldots, r_p^i, t) = -\operatorname{grad} \frac{a(k)}{2} \cdot |s(p, t)|^2.$$

$M_{211} = \widehat{OAS}$.

Admitimos que disponemos de una lista de materiales y que k es el índice correspondiente a uno de ellos. El valor $a(k)$ representa entonces la «constante elástica» característica de ese material particular.

La condición de ligadura correspondiente a la ley de Hooke se refiere precisamente a la invariancia de esa constante de los materiales con respecto al cambio de modelos.

*D*14: El conjunto $X \subseteq \widehat{OAS}$ satisface la ligadura $C_a^=$ si y sólo si:

$$\forall x, x' \in X, \forall k \in \mathbb{N}\colon a_x(k) = a_{x'}(k).$$

Dentro de esta misma línea, otras especializaciones de aplicación completamente distinta son las que se refieren a fenómenos de gravedad y «atracción» entre cuerpos (caída de graves cerca de la superficie de un astro; atracción entre astros; interacción de cargas eléctricas o polos magnéticos, etc.). Primero introducimos la popular ley de Galileo para la caída libre, en una versión generalizada a cualquier astro.

*D*15: x es una *caída libre de partículas* ($x \in \widehat{CLP}$) si y sólo si:

(1) $x \in \widehat{DCP}$.

(2) Existen $g \in \bar{G}$, con $D_I(g) \subseteq \mathbb{R} \times \mathbb{N}$ y $D_{II}(g) \subseteq \mathbb{R}^3$, y $f_i \in \bar{F}$ tales que:

$$\exists r \in \mathbb{R} \exists k \in \mathbb{N} \forall p \in P \, \forall t \in T: f_i(r_1^i, \ldots, r_p^i, t) = \\ = -\text{grad } m(p) \cdot g(r, k) \cdot |s(p, t)|.$$

$M_{212} = \widehat{CLP}$.

g representa la llamada «aceleración de caída»; es sabido que no se trata de una verdadera constante, sino de un parámetro dependiente del planeta (representado por el índice k) y de la latitud del mismo en que tiene lugar la caída (por eso el argumento r). Es de notar que, en esta especialización, tenemos $m \in \bar{G}$, o sea, que la «fuerza conservativa» no sólo depende de la posición.

La condición de ligadura correspondiente atañe, como es de esperar, a la invariancia de la aceleración de caída.

*D*16: El conjunto $X \subseteq \widehat{CLP}$ satisface la ligadura $C_g^=$ si y sólo si:

$$\forall x, x' \in X \, \forall p \in P_x, q \in P_{x'} \, \forall k, k' \in \mathbb{N} \forall t \in T_x, t' \in T_{x'}: \\ r = r' \wedge k = k' \rightarrow g_x(r, k) = g_{x'}(r', k').$$

Quizá la sublínea de especialización más importante y conocida de la mecánica es la constituida por una serie de leyes que postulan una dependencia cuadrático-inversa respecto de la distancia; por supuesto, la más famosa de ellas es la ley de gravitación. Otra que reconstruiremos será la de Coulomb para la electrostática.

*D*17: x es una *dinámica cuadrático-inversa de la distancia* ($x \in \widehat{DCID}$) si y sólo si:

(1) $x \in \widehat{DCP}$.

(2) $\exists f_i, V \in \bar{F} \forall p \in P \forall t \in T: f_i(r_1^i, \ldots, r_p^i, t) = -\text{grad } V$

$$\left(\sum_{\substack{q \in P \\ p \neq q}} \frac{1}{|s(p, t) - s(q, t)|}, \ldots \right).$$

$M_{213} = \widehat{DCID}$.

*D*18: x es una *dinámica gravitacional de partículas* ($x \in \widehat{DGP}$) si y sólo si:

(1) $x \in DCID$.
(2) Existen $G \in \bar{G}$, con $G \in \mathbb{R}^+$, y $f_i \in \bar{F}$ tales que:

$$\forall p \in P \, \forall t \in T: f_i(r_1^i, \ldots, r_p^i, t) = -\text{grad} - G \cdot \sum_{\substack{q \in P \\ p \neq q}} \frac{m(p) \cdot m(q)}{|s(p, t) - s(q, t)|}.$$

G representa la constante de gravitación universal.

$M_{2131} = \widehat{DGP}$.

La condición de ligadura correspondiente a la ley de gravitación es la invariancia de la constante de gravitación.

*D*19: El conjunto $X \subseteq \widehat{DGP}$ satisface la ligadura $C_G^=$ si y sólo si:

$$\forall x, x' \in X : G_x = G_{x'}.$$

Podemos postular, además, otra ligadura especial, que conecta *CLP* con *DGP*.

*D*20: El conjunto $X \subseteq \widehat{CLP} \cup \widehat{DGP}$ satisface la ligadura $C_{g' G}$ si y sólo si:

$$\forall x \in X \, \forall r \in \mathbb{R} \, \forall k \in \mathbb{N}: \text{si } x \in \widehat{CLP}, \text{ entonces existen}$$

$$x' \in \widehat{DGP} \text{ y } p^* \in P_{x'}, \text{ tales que: } P_x \subset P_{x'} \wedge$$
$$p^* \in P_{x'} - P_x \wedge \forall p \in P_x \, \forall t \in T_x \cap T_{x'}:$$
$$g(r, k) = \frac{G_{x'} \cdot m_{x'}(p_{x'}) \cdot (s_{x'}(p, t) - s_{x'}(p^*, t))}{|s_{x'}(p, t) - s_{x'}(p^*, t)|^3}.$$

Es decir, el valor de la aceleración de caída se deriva de la constante de gravitación y de la masa del cuerpo p^* sobre el que se produce la caída.

*D*21: x es una *electrostática de partículas* ($x \in \widehat{ESP}$) si y sólo si:

(1) $x \in DCID$.
(2) Existen Q, $\varepsilon \in \bar{G}$, con $D_I(Q) = P \times T$ y $D_{II}(Q) = \mathbb{R}$, $D_I(\varepsilon) = \mathbb{N}$ y $D_{II}(\varepsilon) = \mathbb{R}$, y $f_i \in \bar{F}$, tales que:

$$\forall p \in P \, \forall t \in T \, \exists k \in \mathbb{N} : f_i(r^i_1, \ldots, r^i_p, t) = -\operatorname{grad} \frac{1}{\varepsilon(k)} \cdot \sum_{\substack{q \in P \\ p \neq q}} \frac{Q(p, t) \cdot Q(q, t)}{|s(p, t) - s(q, t)|}.$$

$M_{2132} = \widehat{ESP}$.

Q es la carga eléctrica de una partícula (la cual puede variar con el tiempo); ε es la llamada «constante dieléctrica», dependiente de un cierto medio k, el cual puede ser distinto para distintas partículas. La ecuación arriba escrita es una reconstrucción de la primera ley de Coulomb.

Las condiciones de ligadura correspondientes a esta subespecialización son: ligaduras de identidad para la carga eléctrica y la «constante» dieléctrica, como es de esperar, más una ligadura de aditividad para Q, puesto que es una magnitud extensiva al igual que la masa.

*D*22: El conjunto $X \subseteq \widehat{ESP}$ satisface las ligaduras $C_{\bar{Q}}$, $C_{\bar{\varepsilon}}$ y C^{+}_{Q} si y sólo si:

(a) $\forall x, x' \in X \, \forall p \in P_x \cap P_{x'} : Q_x(p) = Q_{x'}(p)$.

(b) $\forall x, x' \in X \, \forall k \in \mathbb{N} : \varepsilon_x(k) = \varepsilon_{x'}(k)$.

(c) $\forall x, x', x'' \in X$, $\forall p_x \in P_x$, $p_{x'} \in P_{x'}$, $p_{x''} \in P_{x''}$, $\forall t \in T_x \cap T_{x'} \cap T_{x''}$:
$p_{x''} = p_x \circ p_{x'} \rightarrow Q_{x''}(p_{x''}, t) = Q_x(p_x, t) + Q_{x'}(p_{x'}, t)$.

De una manera enteramente análoga se introduciría la especialización correspondiente a fuerzas magnetostáticas (o sea, la segunda ley de Coulomb), con el predicado «es una magnetostática de partículas» (*MSP*). Bastaría sustituir la noción de carga eléctrica por la de polo magnético. También las condiciones de ligadura correspondientes serían enteramente análogas. Prescindimos de la formulación explícita de esta especialización porque no ofrece ningún aspecto nuevo. El lector puede construirla por sí mismo.

Tercera línea de especialización:
fuerzas dependientes de la velocidad

*D*23: x es una *dinámica dependiente de la velocidad* ($x \in \widehat{DDV}$) si y sólo si:

(1) $x \in \widehat{DNP}$.
(2) $D_t s \in \bar{G}$.

$M = \widehat{DDV}$.

La subespecialización más sencilla e interesante dentro de esta línea es la correspondiente a la ley de fricción simple, aplicable a una partícula que se traslada en el interior de un fluido poco denso (como el aire).

*D*24: x es una *dinámica de fricción simple* ($x \in \widehat{DFS}$) si y sólo si:

(1) $x \in DDV$.
(2) Existen $n \in \mathbb{N}$ y funciones $b \in \bar{G}$, con $D_1(b) = P \times \mathbb{N}$ y $D_{1i}(b) = \mathbb{R}$, y $f_i \in \bar{F}$ tales que:

$$\forall p \in P \, \forall t \in T \exists k \in \mathbb{N} : f_i(r^1_p, \ldots, r^i_p, t) = b(p, k) \cdot D_t s(p, t)^n.$$

$M_{31} = DFS$.

b es un parámetro de fricción, que depende de las características peculiares del medio y de la partícula; el exponente n de la velocidad también depende de las peculiaridades del sistema dinámico considerado.

Es natural introducir una ligadura de identidad para b.

*D*25: El conjunto $X \subseteq \widehat{DFS}$ satisface la ligadura $C^{=}_{b}$ si y sólo si:

$$\forall x, x' \in X, \forall p \in P_x \cap P_{x'} \, \forall k \in \mathbb{N} : b_x(p, k) = b_{x'}(p, k).$$

Cuarta línea de especialización:
fuerzas dependientes del tiempo

*D*26: x es una *dinámica dependiente del tiempo* ($x \in DDT$) si y sólo si:

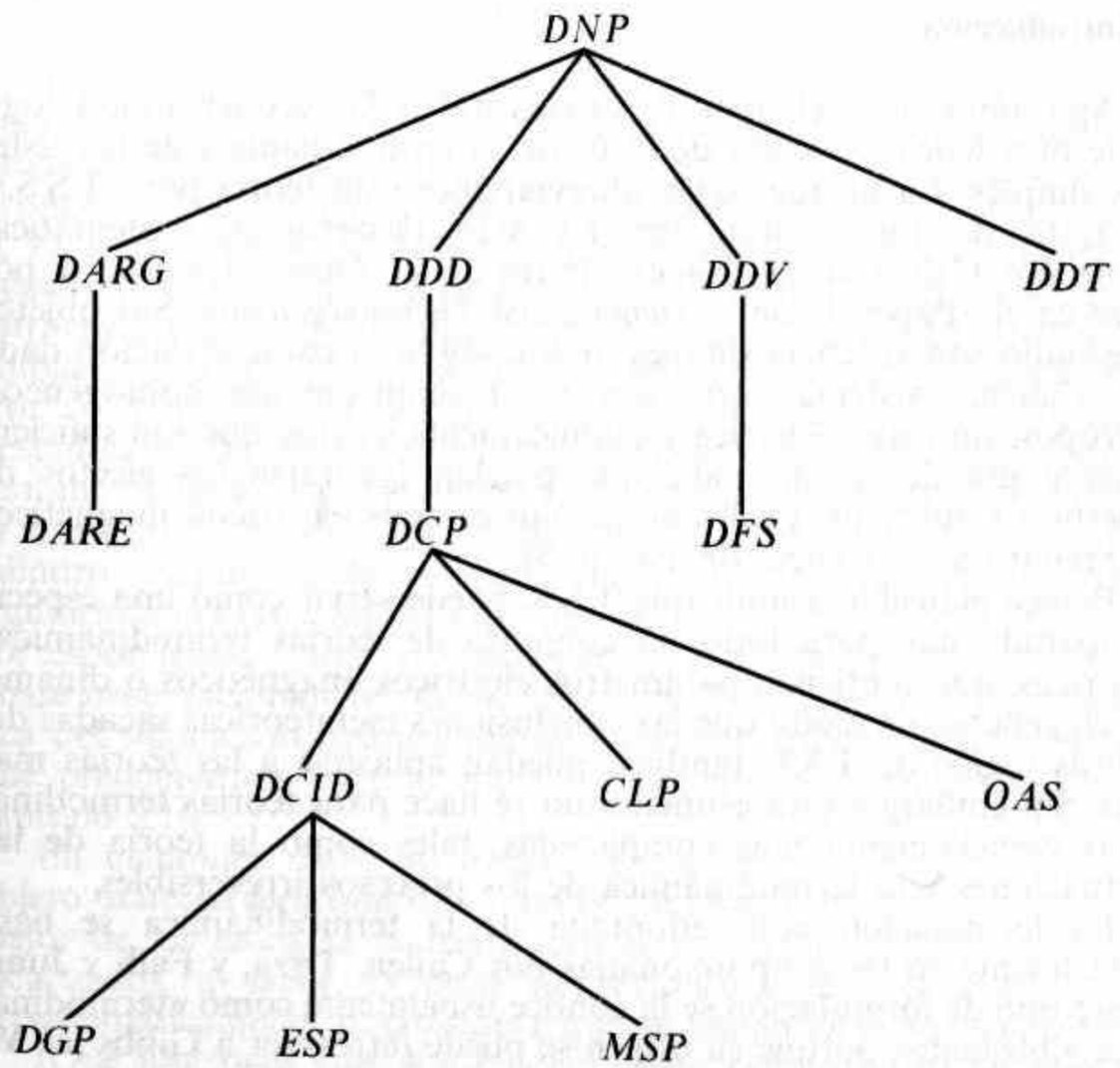
(1) $x \in DNP$.
(2) $t \in \bar{G}$.

$M_4 = \widehat{DDT}$.

La reconstrucción de otras especializaciones, subespecializaciones y combinaciones de los predicados conjuntistas aquí presentados se podría continuar indefinidamente. Para la praxis de la física, y en particular de la física aplicada, todo ese entretejido de innumerables especializaciones es sin duda de gran importancia; para la filosofía de la ciencia, en cambio, el esquema básico aquí expuesto es suficiente para dar una idea de la estructura en red de la mecánica clásica.

Para una mejor visualización de dicha estructura concluimos este capítulo con un gráfico representando la red reconstruida en las definiciones anteriores. Para simplificar la notación sólo se escribirán los nombres de los predicados correspondientes a los modelos de cada elemento teórico, dejando de lado las condiciones de ligadura y los dominios de aplicación. Los trazos indican, de arriba a abajo, especializaciones.

La red de la mecánica clásica de partículas:



2.6. LA RED TEORICA DE LA TERMODINAMICA DE LOS SISTEMAS SIMPLES

0. Introducción

Aplicamos aquí el aparato estructural a la reconstrucción lógica de otra teoría concreta de la física: la termodinámica de los sistemas simples. En lo que sigue, abreviaremos esta teoría por «T.S.S.». Se trata de una teoría que ha sido desarrollada sistemáticamente por Callen en la Primera Parte de su *Thermodynamics* y por Tisza en el «Paper 3» de su *Generalized Thermodynamics*. Sus objetos de estudio son sistemas simples, o sea, según la caracterización dada por Callen, «sistemas que son macroscópicamente homogéneos, isótropos, sin carga eléctrica y químicamente inertes, que son suficientemente grandes como para que puedan ignorarse los efectos de superficie y sobre los cuales no actúan campos eléctricos, magnéticos o gravitatorios» (Callen, *op. cit.*, p. 8).

Parece plausible asumir que T.S.S. puede servir como una especie de «paradigma» para todo un conjunto de teorías termodinámicas más ricas, que contienen parámetros eléctricos, magnéticos o dinámicos en general, de modo que las conclusiones metateóricas sacadas del análisis lógico de T.S.S. también puedan aplicarse a las teorías más ricas. Sin embargo, esta asunción no se hace para teorías termodinámicas esencialmente más complicadas, tales como la teoría de las fluctuaciones y la termodinámica de los procesos irreversibles.

La formulación aquí adoptada de la termodinámica se basa parcialmente en las proporcionadas por Callen, Tisza, y Falk y Jung. A este tipo de formulación se la conoce usualmente como «termodinámica gibbsiana», porque su origen se puede retrotraer a Gibbs (cf. W.

Gibbs, *Collected Works*, t. I). Existe otra tradición en la fundamentación de la termodinámica, a saber, la originada por Carathéodory en su monografía de 1909, en la que trató de reformular las ideas de Kelvin y Clausius de una forma matemáticamente satisfactoria. Esta tradición subsiste hoy día en obras de carácter más o menos sistemático de fundamentos, como en la de Buchdahl de 1966. La relación lógica exacta entre ambas formulaciones no está completamente clara. En principio, uno esperaría que se diera una equivalencia en sentido modelo-teórico, pero parece que la formulación de Carathéodory, incluso si se la precisara adecuadamente, tiene una aplicabilidad más restringida que la termodinámica «neo-gibbsiana». Esta es al menos la conclusión a la que llegan Falk y Jung en su *Axiomatik der Thermodynamik*, y también Tisza (*op. cit.*, p. 50). Naturalmente, la cuestión sólo podría decidirse definitivamente después de un análisis cuidadoso de ambas formulaciones y de sus relaciones interteóricas. Este no es nuestro propósito aquí. De momento, adoptamos simplemente la formulación neo-gibbsiana.

1. Cómo son los modelos de T.S.S.

Vamos a detenernos primero en una descripción informal de la constitución de los modelos de esta teoría. El dominio de cada modelo será un conjunto de *estados*. En principio, y siguiendo la propuesta de Giles en *Mathematical Foundations of Thermodynamics*, p. 93, a tal conjunto lo podemos llamar «un *sistema*». En un capítulo posterior veremos que le podemos dar a la noción de sistema una estructura más compleja que la de un mero conjunto de estados. Sin embargo, en el presente contexto nos basta con la propuesta de Giles. Por lo tanto, en la formulación presente de la termodinámica, «estado» es un término primitivo, mientras que «sistema» es un término definido por el primero. Un ejemplo muy simple de sistema es la sucesión de estados por los que atraviesa un gas contenido en un cilindro cuando se le comprime. Giles ha tratado de revelar el significado operacional del término «estado» dando ciertas «reglas de correspondencia», que supuestamente relacionan la teoría con su «base observacional» (cf. Giles, *op. cit.*, p. 17). Pero toda la problemática del significado operacional de los conceptos físicos no tiene por qué preocuparnos en este contexto. De ella nos ocuparemos en el capítulo 2.10.

En cualquier caso, no debería entenderse un estado como un objeto material *localizado*. Se trata de un objeto «abstracto», que en el mejor de los casos se podría interpretar como una *clase* de objetos localizados, a saber, la clase de regiones espacio-temporales que presistemáticamente se considera que se hallan *en el mismo estado*.

¿Qué funciones vamos a escoger como primitivas en los modelos

del predicado conjuntista de T.S.S.? Los parámetros de la llamada «ecuación fundamental» son, según Callen y Tisza (cf. Callen, *op. cit.*, p. 31; Tisza, *op. cit.*, p. 113), entropía, energía, volumen y los moles de los componentes químicos del sistema. Callen toma temperatura, presión y los potenciales químicos (los llamados «parámetros intensivos») como magnitudes explícitamente definidas en la teoría.

Realmente no veo ningún motivo plausible para dudar de la definibilidad nominal de los potenciales químicos, puesto que este concepto no tiene significado previo a la termodinámica. Pero el caso de la temperatura y de la presión tiene que discutirse más detenidamente. Las conclusiones a las que llegaremos (que son opuestas en uno y otro caso) no están basadas en argumentos formales tajantes, sino más bien en consideraciones de plausibilidad y adecuación. Parece absurdo ir a la búsqueda de «pruebas formales» en esta clase de cuestiones.

En contra de una definición nominal de temperatura dentro de una teoría termodinámica se podría argüir que tenemos una comprensión intuitiva de la temperatura que es previa a la termodinámica. A esta objeción, una primera réplica que probablemente haría un físico es la de que la termodinámica no está interesada en la temperatura «intuitiva», sea ésta lo que fuere, sino en la temperatura *termodinámica*, y esta última se define explícitamente mediante la entropía y la energía (o mediante la entropía y el calor en otras formulaciones de la termodinámica). Si el físico obtiene una contradicción entre la temperatura termodinámica y la intuitiva, mantiene la primera y abandona la segunda.

Es cierto que muchas exposiciones (poco sistemáticas) de la termodinámica empiezan con una discusión de las «propiedades termométricas» de varias substancias (del mercurio, de los gases) antes de que se introduzcan los conceptos de energía y entropía. Tales propiedades termométricas pueden representarse cuantitativamente por medio de una función numérica determinada por una teoría de la metrización derivada (cf., por ejemplo, Stegmüller, *Theorie und Erfahrung*, pp. 61-69). A esta función a veces se la llama «temperatura empírica». Pero no puede *identificarse* con la temperatura termodinámica.

Naturalmente, la «temperatura empírica» (esto es, una magnitud que viene determinada por medio de magnitudes pretermodinámicas tales como presión, volumen o potenciales eléctricos) *coincide* con la temperatura termodinámica (es decir, con la derivada parcial de la energía con respecto a la entropía) en un amplio espectro de aplicaciones de la termodinámica. Por ejemplo, se usa un termómetro de gas bajo condiciones especiales (suponiendo que la presión constante es muy baja y que el punto de condensación está lo bastante alejado) para determinar diferencias de temperatura midiendo las variaciones de volumen del gas en cuestión. Pero de ahí no se sigue

que la temperatura sea *definible* por medio del volumen. Tal inferencia (errónea) sería análoga al argumento presentado por Mach en su artículo inicial de 1868 y resumido en su texto básico sobre la mecánica en favor de una definición cinemática de la masa, basado en el hecho de que las proporciones de las masas de dos cuerpos cualesquiera pueden determinarse mediante las proporciones de las aceleraciones mutuas de los mismos en ciertas aplicaciones de la mecánica[25]. Sneed ha analizado el error lógico de Mach en términos de la metateoría estructural (cf. Sneed, *The Logical Structure of Mathematical Physics*, pp. 59, 135): Mach tomó por definición explícita de la masa lo que sólo es *eliminabilidad tipo Ramsey en sentido débil* de la masa en algunas aplicaciones de la mecánica[26]. Ahora bien, una eliminabilidad tipo Ramsey en sentido débil es una forma de determinación de un concepto mucho más débil que su definición nominal.

Del mismo modo, el hecho de que la temperatura se pueda eliminar tipo Ramsey mediante otras magnitudes como la presión y el volumen en muchas aplicaciones de la termodinámica no implica en absoluto que la temperatura sea definible mediante esas magnitudes en el sentido estricto de «definible». La temperatura queda realmente definida por medio de la entropía y la energía, y luego comprobamos que ese concepto puede eliminarse tipo Ramsey en muchas aplicaciones de la teoría por medio de otras magnitudes.

Estas reflexiones basadas en presentaciones existentes de la termodinámica hacen plausible que: (*a*) el *significado* de la temperatura (termodinámica) no es previo a la termodinámica misma, y (*b*) la temperatura está unívocamente determinada por la energía y la entropía; por lo tanto, la temperatura no es una noción primitiva, sino derivada en esta teoría.

El caso de la presión parece ser algo más claro y opuesto al de la temperatura. La presión se introduce explícitamente como concepto *mecánico*, definido a través de dos funciones (fuerza y superficie) que son previas a la termodinámica. Ciertamente, en algunas presentaciones de la termodinámica (en la de Callen, por ejemplo) podemos leer que «la presión se define como la derivada parcial de la energía respecto del volumen». Sin embargo, en estos casos no se pretende usar el término «definición» en el sentido lógico estricto. El *significado* del concepto de presión se determina en hidrodinámica. Los valores de la presión calculados con el formalismo termodinámico *tienen que* coincidir con los valores de la presión calculados en una teoría hidrodinámica; de lo contrario, suponemos que algo anda mal con la

[25] Debo esta analogía a una conversación con el profesor Sneed.

[26] La noción de eliminabilidad tipo Ramsey a la que aquí me refiero se explica en el libro de Sneed, cap. 3, y en Stegmüller, *Theorienstrukturen und Theoriendynamik*, páginas 71 y ss.

aplicación del formalismo termodinámico. El caso de la temperatura era muy diferente: para la determinación exacta de los valores, la termodinámica siempre tiene prioridad.

Tisza es un poco más cuidadoso que Callen y no habla de una «*definición* termodinámica» de la presión. Sólo dice:

«*Aseveramos* [nótese que no dice: "definimos"] que la intensidad que es la conjugada del volumen es la presión negativa:

$$-P = \left(\frac{\partial U}{\partial V}\right)_{-S,N}$$»;

y añade una observación reveladora:

«Esto se sigue de una especie de *principio de correspondencia* que coordina la relación de *MTE* [es decir, nuestra T.S.S.] con disciplinas vecinas, en particular con la macromecánica del sistema en su totalidad» (cf. Tisza, *op. cit.*, p. 59).

Más adelante analizaremos el *status* de este «principio de correspondencia» con más detalle. De momento basta con tener en cuenta que la presión es una noción cuyo significado queda completamente fijado en una teoría independiente de la termodinámica; de ahí que no parezca adecuado tratar la presión como una función definible termodinámicamente: es un concepto primitivo de la termodinámica.

Si las conclusiones anteriores se aceptan, entonces obtenemos el siguiente conjunto de funciones primitivas en la reconstrucción lógica de la T.S.S. «neo-gibbsiana»: energía, entropía, volumen, moles y presión. Tenemos que preguntarnos ahora cuáles de ellas son *TSS*-teóricas y cuáles *TSS*-no-teóricas.

Como sabemos, el criterio de teoricidad de Sneed, basado en un análisis de los métodos de medición de las funciones, pretende dar una respuesta a esa pregunta para cualquier teoría empírica. Pero, como se puede ver al analizar la situación en una teoría de la complejidad de T.S.S., la aplicación del criterio de Sneed a casos concretos de teorías está plagada de dificultades aún no resueltas. Como solución provisional, propongo aplicar un criterio más burdo, pero también mucho más simple, para decidir si una función dada en una teoría T es T-teórica o no. La idea intuitiva es que la cuestión primordial que hay que plantear no es una cuestión pragmática (sobre *métodos de medición*), sino una cuestión semántica (sobre el *significado* de una función). En términos informales, diremos que f es T-teórica si f no tiene un significado claro previo a T, o en otras palabras, si el concepto de f se aprehende correctamente sólo dentro de T; más brevemente, si f es un concepto *específico* de T. En caso contrario, si f tiene un sentido claro en una teoría T' que puede formularse independientemente de T, diremos que f es T-no-teórica.

Me parece que este «criterio» es una forma ligeramente más desarrollada de la intuición de Putnam: «Un término teórico es uno que proviene de una teoría científica» (cf. Putnam «What Theories Are Not», p. 243). Naturalmente, la propuesta hecha aquí no es formalmente exacta. No obstante, como ya han mostrado los análisis de la presión y la temperatura, parece que este criterio semántico tan intuitivo funciona al menos en cierta medida, sin que nos veamos envueltos en las dificultades características del criterio pragmático de Sneed[27].

Si preguntamos: «¿Es el significado del volumen y los moles específico de la teoría termodinámica?», la respuesta es claramente: «No.» El significado del volumen se fija en la geometría física, el de los moles en la química (estequeometría). La misma respuesta parece ser casi igualmente cierta en el caso de la presión, como he argüido antes.

Por otro lado, parece bastante claro que la entropía no es un concepto que tuviera sentido alguno en un mundo «termodinámicamente ignorante». Históricamente, el concepto de entropía nació con la termodinámica propiamente dicha: es un concepto «específicamente termodinámico», si es que hay alguno que deba ser considerado así.

La decisión parece más difícil de tomar en el caso de la energía. Históricamente, la noción de energía (interna) surgió a partir de una combinación de los conceptos «trabajo» y «calor». Para los sistemas simples, el trabajo es claramente un concepto mecánico, mientras que el calor es no-mecánico. La teoría termodinámica se desarrolló en primer lugar precisamente para dar cuenta del fenómeno «calor». ¿Debemos decir entonces que la energía es parcialmente *TSS*-teórica y parcialmente *TSS*-no-teórica? Esto no tendría ningún sentido en la formulación presente, dado que hemos introducido la energía como *un solo* concepto primitivo (y en ello seguimos la estrategia de la mayoría de las presentaciones cuasi-axiomáticas de la termodinámica). Trabajo y calor pueden introducirse, si se considera conveniente, como conceptos definidos en un estadio posterior.

En su *Thermodynamics* (pp. 15-17), Callen presenta un método para determinar «operacionalmente» la energía en términos de laboratorio. Pero es fácil ver a partir de la discusión que él mismo emprende que la supuesta «mensurabilidad directa» de la energía depende de presupuestos que sólo tienen sentido dentro de la termodinámica. Callen quiere mostrar que, en algunos casos, la energía puede determinarse midiendo únicamente el trabajo mecánico, esto es, a través de un concepto independiente de la termodinámica. Pero esta determinación mecánica de la energía depende de cierto

[27] Si he entendido correctamente la réplica que me hizo el profesor Sneed, parece que él considera que su criterio de los métodos de medición debería ser una *consecuencia* de consideraciones semánticas de un tipo similar a las enunciadas aquí. Sería muy interesante examinar esta relación con más detalle.

número de supuestos cuya formulación exacta sería imposible sin el aparato conceptual termodinámico. Hay que asumir, primero, que algunos sistemas simples pueden encerrarse en paredes *perfectamente adiabáticas*; segundo, que satisfacen una ecuación de la forma

$$(*) \qquad dU = T \cdot dS - P \cdot dV;$$

tercero, que si el proceso se lleva a cabo dentro de las paredes adiabáticas de manera *cuasi-estática*, la variación de la entropía será nula. Por (*) y la condición $dS = 0$, puede determinarse la variación de U midiendo el trabajo de compresión $P \cdot dV$. Pero todo esto significa ni más ni menos que hemos supuesto que las leyes y los conceptos de la termodinámica se aplican al sistema en cuestión. En particular, la ecuación (*) no es sino una forma especial de la ley fundamental de T.S.S., como veremos más adelante. En este caso, el criterio de medición de Sneed puede aplicarse de manera muy plausible y podemos sacar la conclusión general de que *cualquier* determinación de los valores de la energía en un sistema termodinámico se basa en el supuesto específicamente termodinámico de que algo así como la ecuación (*) será válido, y que se pueden aplicar conceptos como «adiabático» y «cuasi-estático». Esto hace plausible la conclusión de que la energía debe considerarse como un concepto termodinámico-teórico.

Por lo tanto, en esta teoría tenemos dos funciones teóricas (entropía y energía) y tres funciones no-teóricas (volumen, moles y presión).

Finalmente, hagamos algunas observaciones sobre una noción muy importante para la termodinámica, la de estado *en equilibrio*. Un sistema termodinámico en T.S.S. es un conjunto de estados que incluye, como subconjunto particular, el conjunto de estados en equilibrio; las leyes de la teoría se aplican solamente a este subconjunto.

A excepción de la rama relativamente reciente de la «termodinámica irreversible», todas las teorías termodinámicas clásicas «tratan» solamente de estados en equilibrio, esto es, sus leyes se enuncian únicamente para estados en equilibrio. Este hecho ha llevado a algunos autores a sostener que los únicos sistemas que se pueden tomar en consideración en la termodinámica «ordinaria» son conjuntos de estados en equilibrio (procesos cuasi-estáticos) (cf., por ejemplo, el epílogo de Max Päsler a las *Vorlesungen über Thermodynamik* de Max Planck, p. 303, o también Callen, *op. cit.*, p. 10). Si se tomara esta idea literalmente, entonces una reconstrucción consecuente de T.S.S. debería tomar el término «estado» como sinónimo de «estado en equilibrio». Curiosamente, incluso autores que parecen tomar este punto de vista en sus observaciones preliminares (como Callen), cuando llega el momento de aplicar los principios generales de la

teoría, asumen más o menos implícitamente la existencia de estados de no-equilibrio. Hacen aseveraciones como: «Si el sistema no estuviera en equilibrio, daría lugar a un proceso tal que...» El *concepto* de no-equilibrio ciertamente no carece de sentido, aunque estemos aplicando la termodinámica del *equilibrio*. Estados de no-equilibrio son inevitables especialmente para una interpretación y aplicación adecuadas del principio extremal de Gibbs y para la determinación de la entropía en sistemas compuestos. La restricción del formalismo a los procesos cuasi-estáticos haría a la teoría inmanejable. Como observa Tisza:

«Si hubiéramos de aceptar esto [es decir, la restricción a procesos cuasi-estáticos] como una limitación de la termostática, seríamos incapaces de formular apropiadamente la propiedad extremal de la entropía, la cual es un elemento indispensable de MTE [=T.S.S.]» (cf. Tisza, *op. cit.*, p. 40).

Tomaremos, pues, el conjunto de los estados en equilibrio como un subconjunto del sistema. En algunas aplicaciones especiales de T.S.S. este subconjunto será un subconjunto impropio; pero esto es sólo un caso límite (procesos cuasi-estáticos).

Otra cuestión que hemos de plantearnos con respecto al concepto de equilibrio es la de si es un concepto *TSS*-teórico o *TSS*-no-teórico. Quizá se estaría dispuesto a creer que es muy fácil decidir si un sistema está en equilibrio o no sin saber nada sobre termodinámica; la cuestión se podría resolver por medios «puramente observacionales»: el sistema está en equilibrio si y sólo si está quieto. Pero esto sería un punto de vista muy ingenuo. Por lo general, no podemos estar seguros de que el sistema esté en equilibrio si no aplicamos la teoría misma. Nuevamente podemos citar una observación pertinente hecha por Tisza:

«Entre los conceptos más importantes de la termodinámica está el de equilibrio, un concepto elusivo, puesto que no existen medios puramente observacionales para decidir si un sistema aparentemente quieto ha alcanzado realmente un equilibrio, o bien se halla varado en un estado de no-equilibrio, aunque derivando imperceptiblemente hacia el equilibrio» (cf. Tisza, *op. cit.*, p. 13).

Y Callen hace notar con respecto a la determinación empírica del equilibrio:

«Si un sistema que no se halla en equilibrio se analiza sobre la base de un formalismo termodinamico bajo el supuesto de equilibrio, aparecen contradicciones en el formalismo, y los resultados predichos difieren de las observaciones experimentales. Este fracaso de la teoría lo usa el físico experimental como criterio *a posteriori* para la detección de estados de no-equilibrio» (cf. Callen, *op. cit.*, p. 13).

Así tenemos un buen ejemplo de un concepto que no es una función (por lo que no tiene que ser medido), y cuya determinación, no obstante, presupone la aplicabilidad de la teoría. Podemos concluir de ahí que el concepto de equilibrio es un concepto primitivo y teórico en T.S.S.

2. El núcleo básico

Después de las conclusiones a las que hemos llegado sobre la composición de los modelos de T.S.S., podemos pasar ya a la definición formal de los modelos parciales, de los modelos potenciales y de los modelos actuales de esta teoría. (Recuérdese que los conceptos TSS-teóricos son los de equilibrio, energía y entropía.)

*D*1: x es un *proceso simple* ($x \in \widehat{PS}$) si y sólo si: existen Z, I, V, N, P, tales que:

(1) $x = \langle Z, I, V, N, P \rangle$.

(2) Z es una variedad diferenciable acotada unidimensional.

(3) I es un subconjunto finito no-vacío de los números naturales.

(4) V es una función de Z en $\mathbb{R}^+$ y V es dos veces continuamente diferenciable.

(5) N es una función de $I \times Z$ en $\mathbb{R}^+$ y N es dos veces continuamente diferenciable con respecto a su segundo argumento.

(6) P es una función de un subconjunto de Z en $\mathbb{R}^+$ y P es dos veces continuamente diferenciable.

$M_{pp}[TSS] = \widehat{PS}$.

Z es el dominio de individuos de T.S.S.: una secuencia de estados. Al ser Z una variedad diferenciable unidimensional, se desprende que Z es biyectable con un intervalo de $\mathbb{R}$. Por lo tanto, podemos ordenar Z de manera correspondiente a como $\mathbb{R}$ está ordenado por $<$. A esta relación de orden inducida entre los estados la podemos designar por « $\prec$» y leer «$z \prec z'$» como «z es anterior a z'».

El requisito de que Z sea una variedad diferenciable unidimensional implica ya una restricción importante a su posible interpretación empírica. Formalmente, necesitamos este requisito tan fuerte para garantizar la diferenciabilidad de las funciones definidas sobre Z. Una desventaja de este postulado, por otra parte, es que dentro de la formulación presente de T.S.S. no podemos decir que el mismo estado aparece dos veces en el proceso representado por Z. No se admiten «puntos dobles» en el continuo Z. Esto puede parecer intuitivamente

inadecuado para la termodinámica, puesto que estamos habituados a la idea de que un mismo sistema puede atravesar por un proceso cíclico después del cual vuelve al estado inicial. No obstante, la posibilidad de usar una multiplicidad de modelos admitida en nuestra noción de teoría compensa este defecto. Podemos modificar nuestro modo de hablar sin que cambie nada sustancialmente: en vez de decir que un sistema dado sufre un proceso cíclico, podemos decir que el proceso cíclico consiste en realidad en dos procesos (=modelos) que poseen el mismo estado «repetido». Nada nos impide hablar así. Naturalmente, los dos procesos que constituyen el proceso cíclico total considerado estarán conectados de una manera natural mediante ligaduras apropiadas.

A pesar de las repetidas aseveraciones de muchos filósofos y científicos de que espacio y tiempo pertenecen necesariamente al marco conceptual de *cualquier* teoría física, aquí tenemos un ejemplo de una teoría física en la que no parece necesario introducir espacio y tiempo entre sus conceptos básicos. Los únicos individuos asumidos aquí son *estados*, y los estados son objetos abstractos en el sentido de no-localizados espaciotemporalmente, como ya hemos argüido. (A lo sumo podrían considerarse como clases de equivalencia de objetos localizados espaciotemporalmente; pero ni siquiera esto es necesario para una formulación elegante de la termodinámica.) En contra de este punto de vista podrían plantearse dos objeciones. Primera, se podría argüir que la sucesión de estados representada por Z *es* de hecho una sucesión temporal; segunda, que una de las funciones básicas de la teoría es el volumen V, el cual *es* un concepto «espacial». A estas objeciones se puede replicar lo siguiente. Puede servir de ayuda intuitiva para la aplicación de la termodinámica, pero no es necesario en absoluto interpretar la sucesión Z de estados como algo que «ocurre en el tiempo»; dicho más exactamente, para la aplicación de la termodinámica no necesitamos una medida cronológica sobre Z. Por lo que respecta al volumen, es cierto que se trata de un concepto tomado de la geometría física, y en *este* sentido presuponemos una noción espacial. Pero no es en este sentido débil e inofensivo que se suele decir que el espacio entra esencialmente en la base conceptual de cualquier teoría física. Lo que se quiere decir normalmente con ello es que las entidades básicas de la teoría se supone que «ocupan algún lugar en alguna parte», que están localizadas en el espacio físico (en el «cosmos», por así decir). Esto es lo que no es cierto de la teoría termodinámica. Para asignar un volumen a un estado no es necesario presuponer que está localizado en alguna parte. Con otras palabras, para formular la termodinámica no es necesario presuponer que disponemos ni de relojes ni de marcos espaciales de referencia. Claro que aún se podría argüir que, sean cuales sean las alternativas formalmente posibles de reconstrucción de una teoría física, los supuestos ontológicos universales de la física son tales que la plena

estructura de espacio y tiempo *debe* entrar en cualquier reconstrucción, incluso a riesgo de aparecer *ad hoc*. Esta posición es manifiestamente una muestra de metafísica apriorística sobre la ciencia. Puede que venga apoyada por autoridades tales como Kant. Pero no veo ninguna necesidad para sustentarla en este caso concreto.

I es un conjunto-índice, es decir, un conjunto de números naturales que se toman como índices de una lista de objetos, en este caso, sustancias químicas. La función moles N está definida sobre índices y estados, indicándonos así cuántos moles de cada sustancia hay en cada estado: «$N(i, z)=r$» significa que el sistema posee r moles de la sustancia i en el estado z. Normalmente, en vez de una única función moles diádica N, se introduce un conjunto indeterminado de funciones moles distintas N_1, ..., N_n (una para cada sustancia química que constituye el sistema). Sin embargo, es más expeditivo «resumir» toda la serie de n funciones N_1, ..., N_n en una única función N que esté definida no sólo sobre estados, sino también sobre índices. (Algo parecido hicimos con los parámetros g_{ij} en la M.C.P.). En cuanto a los valores que toma N, en rigor deberían ser sólo números *racionales* y no reales: los números molares se definen químicamente como el número de moléculas de la sustancia en cuestión que están presentes dividido por el número de Avogadro ($6{,}024 \cdot 10^{23}$); es obvio que este valor nunca puede ser irracional. Sin embargo, las exposiciones sistemáticas de la termodinámica admiten la útil idealización de que las funciones moles sean continuas y diferenciables. En aplicaciones «normales» de la termodinámica esta idealización no es excesivamente cruda, ya que el número de moléculas presentes en el sistema es muy grande.

El dominio de la presión P es un subconjunto del conjunto total de estados; intuitivamente, contiene todos los estados con presión definida o, como a veces se dice, estados con «equilibrio de presión». Este subconjunto puede o no coincidir con la totalidad de Z.

La extensión teórica de los modelos parciales de T.S.S. se obtiene añadiéndoles el concepto de equilibrio (total) y las funciones energía y entropía.

D2: x es un *proceso termodinámico simple* ($x \in \widehat{PTS}$) si y sólo si existen Z, I, V, N, P, Z^e, U, S, tales que:

(1) $x=\langle Z, I, V, N, P, Z^e, U, S\rangle$.

(2) $\langle Z, I, V, N, P\rangle \in \widehat{PS}$.

(3) Z^e es un subconjunto finito no-vacío de $\mathscr{P}(Z)$ y, para todo $e \in Z^e$, e es un abierto no-vacío.

(4) U es una función de Z en $\mathbb{R}$ y U es dos veces continuamente diferenciable.

(5) S es una función de Z en $\mathbb{R}$ y S es dos veces continuamente diferenciable.

$M_p[TSS]=PTS$.

En los modelos, U representará la energía interna y S la entropía del sistema. Z^e representa el conjunto de todas las secuencias de estados en Z que se hallan en *equilibrio* (total) (con respecto al formalismo de T.S.S.). Esto es, el proceso entero representado por Z puede atravesar alternativamente, un número finito de veces, fases de equilibrio y de no-equilibrio. Se exige que haya al menos una fase de equilibrio. Se admite que cada fase de equilibrio del sistema es un subproceso del proceso total Z. Es decir, no se admiten puntos de singularidad entre los estados de equilibrio. Esto se desprende del requisito de que cada $e\in Z^e$ sea un conjunto *abierto*. Esta condición es necesaria para garantizar la aplicabilidad de las técnicas de diferenciación a los estados de equilibrio. Desde un punto de vista físico no significa ninguna restricción grave del alcance de la teoría. Sólo nos interesan situaciones de equilibrio que, por así decir, «duran un rato».

El conjunto de todos los estados de equilibrio viene dado por $\cup Z^e$. Dado que vamos a usar este conjunto con bastante frecuencia, es conveniente darle un nombre especial. Llamémosle E.

Al ser el conjunto básico Z una variedad diferenciable unidimensional y cada $e\in Z^e$ un abierto en Z, ambos serán isomorfos a intervalos de números reales. Esto significa que los elementos de Z (así como los de cada e) vendrán ordenados del mismo modo en que los números reales vienen ordenados por la relación $<$. Vamos a denotar la relación correspondiente entre estados por «$\prec$». Podemos leer «$z\prec z'$» como «el estado z' es posterior al estado z». Esto es simplemente un modo de hablar.

Para introducir el siguiente predicado conjuntista necesitamos las siguientes nociones técnicas. Diremos que una función real f que tiene números por argumentos es *monótona*, respectivamente *estrictamente monótona* con respecto a sus argumentos siempre que se cumpla $f(x)\leqslant f(x')$, respectivamente $f(x)<f(x')$ para $x<x'$. En caso de que f no sea una función de un solo argumento, sino función de un tuplo $\langle x_1, ..., x_n\rangle$, convenimos en decir que f es monótona, respectivamente estrictamente monótona con respecto a cada uno de los x_i siempre que se cumpla $f(x_1, ..., x_i, ..., x_n)\leqslant f(x'_1, ..., x'_i, ..., x'_n)$, respectivamente $f(x_1, ..., x_i, ..., x_n)<f(x'_1, ..., x'_i, ..., x'_n)$ para $x_i<x'_i$ *bajo el supuesto* de que $x_j=x'_j$ para todo $j\neq i$. Por extensión, podemos decir que una función real f es monótona con respecto a otra función g siempre que se cumpla $f(x)\leqslant f(x')$, respectivamente $f(x)<f(x')$ para $g(x)<g(x')$.

D3: x es una *termodinámica gibbsiana simple* ($x\in\widehat{TGS}$) si y sólo si existen Z, I, V, N, P, Z^e, U, S, f^S, tales que:

(1) $x=\langle Z, I, V, N, P, Z^e, U, S\rangle$.

(2) $x\in\widehat{PTS}$.

(3) f^S es una función de $\mathbb{R}^{|I|+2}$ en $\mathbb{R}$, y f^S es dos veces continuamente diferenciable, y f^S es estrictamente monótona con respecto a cada uno de sus argumentos, y S es débilmente monótona con respecto a la composición funcional de f^S con U, V, N, y $\forall z(z\in E\leftrightarrow S(z)=f^S(U(z), V(z), N(i_1, z), ..., N(i_{|I|}, z)))$.

(4) $\forall z(z\in E\rightarrow P(z)\cdot D_U f^S(z)=D_V f^S(z))$.

$M[\text{T.S.S.}]=\widehat{TGS}$.

El bicondicional de la condición (3) de la definición anterior expresa la ley fundamental de T.S.S. Formaliza la correlación fundamental entre la entropía y el resto de los parámetros termodinámicos: en cada estado de equilibrio, la entropía viene determinada, a través de un funcional f^S no-especificado, por la energía, el volumen y el número de moles de cada sustancia. En los libros de texto corrientes esta correlación funcional se escribe en forma de la ecuación «$S=S(U, V, N_i)$», la cual, desde un punto de vista estrictamente formal, es un absurdo, dado que la entropía no puede ser a la vez una función monádica de estado y un funcional de otras funciones de estado. Lo que se quiere decir, naturalmente, es que los valores que toma la entropía (en los estados de equilibrio) dependen de los valores que toman la energía, el volumen y el número de moles. En nuestra formulación, esta dependencia viene dada por el funcional f^S. A esta función –cuyos argumentos no son estados, sino tuplos de números reales– la llamamos una «determinación entrópica». Podríamos decir que la entropía viene determinada por los valores de la energía, etc., «al modo f^S». La forma particular de f^S puede ser diferente para diferentes clases de sistemas (veremos algunos ejemplos en lo que sigue), pero es siempre la misma para *un* sistema dado. Las formas particulares de f^S se especifican mediante leyes especiales de la teoría. La condición (o ley) *D*3-(3), tal como se enuncia en el núcleo de T.S.S., no especifica nada. Es un «principio-guía». De un modo análogo al del «Segundo Principio» de Newton en M.C.P., esta ley de marco termodinámica no es «falseable». Dado un sistema simple, tratamos de aplicarle la teoría con una f^S específica; si no obtenemos los resultados deseados, no decimos que la teoría ha sido falseada: simplemente hacemos el intento con otra forma específica de determinación entrópica. La ley fundamental sólo nos dice que existe *alguna* forma, no cuál.

Es de notar la importante asimetría que existe entre la ley expresada en la condición *D*3-(3) y la expresada en la condición *D*3-(4). Desde un punto de vista formal, vemos que el último axioma

incluye solamente tres conceptos primitivos de la teoría: energía, volumen y presión. A esta relación tripartita frecuentemente se la considera como una «definición termodinámica» de la presión. Pero ya hemos argüido antes que no es conveniente tomar esta «definición» en el sentido del físico como una definición genuina en el sentido lógico. La presión P debe tomarse como un concepto primitivo en T.S.S., y además como uno no-teórico, porque es una magnitud específicamente *mecánica*. Podemos identificar algunos valores medidos en un sistema físico como valores de la presión de ese sistema, incluso si no sabemos nada de termodinámica, ni nada sobre la relación entre la presión por un lado y la energía y el volumen por otro. Esta relación es una *ley empírica*, no una definición nominal. Y es sin duda una ley general de la termodinámica: si un sistema no la cumple, no puede considerarse como una aplicación de T.S.S.

Por otra parte, parece intuitivamente claro que la condición $D3$-(3) es «más fundamental» que $D3$-(4) para T.S.S. $D3$-(3) es una aserción extremadamente general sobre la relación básica entre los parámetros que describen un sistema simple, una aserción que contiene todas las informaciones «esenciales» necesarias para la aplicación subsiguiente de la teoría a casos especiales. Tisza observa:

«... la ecuación fundamental [a saber, en nuestra reconstrucción: el miembro derecho del bicondicional de $D3$-(3)] realmente merece este nombre. Su propiedad más notable es que contiene un enunciado completo de la información termodinámica concerniente a todos los estados de equilibrio del sistema en cuestión. Expresamos este hecho importante diciendo que la ecuación fundamental, junto con sus ingredientes, la energía y la entropía, proporciona una *parametrización de todos los estados de equilibrio*» (Tisza, *op. cit.*, p. 40).

En otras palabras, $D3$-(3) es un *principio-guía* de T.S.S. en el sentido en que hemos introducido este concepto en el capítulo 2.3. En cambio, $D3$-(4) ciertamente *no* es una ley fundamental o principio-guía en este sentido; presistemáticamente, debería ser considerada más bien como una relación interteórica (un «principio de correspondencia», como dice Tisza), un puente entre la termodinámica y la mecánica. En un trabajo futuro trataremos de desarrollar un aparato conceptual apropiado para reconstruir esta clase de principios interteóricos, por el que se verá claramente que $D3$-(3) y $D3$-(4) no caen bajo la misma categoría metateórica. Pero, de momento, aún no disponemos de tal aparato. La solución más adecuada para la reconstrucción presente parece ser la de introducir la relación entre presión, volumen y energía también en la estructura matemática del núcleo básico de T.S.S. En cualquier caso, la discusión del «caso de la presión» muestra, a mi entender, claramente que no es posible una reconstrucción *completamente* satisfactoria de una teoría física sin analizar sus relaciones «externas» con otras teorías.

Al poner la entropía en una dependencia funcional respecto de las otras primitivas de T.S.S. a través de f^S decimos que hemos escogido la «representación entrópica» de la termodinámica. Pero los requisitos de diferenciabilidad continua y de monotonía estricta de f^S nos permiten tomar cualquier otra función primitiva y ponerla en la correlación funcional apropiada con la entropía y las restantes funciones (en el caso de estados en equilibrio). En particular, podemos poner la energía en una dependencia funcional de las demás funciones. Esto sería la «representación energética» de la termodinámica.

Vamos a expresar lo anterior más exactamente. En general, tenemos la función compuesta $\psi(z)=f^S(U(z), V(z), ...)$. En el caso de estados en equilibrio, tiene que ser $\psi(z)=S(z)$. Ahora bien, partiendo de la diferenciabilidad y monotonía de f^S con respecto a $U(z)$, etc., resulta que, para todo z, $D_U f^S(z)\neq 0$. Por un teorema general del análisis matemático (el teorema de las funciones implícitas), esto, a su vez, implica que existe un funcional unívocamente definido f^U tal que $U(z)=f^U(\psi(z), V(z), ...)$. En el caso de estados en equilibrio, dado que $\psi(z)=S(z)$, obtenemos $U(z)=f^U(S(z), V(z), ...)$. Esta es la representación energética de la ecuación fundamental de T.S.S. Tanto la representación energética como la entrópica se usan por igual en las aplicaciones de T.S.S. Podemos pasar de la una a la otra según parezca más conveniente para nuestras necesidades de cálculo. Son empíricamente equivalentes. Vamos a expresar todo esto en un teorema. (Asumimos aquí y en el resto del capítulo que estamos tratando con funciones que pertenecen a algún modelo actual de T.S.S.)

*T*1: Para cualquier $x\in\widehat{TGS}$, existe una función f^U unívocamente definida, tal que:

$$\forall z(z\in E\leftrightarrow U(z)=f^U(S(z), V(z), N(i_1, z), ..., N(i_{|I|}, z))).$$

Por la propiedad $D_U f^S(z)\neq 0$ que se desprende de *D*3-(3) también podemos deducir a partir de teoremas del análisis que:

T 2: Para todo $z\in E$: $D_S f^U(z)=\dfrac{1}{D_U f^S(z)}$.

(Esta propiedad nos será útil más adelante.)

Diferenciando la ley fundamental *D*3-(3) en la representación entrópica o en la energética respectivamente (recuérdese que cada $z\in E$ pertenece a un conjunto abierto de estados en equilibrio), obtenemos las ecuaciones diferenciales usuales. Para que la fórmula sea más perspicua, escribiremos $DN_i(z)$ en vez de $D_z N(i, z)$:

T 3: a) $\forall z(z\in E\rightarrow DS(z)=D_U f^S(z)\cdot DU(z)+D_V f^S(z)\cdot DV(z)+$

$$+\sum_{i\in I} D_{N_i} f^S(z)\cdot DN_i(z)).$$

b) $\forall z(z\in E\rightarrow DU(z)=D_S f^U(z)\cdot DS(z)+D_V f^U(z)\cdot DV(z)+$

$$+\sum_{i\in I} D_{N_i} f^U(z)\cdot DN_i(z)).$$

Otra propiedad importante de los modelos de T.S.S. (importante porque representa una conexión directa con teorías mecánicas) es la que se suele expresar en los libros de texto diciendo que «la presión negativa de un sistema en el que no ocurren cambios caloricos ni químicos es la derivada de la energía respecto al volumen». Esto es, por supuesto, una consecuencia de nuestro axioma *D*3-(4).

*T*4: Cuando en un sistema en equilibrio no se dan variaciones caloricas ni químicas, la presión negativa del sistema es igual a la derivada de la determinación energética respecto del volumen.

Prueba: Dado que suponemos el sistema en equilibrio, podemos escribir las dos ecuaciones fundamentales empíricamente equivalentes:

(1) $DS(z)=D_U f^S(z)\cdot DU(z)+D_V f^S(z)\cdot DV(z)+\sum_{i\in I} D_{N_i} f^S(z)\cdot DN_i(z).$

(2) $DU(z)=D_S f^U(z)\cdot DS(z)+D_V f^U(z)\cdot DV(z)+\sum_{i\in I} D_{N_i} f^U(z)\cdot DN_i(z).$

El supuesto de que no se dan variaciones caloricas ni químicas significa en nuestro formalismo que

(3) $DS(z)=DN_i(z)=0$ para todo $i\in I$.

De ahí resulta:

(4) $0=D_U f^S(z)\cdot DU(z)+D_V f^S(z)\cdot DV(z)$ por (1), (3).

(5) $-\dfrac{D_V f^S(z)}{D_U f^S(z)}=\dfrac{DU(z)}{DV(z)}$ por (4).

(6) $-P(z)=\dfrac{DU(z)}{DV(z)}$ por (5) y *D*3-(4).

(7) $DU(z)=D_V f^U(z)\cdot DV(z)$ por (2), (3).

(8) $\dfrac{DU(z)}{DV(z)} = D_V f^U(z)$ por (7).

(9) $-P(z) = D_V f^U(z)$ por (6), (8).

Definimos ahora algunas importantes magnitudes derivadas.

D4: La *temperatura* (absoluta) T es una función de $E \subseteq Z$ en $\mathbb{R}$ tal que:

$$T(z) =: \frac{1}{D_U f^S(z)} \; (= D_S f^U(z) \text{ por } T2).$$

D5: El *potencial químico* μ es una función de $I \times E$ en $\mathbb{R}$ tal que:
$\mu(i, z) =: -T(z) \cdot D_{N_i} f^S(z)$ $(= D_{N_i} f^U(z)$, por una prueba análoga a la de $T4)$.

(Para cada sustancia química i, $\mu(i, z)$ se llama «el potencial químico de la sustancia i».)

Si introducimos P, T, μ, en la ecuación diferencial de $T3$-(b), obtenemos la forma canónica de la ecuación fundamental de T.S.S.:

T 5: $\forall z (z \in E \rightarrow DU(z) = T(z) \cdot DS(z) - P(z) \cdot DV(z) + \sum_{i \in I} \mu(i, z) \cdot DN_i(z))$.

Pasamos ahora a la formulación de las condiciones generales de ligadura de T.S.S. En esta teoría, las ligaduras entre los distintos modelos juegan un papel especialmente importante para darle contenido empírico a la teoría y obtener propiedades termodinámicas generales de los sistemas simples.

La mayoría de condiciones de ligadura de la termodinámica tienen que formularse mediante una operación de *concatenación* «o», cuyas propiedades formales y función son similares a las de la operación de concatenación de partículas introducida en la mecánica para expresar la aditividad de la masa. Pero la concatenación parece ser mucho más significativa en la termodinámica que en la mecánica, por lo que merece una discusión detallada.

Consideremos primero la estructura y función de la concatenación termodinámica en términos fisicamente intuitivos. Supóngase que el sistema descrito por un modelo x consiste en un gas encerrado dentro de un recipiente. Solamente una de las paredes del recipiente no es totalmente impermeable. Supóngase que lo mismo es válido del sistema correspondiente a otro modelo x'. Vamos a concentrarnos en la consideración de las funciones energía U y U'. Si el sistema se halla

en un estado $z \in Z$ de energía $U(z)$ podemos imaginar un estado posterior z_1 cuya energía sea menor: $U(z_1) \leqslant U(z)$. Esto puede haber ocurrido porque algunas de las moléculas hayan abandonado el recipiente al pasar a través de la pared permeable. Un proceso similar es posible para Z' de x'. Es obvio que, en principio, no tiene por qué haber ninguna conexión entre lo que ocurre en Z y en Z'. Sea cual sea la energía $U(z)$ de Z, ella no nos da ninguna información sobre la energía $U'(z')$ de Z'. Además, no hay modo de saber si, y en qué medida, la variación de energía en x depende de la variación de energía en x'.

Ahora bien, supongamos ahora que ambos sistemas están conectados de tal manera que las dos paredes permeables de uno y otro coinciden exactamente superpuestas. Intuitivamente, podemos decir que el sistema total, que resulta de la combinación de los dos primeros, es un sistema cerrado. Su energía total no varía y es simplemente la suma de las energías de ambos subsistemas. En tal caso, los cambios de energía en un subsistema dependerán unívocamente de los cambios de energía en el otro.

Para expresar todo esto más precisamente introducimos una operación de concatenación para los estados de modelos distintos. A partir de los dos estados z y z' de los modelos x y x', la operación de concatenación (que denotamos por «o») contituye un nuevo estado $z \circ z'$. Intuitivamente, éste es el estado del sistema que resulta de la combinación de los dos sistemas originales. Lo más natural es imaginar que la combinación de los dos sistemas ocurre en forma de yuxtaposición espacial de modo que se da algún tipo de interacción (pongamos por caso, la resultante de la superposición de dos paredes permeables como en el ejemplo anterior). Así, dos estados cualesquiera de los sistemas originales determinan unívocamente por concatenación un estado correspondiente en el sistema total.

Es preciso hacer notar que el ejemplo intuitivo anterior de concatenación sólo cubre un caso muy especial. La concatenación de dos sistemas no tiene por qué consistir en una yuxtaposición en el espacio. El proceso abstracto general es el siguiente. Tenemos dos sistemas Z y Z' (que pueden o no interaccionar en algún sentido «real»). Lo que de hecho nos interesa es poder construir conceptualmente la concatenación de los estados de ambos sistemas. Queremos que esto sea posible en cualquier caso, incluso si Z y Z' no interactúan físicamente. Esto es, de lo que se trata es de establecer una función diádica o de $Z \times Z'$ en un nuevo conjunto de estados Z''. La dificultad que aparece a primera vista es la de caracterizar exactamente este Z''. No hay modo de *definirlo* o explicitarlo en términos físicos a partir de Z y Z', puesto que la operación o es una noción «abierta». Lo que debemos hacer es tomar o, y por tanto Z'' mismo, como un concepto primitivo, no-definible ni ulteriormente caracterizable de T.S.S. Simplemente asumimos (como en el caso de la

mecánica) que o es una operación que puede definirse sobre la unión de todos los estados de todos los modelos potenciales de T.S.S., de tal modo que a dos estados cualesquiera les asigna su concatenación. Formalmente, establecemos el siguiente postulado sobre la totalidad de los modelos de T.S.S.:

*P*1: Existe una operación $\circ$ de $\mathfrak{Z} \times \mathfrak{Z}$ en $\mathfrak{Z}$, donde $\mathfrak{Z} =: \cup\{\pi_1 x / x \in M_p(\text{T.S.S.})\}$, y tal que $\circ$ es asociativa y conmutativa.

Retomemos ahora nuestro ejemplo anterior. En ese caso lo que hacemos es tomar dos sistemas que suponemos que son modelos potenciales de T.S.S. y que se hallan respectivamente en los estados z y z', y concatenarlos para formar un sistema «total», que también suponemos que es un modelo potencial de T.S.S., y que se hallará en el estado z'', resultado de la concatenación de z y z'. Entonces, es un postulado esencial de la teoría que la energía del estado «total» z'' será la suma de las energías parciales de z y z'. Dicho más formalmente: si $z'' = z \circ z'$, entonces $U''(z'') = U(z) + U(z')$. Así podemos formular con precisión (mediante el concepto de concatenación de estados) el postulado de la aditividad de la energía en la termodinámica clásica.

Dado que en la formulación de este postulado intervienen tres modelos potenciales distintos (que corresponden, respectivamente, a los estados z, z' y z''), es decir, conectamos con él las funciones energía de tres aplicaciones distintas de la teoría, se tratará evidentemente de una condición de ligadura de la teoría. Conexiones del tipo ejemplificado son absolutamente esenciales en la teoría termodinámica. Pues si no pudiéramos concatenar diferentes sistemas termodinámicos de diferentes modos y hacer afirmaciones sobre la conexión entre las funciones de los sistemas concatenados, la teoría entera se convertiría en un formalismo muy pobre en contenido, de muy escasa aplicabilidad. Las ligaduras son absolutamente indispensables para la aplicación de la termodinámica. Debemos incluirlas con todo el detalle necesario dentro del concepto de la teoría.

Para formular las ligaduras de T.S.S. necesitamos introducir cierto número de nociones técnicas auxiliares. En primer lugar, es preciso establecer la distinción entre funciones *extensivas* e *intensivas*.

*D*6: g es una *función extensiva* ($g \in EXT$) si y sólo si: $g \in \{U, S, V, N\}$. h es una *función intensiva* ($h \in INT$) si y sólo si existe $g \in EXT$ tal que $h = \pm D_g f^S$, o bien $h = \pm D_g f^U$.

Otra noción que precisamos es la de concatenación de dos sistemas, definida a partir de la concatenación de los estados respectivos.

*D*7: $Z_x \bar{\circ} Z_{x'} =: \{z \circ z' / z \in Z_x \wedge z' \in Z_{x'}\}$.

Una clase particular de sistemas compuestos de T.S.S. son aquellos cuyos subsistemas componentes se hallan en equilibrio mutuo parcial (abreviadamente: en *equilibrio-i*). Decimos que dos sistemas se hallan en un equilibrio-*i* si hay al menos una función *intensiva* que toma el mismo valor en ambos sistemas. Un ejemplo es el de un sistema compuesto cuyas partes componentes se hallan en equilibrio térmico entre sí, es decir, en ellas la temperatura toma el mismo valor (aunque quizá la presión y los potenciales químicos tomen valores distintos en las distintas partes). Otro ejemplo sería el de un sistema compuesto cuyas partes componentes tienen la misma presión. A tales sistemas compuestos en equilibrio parcial se les llama a veces «sistemas reducidos». A ellos los vamos a definir formalmente como un conjunto de estados concatenados de ambos sistemas para los que una función intensiva determinada toma el mismo valor en uno y otro modelo.

*D*8: Sea $h \in INT$.

$[Z_x \bar{\circ} Z_{x'}]_h =: \{z \circ z' \in Z_x \bar{\circ} Z_{x'} / h_x(z) = h_{x'}(z')\}$.

Finalmente, recordemos algunas nociones del análisis matemático que aquí necesitamos para la formulación de algunas ligaduras.

*D*9: Sea f una función cualquiera y z un argumento suyo.
«relmax (f, z)» significa que f posee un máximo relativo en z.
«relmin (f, z)» significa que f posee un mínimo relativo en z.
«$\Delta f(z) = 0$» significa que f es constante en un entorno topológico de z.

Postulamos a continuación las condiciones generales de ligadura.

*D*10: $X \subseteq \widehat{PTS}$ satisface la ligadura C_E si y sólo si:

$$\forall x, x', z(x, x' \in X \wedge z \in E_x \wedge z \in Z_{x'} \rightarrow z \in E_{x'}).$$

Este requisito es una especie de «ligadura de identidad» para el concepto de equilibrio general: la propiedad de ser un estado de equilibrio es independiente del sistema al que el estado en cuestión pertenezca.

*D*11: $X \subseteq \widehat{PTS}$ satisface la ligadura C_E^{INT} si y sólo si:

$$\forall x, x', x'' \forall z, z'(x, x', x'' \in X \wedge Z_x'' = Z_x \bar{\circ} Z_{x'} \wedge z \circ z' \in Z_{x''} \rightarrow$$
$$\rightarrow (z \circ z' \in E_{x''} \leftrightarrow \forall_h \in INT(h_x(z) = h_{x'}(z')))).$$

Esta ligadura conecta la noción de equilibrio total con la de sistema reducido: un sistema compuesto se halla en un estado de equilibrio (total) si y sólo si ha sido «reducido» con respecto a todas su funciones intensivas. Dicho de otro modo, el equilibrio total se consigue a través de la realización simultánea de todos los equilibrios parciales.

*D*12: Sean $x, x', x'' \in \widehat{PTS}$ y $Z_{x''} = Z_x \bar{\circ} Z_{x'}$. Sea $h \in INT$. Si tomamos $h^*_{x''}$ como la restricción de $h_{x''}$ definida sólo sobre $[Z_{x''}]_h$, tal que, para todo $z \in Z_x$, $z' \in Z_{x'}$: $h^*_{x''}(z \circ z') = h_x(z) = h_{x'}(z')$, entonces diremos que $X \subseteq \widehat{PTS}$ satisface la ligadura C^{INT} si y sólo si:

$$\forall x, x', x'', z, z' (x, x', x'' \in X \wedge Z_{x''} = Z_x \bar{\circ} Z_{x'} \wedge$$
$$z \circ z' \in E_{x''} \subseteq Z_{x''} \rightarrow h_{x''}(z \circ z') = h^*_{x''}(z \circ z')).$$

Es decir, los valores de la función intensiva h definida sobre el sistema compuesto $Z_{x''}$ tiene que coincidir con los valores de h en cada uno de los subsistemas Z_x y $Z_{x'}$.

*D*13: $X \subseteq \widehat{PTS}$ satisface la ligadura $C^{U,S}_E$ si y sólo si:

$$\forall x, x', x'', z, z' \forall h \in INT (x, x', x'' \in X \wedge z \in Z_x \wedge z' \in Z_{x'} \wedge$$
$$z \circ z' \in Z_{x''} = Z_x \bar{\circ} Z_{x'} \wedge z \circ z' \notin [Z_{x''}]_h \rightarrow (((\Delta S_{x''}(z \circ z') = 0$$
$$\rightarrow \exists z_e (z_e \in Z_{x''} \wedge z \circ z' \prec z_e \wedge \Delta S_{x''}(z_e) = 0 \wedge z_e \in [Z_{x''}]_h)$$
$$\wedge (\Delta U_{x''}(z \circ z') = 0 \rightarrow \exists z'_e (z'_e \in Z_{x''} \wedge z \circ z' \prec z'_e \wedge$$
$$\Delta U_{x''}(z'_e) = 0 \wedge z'_e \in [Z_{x''}]_h))).$$

$\Delta S_{x''}(z \circ z') = 0$ y $\Delta U_{x''}(z \circ z') = 0$ expresan dos posibles «condiciones de aislamiento» del sistema compuesto $Z_{x''}$: una adiabática y una energética, respectivamente. La ligadura postula entonces que un sistema compuesto que se mantiene en una de esas dos condiciones de aislamiento y *no* se halla en un estado de equilibrio-*i* con respecto a alguna función h (=que no ha sido «reducido» respecto a la función intensiva h), «tiende» a un estado «posterior» z_e de equilibrio-*i* con respecto a h, el cual se hallará también en aislamiento adiabático o energético, respectivamente. Esta condición de ligadura fija la dirección de los procesos termodinámicos cuando se considera la combinación de sistemas. También significa que T.S.S. no está interesada en procesos que no tienden a un equilibrio-*i* bajo las condiciones especificadas (por ejemplo, procesos que en aislamiento conducen a una explosión); tales procesos quedan eliminados del alcance de la teoría por esa ligadura.

Las exposiciones corrientes de la termodinámica asumen un postulado similar sobre la dirección de los procesos termodinámicos más o menos implícitamente en conexión con el llamado «Segundo Principio de la Termodinámica». A veces se supone que la formula-

ción gibbsiana del segundo principio ya incluye o implica aquel postulado. Sin embargo, el principio de Gibbs, si se lo formula correctamente, solamente fija las propiedades de extremo de la entropía y la energía en un equilibrio-*i*; no afirma que los sistemas «tiendan» a un equilibrio tal. En rigor se trata de un postulado adicional de tipo direccional. Veremos su función más abajo en la consecución de importantes resultados.

*D*14: $X \subseteq \widehat{PTS}$ satisface la ligadura $C^{=}_{U,S}$ si y sólo si:

$$\forall x, x', z(x, x' \in X \wedge z \in Z_x \cap Z_{x'} \rightarrow U_x(z) = U_{x'}(z) \wedge S_x(z) = S_{x'}(z)).$$

Esta es la ligadura de identidad para la energía y la entropía, análoga a la ligadura de identidad para la masa en M.C.P.

*D*15: $X \subseteq \widehat{PTS}$ satisface la ligadura $C^{+}_{U,S}$ si y sólo si:

$$\forall x, x', x'', z, z'(x, x', x'' \in X \wedge z \in Z_{x'} \wedge z' \in Z_{x''} \wedge$$
$$z \circ z' \in Z_{x'''} = Z_{x'} \bar{\circ} Z_x \rightarrow U_{x'''}(z \circ z') = U_x(z) + U_{x'}(z') \wedge$$
$$S_{x'''}(z \circ z') = S_x(z) + S_{x'}(z')).$$

Se trata aquí de la ligadura de aditividad para la energía y la entropía bajo concatenación.

*D*16: $X \subseteq \widehat{PTS}$ satisface la ligadura C_U si y sólo si:

$$\forall x, z(x \in X \wedge z \in Z_x \wedge DU_x(z) \neq 0 \rightarrow \exists x', x'', z'(x', x'' \in X \wedge z' \in Z_{x''}$$
$$\wedge z \circ z' \in Z_{x'''} = Z_x \bar{\circ} Z_{x'} \wedge DU_{x'''}(z \circ z') = 0)).$$

Esta es nuestra formulación del «postulado de la conservación de la energía»: si un sistema Z_x experimenta un cambio de energía, entonces se puede encontrar otro sistema $Z_{x'}$ tal que concatenado con el primero «compensa» su variación de energía, de modo que el sistema total posea energía constante. Nótese que $Z_{x'}$ mismo puede ser ya un sistema compuesto y «tan grande» como se quiera. En cuanto a $Z_{x''}$ en esta formulación sólo sirve como marco para expresar la energía compuesta. (A veces se interpreta $Z_{x''}$ como «el universo», pero no necesitamos hacer uso de este modo de hablar aquí.)

Nuestra última condición de ligadura será el principio extremal de Gibbs, que con frecuencia se toma como «empíricamente equivalente» al Segundo Principio de la Termodinámica en la formulación clásica de Kelvin-Clausius. El principio de Gibbs consta de dos partes: la primera afirma que la entropía toma un valor máximo (relativo) en un sistema compuesto cuyas partes se hallan en equilibrio-*i* bajo ciertas

condiciones; la segunda, que la energía toma un valor mínimo bajo condiciones análogas.

D17: $X \subseteq \widehat{PTS}$ satisface C_G si y sólo si: para todo x, x', $x'' \in X$, $z \in Z_x$, $z' \in Z_{x'}$, $h \in INT$, si $z \circ z' \in Z_{x''} = Z_x \bar{\circ} Z_{x'} \wedge \exists g \in EXT$

$(g \neq U \wedge g \neq S \wedge (h = \pm D_g f^S \vee h = \pm D_g f^U) \wedge \Delta g_{x''}(z \circ z') \neq 0) \wedge$
$\forall g' \in EXT (g' \neq g \rightarrow \Delta g'_x(z) = \Delta g'_{x'}(z') = 0)$, entonces:

(a) si $\Delta U_{x''}(z \circ z') = 0$, entonces $(h_x(z) = h_{x'}(z') \leftrightarrow \text{relmax}\ (S_{x''}, z \circ z')) \wedge$

(b) si $\Delta S_{x''}(z \circ z') = 0$, entonces $(h_x(z) = h_{x'}(z') \leftrightarrow \text{relmin}\ (U_{x''}, z \circ z'))$.

Esto es, si la energía tiene un valor constante en un entorno topológico del estado compuesto $z \circ z'$ (aislamiento energético del sistema compuesto) y si dejamos variar sólo una función g, manteniendo constantes las demás funciones en cada subsistema, entonces el sistema compuesto alcanzará el equilibrio-*i* correspondiente a g si y sólo si la entropía total toma un máximo relativo. Una aserción análoga es válida para la energía: si la entropía total se mantiene constante, entonces el sistema alcanza un equilibrio-*i* si y sólo si la energía toma un mínimo relativo.

Este principio extremal sólo tiene sentido si se refiere a sistemas compuestos. Este punto, que se suele ignorar, ha sido argumentado convincentemente por Tisza en su análisis crítico del principio de Gibbs (Tisza, *Generalized Thermodynamics*, p. 41); los procesos que ocurren en sistemas que «tienden» a un máximo de entropía (o a un mínimo de energía) son procesos que se dan dentro de un sistema compuesto.

Muchas presentaciones de la termodinámica toman el principio maximal de la entropía y el principio minimal de la energía como dos enunciados lógicamente equivalentes en el formalismo de la teoría; sin embargo, como han señalado correctamente Falk y Jung (*Axiomatik der Thermodynamik*, p. 153), la equivalencia asumida es válida solamente si se admiten ciertos supuestos implícitos adicionales. Un análisis cuidadoso de las «pruebas» de esta equivalencia muestra que no sólo se presupone la positividad de la temperatura (la cual es un *teorema* en la reconstrucción presente), sino también la posibilidad de ciertos «procesos cíclicos». La formulación explícita de estos supuestos daría lugar a una reconstrucción mucho más complicada de la teoría.

A la intersección conjuntista de todas las ligaduras termodinámicas expuestas hasta aquí, C_E, C_E^{INT}, C^{INT}, $C_{E,S}^{U,S}$, $C_{\bar{U},S}$, $C_{\dot{U},S}^{+}$, C_U, C_G, la llamaremos «condición de ligadura básica de T.S.S.» y la designaremos por $C_B[TSS]$. Así, pues, el núcleo básico de T.S.S. es la estructura: $K_B[TSS] = \langle PTS, TGS, PS, C_B[TSS] \rangle$.

Lo anterior completa la descripción del elemento teórico básico de T.S.S. Esta estructura contiene ya propiedades termodinámicas importantes, que se pueden derivar como teoremas; por ejemplo, la positividad de la temperatura y los llamados principios de Kelvin y Clausius. En lo que sigue vamos a presuponer siempre que todos los estados y funciones consideradas (ya sean de sistemas simples o compuestos) pertenecen a modelos *actuales* de T.S.S. También presuponemos que los *conjuntos* de modelos considerados satisfacen todas las condiciones de ligadura enunciadas hasta aquí.

El próximo teorema afirma la positividad de la temperatura (en *esta* teoría termodinámica –puede que haya otras teorías termodinámicas donde este teorema ya no se cumple). La afirmación de la positividad de la temperatura se obtiene bajo ciertos supuestos: que para cada estado en un sistema hay otro estado en otro sistema de temperatura distinta tal que podemos iniciar un proceso en el que el sistema compuesto se halle adiabática y energéticamente aislado, y de modo que todos los demás parámetros en cada sistema componente se mantienen constantes. Formalmente, estos supuestos parecen ser bastante fuertes, de modo que la afirmación de la positividad de la temperatura no sería universalmente válida. No obstante, desde el punto de vista empírico (experimental) sabemos que estas condiciones siempre se pueden realizar de hecho en aquellos sistemas de los que trata la termodinámica clásica. Hablando con un poco más de precisión, las *aplicaciones propuestas* de esta teoría son tales que estas condiciones siempre son realizables.

T6: Para cualquier $z_1 \in E_1 \subseteq Z_1$, si existe un $z_2 \in E_2 \subseteq Z_2$ con $z_1 \circ z_2 \in Z = Z_1 \bar{\circ} Z_2$, tal que $(T(z_1) \neq T(z_2) \wedge \Delta U_Z(z_1 \circ z_2) = 0 \wedge \Delta S_Z(z_1 \circ z_2) = 0 \wedge \forall F_Z(U_Z \neq F_Z \neq S_Z \rightarrow \Delta F_Z(z_1) = \Delta F_Z(z_2) = 0)$, entonces $T(z_1) > 0$.

Prueba: Las premisas son:

(1) $z_1 \in E_1 \wedge z_2 \in E_2$.
(2) $T(z_1) \neq T(z_2)$.
(3) $\Delta U_Z(z_1 \circ z_2) = 0$.
(4) $\Delta S_Z(z_1 \circ z_2) = 0$.
(5) $\forall F_Z(U_Z \neq F_Z \neq S_Z \rightarrow \Delta F_Z(z_1) = \Delta F_Z(z_2) = 0)$.

(Nótese que estas condiciones siempre son realizables en el laboratorio bajo condiciones normales.) De ellas se sigue:

(6) $DS(z_1) = \dfrac{1}{T(z_1)} \cdot DU(z_1) \wedge DS(z_2) = \dfrac{1}{T(z_2)} \cdot DU(z_2)$ por (1), (5), *T3*, *D4*.

(7) $z_1 \circ z_2 \notin [Z]^1_{\frac{1}{T}}$ por (2), *D*8.

(8) $\neg$ relmax $(S, z_1 \circ z_2)$ por (3), (5), (7), *D*17.

(9) $\exists z_e (z_1 \circ z_2 \prec z_e \wedge z_e \in [Z]^1_{\frac{1}{T}} \wedge \Delta U(z_1 \circ z_2) = \Delta U(z_e) = 0)$ por (3), (7), *D*13.

(10) relmax (S, z_e) por (3), (5), (9), *D*17.

(11) $\Delta f^S(z_1 \circ z_2) = 0$ por (3), (5) y la monotonía estricta de f^S con respecto a sus argumentos (*D*3-(3)).

(12) $DS(z_1 \circ z_2) \geqslant 0$ por (11) y la monotonia débil de S con respecto a f^S (*D*3-(3)).

(13) $DS(z_1 \circ z_2) = \frac{1}{T(z_1)} \cdot DU(z_1) + \frac{1}{T(z_2)} \cdot DU(z_2) \geqslant 0$ por (6), (12).

(14) $DU(z_1) = -DU(z_2)$ por (3), *D*15.

(15) $-\frac{1}{T(z_1)} \cdot \frac{T(z_2) - T(z_1)}{T(z_2)} \cdot DU(z_2) \geqslant 0$ por (13), (14).

Hasta este punto hemos hecho uso de la condición (3) como base del argumento. Vamos a basarnos ahora en la condición (4). Aplicando la ley fundamental en la representación *energética* diferencial, podríamos desarrollar un argumento enteramente análogo al precedente, hasta llegar a la desigualdad:

(16) $DU(z_1 \circ z_2) \leqslant 0$ (en vez de (13) y por las razones correspondientes).

(17) $DS(z_1) = -DS(z_2)$ por (4), *D*15.

(18) $(T(z_2) - T(z_1)) \cdot DS(z_2) \leqslant 0$ por (6), (16), (17), *D*15.

(19) $\frac{T(z_2) - T(z_1)}{T(z_2)} \cdot DU(z_2) \leqslant 0$ por (6), (18).

Si comparamos (15) con (19) es fácil ver que ambas desigualdades sólo son válidas en caso de que:

(20) $T(z_1) \geqslant 0$.

Dado que sabemos que una de las propiedades implicadas analíticamente por las propiedades de f^S postuladas en *D*3-(3) es la de que $D_U f^S(z) \neq 0$ para todo z, y además por definición $T(z) = \frac{1}{D_U f^S(z)}$, se sigue de ahí que $T(z) \neq 0$ para todo z, y por tanto:

(21) $T(z_1)>0$.

Como corolario de este teorema obtenemos el *Principio de Kelvin:* si todas las funciones extensivas, excepto energía y entropía, se mantienen constantes en un sistema dado en equilibrio, es imposible sacar sólo energía del sistema, pero no entropía. En efecto, $DU(z_1) = T(z_1)\cdot DS(z_1)$ y $T(z_1)>0$ implican que $DS(z_1)\geqslant 0$ es imposible si $DU(z_1)<0$.

A continuación obtenemos el *Principio de Clausius:* si en un sistema compuesto bajo aislamiento energético, el volumen y los moles se mantienen constantes, entonces el subsistema con mayor temperatura pierde energía en beneficio del subsistema con menos temperatura.

T7: Para todo $z_1\in Z_1$, $z_2\in Z_2$ con $z_1\circ z_2\in Z_1\bar{\circ}\ Z_2$ se cumple:

$$\Delta U(z_1\circ z_2)=0\wedge\forall g(g\neq S\wedge g\neq U\rightarrow\Delta g(z_1)=\Delta g(z_2)=0)$$
$$\wedge\, T(z_1)>T(z_2)\rightarrow DU(z_1)\leqslant 0\wedge DU(z_2)\geqslant 0.$$

Prueba: Las premisas son:

(1) $\Delta U(z_1\circ z_2)=0$.
(2) $\forall g(g\neq S\wedge g\neq U\rightarrow\Delta g(z_1)=\Delta g(z_2)=0)$.
(3) $T(z_1)>T(z_2)$.

De la premisa (3) se desprende:

(4) $z_1\circ z_2\notin[Z_1\bar{\circ}Z_2]_{\frac{1}{T}}$.

A partir de (4) y de las premisas se puede derivar, siguiendo el mismo argumento que hemos usado en *T6*, que:

(5) $DS(z_1\circ z_2)\geqslant 0$.

(6) $DS(z_1\circ z_2)=\dfrac{1}{T(z_1)}\cdot DU(z_1)+\dfrac{1}{T(z_2)}\cdot DU(z_2)$, por (2), *T3*, *D4*, *D15*

(7) $\left(\dfrac{1}{T(z_1)}-\dfrac{1}{T(z_2)}\right)\cdot DU(z_1)\geqslant 0$ por (1), (5), (6), *D15*.

(8) $\dfrac{1}{T(z_1)}-\dfrac{1}{T(z_2)}<0$ por (3).

(9) $DU(z_1)\leqslant 0\wedge DU(z_2)\geqslant 0$, por (7), (8), *D15*.

Un análisis detallado de las pruebas de los últimos teoremas mostraría claramente lo importantes que son las condiciones de ligadura en T.S.S. para derivar propiedades esenciales del núcleo de T.S.S.

3. Especializaciones

Como es de esperar, la mayoría de especializaciones del núcleo básico de T.S.S. son especificaciones de la función «abstracta» f^S de determinación entrópica (o bien, alternativamente, de la determinación energética f^U) que aparece en la ecuación fundamental de T.S.S. ($D3$-(3)). Al especificar sucesivamente la forma de f^S (o de f^U) se obtienen diversas líneas de especialización para la aplicación de T.S.S. a casos particulares. No obstante, la primera especialización que consideramos *no* es de la forma antedicha: la ley de Nernst, llamada también «Tercer Principio de la Termodinámica» en la literatura usual, no proviene de ninguna especificación de f^S, sino de una condición que relaciona los estados de energía mínima con los de entropía también mínima.

De manera análoga al caso de la «Tercera Ley de Newton» en mecánica, algunos lectores familiarizados con la tradición termodinámica puede que se pregunten por qué no consideramos el «Principio de Nernst» como ley igual de fundamental que la ecuación de estado, o como el Principio de Gibbs, y por tanto como perteneciente al núcleo básico de esta teoría. Dejando a un lado las evaluaciones tradicionales de esta ley, el hecho es que el alcance de la aplicación de la ley de Nernst (que originariamente sólo se proponía explicar la conducta de sistemas a muy bajas temperaturas) es mucho más restringido que el de los axiomas postulados en el núcleo de T.S.S. En realidad, se pueden conseguir una gran cantidad de resultados dentro de la termodinámica del equilibrio sin preocuparse en absoluto por la ley de Nernst. Históricamente, también es un hecho que la termodinámica se desarrolló bastante bien durante varias décadas sin tener conocimiento de esta ley. Hay muchas aplicaciones de la termodinámica que no presuponen la ley de Nernst. Más aun, existen leyes especiales de la termodinámica, como la de Gay-Lussac, que son incompatibles con la ley de Nernst (lo cual no impide a los practicantes de la termodinámica aplicar la ley de Gay-Lussac siempre que lo consideran conveniente). Finalmente, algunos resultados recientes que se encuentran en la literatura parecen señalar que hay algunas aplicaciones propuestas para las que la validez de la ley de Nernst, incluso para temperaturas bajas, aparece como muy dudosa[28].

[28] Una consideración de estos resultados se hallará en H. W. Buchdahl, *The Concepts of Classical Thermodynamics*, pp. 110-112.

En consecuencia, parece más plausible tomar la ley de Nernst como una especialización y no como una ley fundamental de T.S.S. De todos modos, si así lo deseamos, podemos llamarla una «especialización muy general», como en el caso del principio de acción y reacción de Newton.

Para formalizar la ley de Nernst dentro del presente marco necesitamos dos definiciones auxiliares.

*D*18: Para cualquier $x \in \widehat{PTS}$ definimos:

(a) $Z_{S^o} = \{z/z \in Z_x \wedge \forall z'(z' \in Z_x \rightarrow S_x(z) \leqslant S_x(z'))\}$.
(b) $Z_{U^o} = \{z/z \in Z_x \wedge \forall z'(z' \in Z_x \rightarrow U_x(z) \leqslant U_x(z'))\}$.

Es decir, los estados en Z_{S^o}, respectivamente en Z_{U^o}, son los de entropía mínima, respectivamente, de energía mínima.

*D*19: x es una *termodinámica simple de Nernst* ($x \in \widehat{TSN}$) si y sólo si:

(1) $x \in \widehat{TGS}$.
(2) $\varnothing \neq Z_{U^o} \subseteq Z_{S^o}$.

Es decir, en una termodinámica de Nernst *hay* estados de energía mínima, y éstos son también estados de entropía mínima[29].

La ligadura especial asociada a esta ley es una condición que aparece explícitamente enunciada en la literatura: todos los estados «mínimos» poseen los mismos valores de la energía y de la entropía en todos los sistemas.

*D*20: El conjunto $X \subseteq \widehat{TSN}$ satisface la ligadura $C_{\bar{N}}$ si y sólo si:

$$\forall x, x' \in X \, \forall z \in Z_x, z' \in Z_{x'} : z \in Z_{U^o_x} \wedge z' \in Z_{U^o_{x'}} \rightarrow$$
$$U_x(z) = U_{x'}(z') \wedge S_x(z) = S_{x'}(z').$$

Obtenemos una especialización muy distinta de la anterior al introducir la ley de los gases ideales, conocida también como ley de Gay-Lussac. Ya hemos señalado que esta ley, aunque se usa ampliamente en la termodinámica simple, de hecho es incompatible con la ley de Nernst. Sus aplicaciones propuestas también son muy diferentes de las de la ley de Nernst: gases homogéneos a temperaturas altas.

---

[29] La presente formulación está basada en la de Falk en su «On the Third Law of Thermodynamics», la cual a su vez es una generalización del postulado original de Nernst y Planck.

*D*21: x es un *gas ideal* ($x \in \widehat{GI}$) si y sólo si existen $i \in \mathbb{N}$, $R \in \mathbb{R}$, $z_o \in Z_x$, g^S, tales que:

(1) $x \in \widehat{TGS}$ con $I_x = \{i\}$.

(2) g^S es una función de $\mathbb{R}$ en $\mathbb{R}$ tal que g^S es dos veces continuamente diferenciable, y g^S es estrictamente monótona con respecto a sus argumentos, y S es débilmente monótona con respecto a g^S, y $g^S\left(\frac{U(z_o)}{N(i, z_o)}\right) = 0$.

(3) $\forall z(z \in E \leftrightarrow f^S(U(z), V(z), N(i, z)) = \frac{N(i, z)}{N(i, z_o)} \cdot S(z_o) + N(i, z) \cdot g^S\left(\frac{U(z)}{N(i, z)}\right) + N(i, z) \cdot R \cdot \log_n\left(\frac{V(z)}{V(z_o)} \cdot \frac{N(i, z_o)}{N(i, z)}\right)\right)$.

z_o es un estado «fiducial» escogido por convención; g^S y z_o deben elegirse «en mutua armonía», de modo que se cumpla la ecuación $g^S\left(\frac{U(z_o)}{N(i, z_o)}\right) = 0$. R es la constante universal de los gases; su valor viene normalmente dado por $R = 1{,}968$ cal/mol °K. «$\log_n$» significa «el logaritmo natural».

Con esta ley especial obtenemos un primer ejemplo de una forma particular de la ecuación fundamental contenida en la condición *D*3-(3). El nuevo axioma *D*21-(3) resulta de una especificación de la función general f^S de la ley fundamental. Pero la especificación todavía es incompleta, ya que en *D*21-(3) aparece una *variable* funcional g^S; g^S postula una relación entre U y N, pero no especifica esta relación. Una especificación completa de g^S resultará en una ley más especial, como veremos más abajo.

Veamos antes cómo la forma usual de la ley de los gases ideales, la que propiamente puede llamarse *Ley de Gay-Lussac*, se deduce de la ecuación de estado *D*21-(3).

*T*8: $\forall x \in GI$, $\forall z \in E_x$: $P_x(z) \cdot V_x(z) = N_x(i, z) \cdot R \cdot T_x(z)$.

Prueba:

(1) $P(z) = \frac{D_V f^S(z)}{D_U f^S(z)}$, por *D*3-(4).

(2) $T(z) = \frac{1}{D_U f^S(z)}$, por *D*4.

(3) $\frac{P(z)}{T(z)} = D_V f^S(z)$, por (1) y (2).

(4) $D_V f^S(z) = N(i, z) \cdot R \cdot D_V \log_n \left(\frac{V(z)}{V(z_o)} \cdot \frac{N(i, z_o)}{N(i, z)} \right) = N(i, z) \cdot R \cdot \frac{1}{V(z)}$, por premisas y *D*21-(3).

(5) $P(z) \cdot V(z) = N(i, z) \cdot R \cdot T(z)$, por (3) y (4).

Una ligadura natural que corresponde a esta ley es la ligadura de identidad para la constante R; llamémosla $C_R^=$. Nos ahorramos su formulación explícita. Otra condición de ligadura algo más compleja correspondiente a esta especialización es la siguiente.

*D*22: El conjunto $X \subseteq GI$ satisface la ligadura $C_g^=$ si y sólo si:

$\forall x, x' \in X: i_x = i_{x'} \rightarrow g_x^S = g_{x'}^S$.

Intuitivamente, esto significa que la función de determinación g^S de dos gases constituidos por la misma sustancia química i debe ser la misma en ambos gases.

La especialización precedente puede especificarse aún más hasta llegar a especializaciones «terminales» al estipular de modo completamente específico la forma de g^S, es decir, al dar los términos concretos de la determinación entrópica. Un ejemplo de tales especializaciones terminales se obtiene al especificar una función g^S que sólo se aplica a gases ideales *monoatómicos.*

*D*23: x es un *gas ideal monoatómico* ($x \in \widehat{GIM}$) si y sólo si:

(1) $x \in \widehat{GI}$.

(2) $\forall z (z \in E \leftrightarrow g^S \left(\frac{U(z)}{N(i, z)} \right) = \frac{3}{2} \cdot R \cdot \log_n \left(\frac{T(z)}{T(z_o)} \right)$.

Un simple cálculo nos permite derivar la forma explícita de la ecuación fundamental de T.S.S. para este caso:

$$T9: \forall x \in \widehat{GIM}, \forall z \in E_x: S(z) = \frac{N(i, z)}{N(i, z_o)} \cdot S(z_o) + N(i, z) \cdot R \cdot \log_n \left[\left(\frac{U(z)}{U(z_o)} \right)^{3/2} \cdot \frac{V(z)}{V(z_o)} \cdot \left(\frac{N(i, z_o)}{N(i, z)} \right)^{5/2} \right].$$

Así obtenemos un ejemplo concreto de una determinación entrópica completamente especificada.

En este punto debería estar claro ya que el procedimiento de obtener especializaciones termodinámicas especificando f^S (o f^U) de manera más o menos completa puede continuarse en muchas direcciones diversas, las cuales representan diferentes modos de aplicar el formalismo termodinámico a la explicación del comportamiento de

los sistemas simples. El gas ideal es sólo un ejemplo particularmente sencillo. Un caso más complicado es el de los llamados gases «reales», cuyo comportamiento se desvía considerablemente del de los gases ideales, en particular cuando el volumen disminuye en gran medida. Una especie de especialización «general» que se usa para estos casos es la llamada «ecuación virial de estado», o «expansión virial». La determinación entrópica también se puede expresar en este caso por medio de un funcional g^S con propiedades semejantes a las del caso del gas ideal, pero con una forma más compleja.

*D*24: x es una *termodinámica simple virial* ($x \in \widehat{TSV}$) si y sólo si: existen $i \in \mathbb{N}$, $R \in \mathbb{R}$, $z \in Z_x$, g^S, a_1, ..., $a_n \in \mathbb{R}$ tales que:

(1) $x \in \widehat{TGS}$ con $I_x = \{i\}$.

(2) g^S cumple *D*21-(2).

(3) $\forall z(z \in E \leftrightarrow f^S(U(z), V(z), N(i, z)) = \dfrac{N(i, z)}{N(i, z_o)} \cdot S(z_o) +$

$$+ N(i, z) \cdot g^S\left(\frac{U(z)}{N(i, z)}\right) + N(i, z) \cdot R \cdot \left[\log_n\left(\frac{V(z)}{V(z_o)} \cdot \frac{N(i, z_o)}{N(i, z)}\right) - \right.$$

$$\left. - \left(\frac{a_1}{V(z)} + \frac{a_2}{2 \cdot V(z)^2} + \ldots + \frac{a_n}{n \cdot V(z)^n}\right)\right].$$

Los valores a_1, ..., a_n, son los llamados «coeficientes viriales». Para simplificar la exposición se pueden tomar como constantes, tal como se los usa a veces en presentaciones no muy exactas de esta especialización; en una presentación más estricta, sin embargo, deberían aparecer como funcionales de la temperatura. Cuanto más aumentemos el número de estos coeficientes, tanto mejor será la aproximación que obtendremos a los resultados experimentales.

Mediante un simple cálculo obtendríamos a partir de la ecuación de *D*24-(3) la forma usual de la expansión virial:

$$\frac{P(z) \cdot V(z)}{R \cdot T(z)} = 1 + \frac{a_1}{V(z)} + \frac{a_2}{V(z)^2} + \ldots + \frac{a_n}{V(z)^n}.$$

Cuando tomamos $n = 1$, es decir, un solo coeficiente virial, y estipulamos $a_1 = b - \dfrac{a}{R \cdot T(z)}$ para ciertas constantes especificadas a y b, obtenemos una aproximación de la tan conocida y utilizada ley de Van der Waals:

$$\left(P(z) + \frac{a}{V(z)^2}\right) \cdot (V(z) - b) = R \cdot T(z).$$

Podemos considerar que esta ecuación correspondería a otra especialización particular, llamémosla *TSVW*, de la especialización virial general *TSV*.

Otro tipo de especialización de T.S.S. se sugiere al considerar la llamada «radiación de un cuerpo negro» desde un punto de vista termodinámico. Este caso es metodológicamente de especial interés, porque con él obtenemos inmediatamente una especialización terminal muy sencilla de la ecuación fundamental. Para expresarla es más conveniente utilizar la representación energética en vez de la entrópica. La ley especial que así se obtiene se llama a veces «ley de Stefan».

D25: x es un *cuerpo negro* ($x \in \widehat{CN}$) si y sólo si existen $i \in \mathbb{N}$, $a \in \mathbb{R}$, tales que:

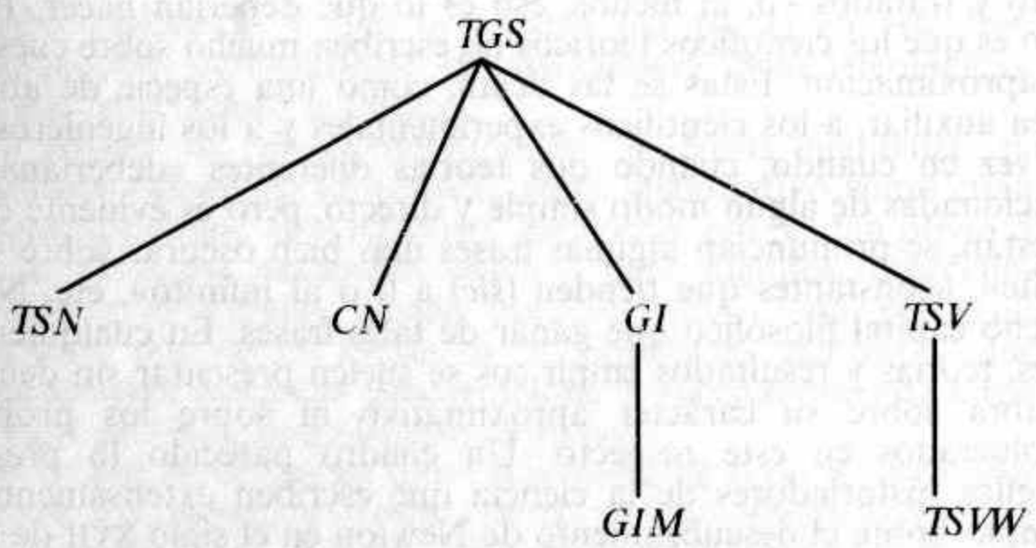
(1) $x \in \widehat{TGS}$, con $I_x = \{i\}$.
(2) $\forall x \in Z$: $N(i, z) = 0$.
(3) $\forall z \in Z$: $f^U(S(z), V(z), N(i, z)) = a \cdot V(z) \cdot (D_S f^U(z))^4$.

La condición *D25*-(2) expresa que la radiación de un cuerpo negro tiene lugar en el vacío: no se toma en consideración la masa de ninguna sustancia química. Puede comprobarse fácilmente que la condición *D25*-(3) es equivalente a la forma usual de la ley de Stefan:

$$\frac{U(z)}{V(z)} = a \cdot T(z)^4 .$$

Si resumimos esquemáticamente las interconexiones de todas las especializaciones de T.S.S. consideradas hasta aquí obtenemos el siguiente gráfico arbóreo:

La red de la termodinámica de sistemas simples:



2.7. APLICACION APROXIMATIVA DE TEORIAS EMPIRICAS

0. Introducción

La función y estructura de la aproximación en las ciencias empíricas es un campo desatendido por la filosofía de la ciencia. No conozco ni siquiera una relación sistemática de las diferentes clases de problemas que surgen en este campo. Todavía hay que hacer el esfuerzo por plantear claramente los problemas que se supone que hay que resolver aquí.

La responsabilidad por esta negligencia no debería adscribirse a los filósofos únicamente. A los científicos teóricos les corresponde parte de la culpa. Los filósofos de la ciencia normalmente toman en consideración lo que los científicos teóricos escriben en sus libros de texto y tratados –o, al menos, eso es lo que deberían hacer. Pero el caso es que los científicos teóricos no escriben mucho sobre cuestiones de aproximación. Estas se las dejan, como una especie de aburrida tarea auxiliar, a los científicos experimentales y a los ingenieros. Sólo de vez en cuando, cuando dos teorías diferentes «deberían» estar relacionadas de algún modo simple y directo, pero es evidente que no lo están, se pronuncian algunas frases más bien oscuras sobre «casos límite», «constantes que tienden (*sic*) a 0 o al infinito», etc. No hay mucho capital filosófico que ganar de tales frases. En cualquier caso, leyes, teorías y resultados empíricos se suelen presentar sin decir una palabra sobre su carácter aproximativo ni sobre los problemas involucrados en este respecto. Un cuadro parecido lo presentan aquellos historiadores de la ciencia que escriben extensamente, por ejemplo, sobre el descubrimiento de Newton en el siglo XVII de que el

sistema planetario «obedece» la ley de la gravitación, pero generalmente se olvidan de relatarnos los penosos esfuerzos hechos durante el siglo XVIII y posteriormente por establecer fuera de toda duda que la ley de Newton realmente encaja con las observaciones astronómicas «con bastante buena aproximación».

Naturalmente, todo el mundo sabe –o debería saber– que las leyes y teorías científicas sólo se pueden aplicar a la realidad con cierto grado de aproximación, y que la competencia entre diversas leyes, teorías y «programas de investigación» no es, en muchos casos, más que una competencia por lograr un más alto grado de aproximación –y no la «verdad», cualquiera que ésta sea. Sin embargo, hasta época muy reciente la inmensa mayoría de las exposiciones sistemáticas en filosofía de la ciencia han tomado esto como un rasgo accidental y poco interesante de la empresa científica, algo que no pertenece realmente a la «esencia» de la ciencia. Considero que esta visión de las cosas es básicamente errónea. Es un error dar por supuesto que la situación canónica de la que el filósofo debe dar cuenta es aquella en que una teoría científica con conceptos perfectamente claros encaja perfectamente bien con los hechos que pretende sistematizar. En la actualidad no existe ninguna teoría importante en la que esto realmente ocurra, y no veo ninguna razón para suponer que esto llegue a suceder algún día. Especular sobre esta posibilidad me parece ser una pérdida de tiempo. Y proponer una ciencia empírica sin aproximación como el ideal que deben alcanzar futuras generaciones de científicos no es más que parte de la teleología especulativa que Stegmüller ha descrito como «el mito teleológico de la verosimilitud creciente»: la idea de que hay una única meta absoluta que debe ser alcanzada por la ciencia algún día en el futuro (cf. Stegmüller [1978], pp. 63 y ss.). Desgraciadamente, el término «aproximación» en sí mismo contribuye a sustentar el mito teleológico de que la ciencia se acerca cada vez más a una Meta Absoluta que nos está esperando en algún lugar. Pero la elucidación de las relaciones aproximativas en la que estoy pensando y que se expone (en parte) en este libro es completamente independiente de cualesquiera ideas teleológicas que se puedan tener sobre la ciencia.

No parece ser una generalización temeraria mantener que el conocimiento científico en general, tanto cuantitativo como cualitativo, es esencialmente aproximativo. De hecho, esto es tan obvio que a nadie le preocupa realmente mucho. Para algunos, especialmente científicos profesionales, puede aparecer como una perogrullada. Ahora bien, una de las tareas de la filosofía es analizar y elucidar cuidadosamente supuestas perogrulladas. Y no creo que esta tarea sea, en el caso presente, completamente trivial.

Por lo que sé, el primer autor que tomó en serio la aproximación como una característica *esencial* de las teorías empíricas fue Ludwig en su *Deutung des Begriffs 'physikalische Theorie'* y en escritos

posteriores. El segundo capítulo de dicho libro (reproducido en su posterior *Grundstrukturen*) incluye un esbozo de una elucidación semiformal de las estructuras de una teoría física que contenga un elemento aproximativo como parte constitutiva. Ciertamente, Ludwig sólo se ocupa de teorías físicas. Sin embargo, su elucidación está enunciada en términos tan generales que debería ser aplicable a cualquier teoría cuyos axiomas estuvieran expresados en «lenguaje matemático». Dado que la teoría de conjuntos puede incluirse en el lenguaje matemático, se sigue que la elucidación de Ludwig es relevante para prácticamente cualquier teoría empírica significativa.

Prescindiendo de la elucidación general ludwigiana del concepto de teoría, que no es nuestro tema ahora, el punto central que quisiera subrayar es su idea (implícita) de que cualquier elucidación del concepto de teoría empírica debe dar cuenta del componente aproximativo; de lo contrario, la elucidación será necesariamente incompleta. Dicho crudamente: un concepto realista de teoría empírica necesariamente incluye un concepto de aproximación. Creo que ésta es una idea filosófica importante. Fue una de las fuentes que inspiraron las ideas que se exponen a continuación, a pesar de que las formulaciones concretas que se darán de ellas no pertenecen al marco conceptual de Ludwig. El programa es, pues, definir conceptos y estructuras aproximativas apropiadas para ser incluidas en estructuras teóricas adecuadamente definidas.

Una reacción posible en contra de este programa sería la siguiente objeción. Al hablar de aproximación no nos estamos refiriendo a la teoría misma, sino a sus aplicaciones; las nociones aproximativas deberían separarse de las estructuras teóricas. Esta objeción se basa en el supuesto de que las aplicaciones de una teoría no pertenecen conceptualmente a la noción de teoría misma. Esto es, una teoría empírica debería ser tratada al igual que una teoría matemática, como una mera estructura formal (un «formalismo matemático») determinada mediante ciertos axiomas. En capítulos anteriores ya hemos expuesto por qué esta noción puramente «formalista» de teoría empírica es inadecuada. En nuestra concepción, una teoría empírica es una entidad compuesta *a la vez* de (por lo menos) una estructura conceptual formal y su dominio de aplicaciones. Las nociones aproximativas, por lo tanto, deberán elucidarse como algo que está en relación tanto con la estructura formal como con el dominio de aplicaciones de la teoría.

1. Diferentes clases de aproximación

En las observaciones introductorias anteriores se ha hablado sobre la aproximación en términos bastante vagos. Ha llegado el momento de ser más sistemático. En primer lugar, digamos que la forma lógica

del concepto de aproximación parece ser la de una relación diádica: algo es una aproximación de otra cosa. (Puede que en una reconstrucción formal nos veamos forzados a introducir más términos en la relación; pero entonces ya veremos.) También parece bastante obvio que la investigación científica emplea muchas clases diversas de aproximación, en contextos muy diversos. Pero quizá no sean tantos como parecen a primera vista. Tratemos de establecer una primera clasificación tipológica de los diversos casos de aproximación. Por lo menos en una primera exploración veo cuatro tipos principales.

(a) Está ante todo el caso de que tratamos de sistematizar ciertos datos empíricos dentro de un marco conceptual dado, y en ese proceso tenemos que hacer algunas «idealizaciones» y «simplificaciones» (léase aproximaciones) para obtener un modelo utilizable. Por ejemplo, aproximamos el movimiento de un cuerpo macroscópico considerándolo como una partícula sobre una trayectoria continua; o a partir de un número finito de mediciones geodésicas concluimos que la Tierra es un elipsoide (olvidándonos de todos sus montes y valles); o consideramos un rayo visible de luz en nuestra habitación como una línea recta geométrica; y así sucesivamente. Este nivel de aproximación (o «idealización», si se prefiere) puede denominarse «construcción de modelos». Constituye las raíces de la ciencia, pero no nos vamos a ocupar de él aquí. La razón es simplemente que no tengo nada que decir acerca de su estructura formal[30].

(b) En el nivel siguiente aplicamos aproximativamente alguna ley o teoría a un «modelo construido», es decir, tratamos de subsumir una colección de datos conceptualmente sistematizados bajo alguna ley o teorías propuestas. Por ejemplo, tratamos de hacer plausible que un cuerpo que está moviéndose en nuestra cercanía es un caso de la ley de Galileo, o que el movimiento de la Tierra alrededor del Sol obedece la ley de gravitación, o que podemos aplicar la ley de reflexión a un rayo (idealizado) de luz solar.

(c) Cuando consideramos que una ley es una aproximación de otra ley más complicada de la misma teoría (es decir, suponiendo que las dos leyes pertenezcan al mismo marco conceptual), tenemos una relación aproximativa a un nivel puramente teórico. En tales casos no nos preocupamos mucho por los hechos empíricos externos a la teoría: estamos solamente practicando matemáticas dentro de la teoría. Esta comparación de leyes es muy corriente en física. Por ejemplo: la ley de Galileo se toma como una aproximación de la ley de la gravitación; o la ley de los gases ideales se toma como «aproximadamente válida» con respecto a la ley de Van der Waals, y

[30] Un primer paso hacia un análisis lógico de este nivel lo ha dado Suppes en su artículo «Models of Data», 1962.

ésta a su vez como una aproximación de una expresión aún más complicada, la ecuación virial de estado; o bien la ley del péndulo simple:

$$\tau = 2 \cdot \pi \cdot \sqrt{\frac{l}{g}}$$

es una aproximación de

$$\tau = 2 \cdot \pi \cdot \sqrt{\frac{l}{g}} \cdot \left(1 + \frac{1}{4} \cdot \operatorname{sen}^2 \frac{\alpha}{2}\right),$$

la cual, a su vez, es una aproximación de una serie potencial. Se pueden encontrar muchos ejemplos de esta clase en cualquier libro de texto de física. Es frecuente que en tales casos la ley más simple o «menos exacta» se derive de la más complicada al sustituir un parámetro (una función constante o real) que aparece en la ley aproximada por un valor numérico fijo (por ejemplo, 0 ó 1). En el caso de la ecuación de Van der Waals:

$$\left(p + \frac{a}{V^2}\right) \cdot (V - b) = R \cdot T;$$

ésta se reduce a la ley de los gases ideales al tomar $a = b = 0$. Naturalmente, no todos los casos son tan simples. Pero generalmente el «truco aproximativo» consiste en cambiar (y fijar) el valor de algún parámetro de la ley.

Se habla aquí de aproximar una ley dada mediante otra ley dada. Pero, en muchos casos, las expresiones que se aproximan o que son aproximadas no merecen realmente el título de «leyes». No aparecen en las exposiciones sistemáticas de la teoría, son simplemente instrumentos de cálculo que se introducen con algún propósito particular, frecuentemente propósitos de aproximación. Formalmente no hay una distinción tajante entre leyes e instrumentos especiales de cálculo; pero a veces realmente no quisiéramos considerar como ley a alguna ecuación especial. Considérese el caso de la sistematización del movimiento de un vehículo a motor en el que se toma la masa como dependiente del tiempo transcurrido (se supone que se pierde masa del combustible). ¿Acaso quisiéramos considerar como ley física la ecuación dinámica resultante?

Como convención terminológica usemos en general el término «sistematizaciones teóricas» de estados de cosas empíricos descritos dentro de cierto marco conceptual previamente dado. Propongo este rótulo tanto para las leyes «serias» como para los instrumentos de

cálculo introducidos *ad hoc*, para aplicaciones particulares de la teoría. En ejemplos del tipo de los que se acaban de describir, diremos que una sistematización teórica es una aproximación de otra. (En lo que sigue abreviaremos «sistematización teórica» por «s.t.».)

Como ya se ha observado, la aproximación entre s.t.s. es, en cierto sentido, una empresa puramente teórica, una operación de papel y lápiz, como diría P. W. Bridgman. No obstante, este hecho no nos impide admitir que hay una relación íntima entre esta clase de aproximación y la aproximación del segundo tipo antes mencionado, es decir, la aplicación aproximativa de una s.t. a un modelo empírico construido. En efecto, supóngase que hemos aplicado alguna ley L_1 a un dominio empírico D con cierto éxito, pero aún no estamos muy satisfechos con el grado de aproximación obtenido; podemos intentar construir una segunda ley L_2 tal que L_1 resulte ser una aproximación de L_2 y de modo que L_2 aplicada a D dé un grado mayor de aproximación. O, a la inversa, L_1 se aplica «bastante bien» a D, pero los cálculos con L_1 son muy difíciles: nos gustaría disponer de una manera más simple, aunque más burda, de sistematizar D. Por ello buscamos una L_0 que, aun siendo una aproximación más simple a L_1, puede aplicarse a D con un grado de aproximación que es suficiente para nuestros propósitos. Más abajo daremos una descripción más formal de esta relación entre ambos niveles de aproximación.

(d) Finalmente, existe un tipo de relación aproximativa al máximo nivel teórico que no parece ser reducible a la relación aproximativa entre s.t.s. Se trata de la aproximación entre dos teorías generales, es decir, entre estructuras globales con un marco conceptual diferente y con distintas leyes fundamentales. Llamémosla «*aproximación interteórica*». (Los tipos (b) y (c) podrían describirse como «aproximación *intra*teórica», mientras que (a) sería una «aproximación *pre*teórica».) Es probable que la relación de aproximación interteórica se dé entre la mecánica clásica y la relativista especial, entre la termodinámica fenomenológica y la mecánica estadística, entre la óptica geométrica y la electrodinámica. A estos ejemplos se les suele subsumir bajo el rótulo «reducción». Pero parece bastante claro que estos casos no pueden subsumirse bajo la reducción *exacta*, tal como se ha venido entendiendo esta relación en la literatura estándar (incluyendo las reconstrucciones formales más elaboradas de Adams [1959] y de Sneed en su *Logical Structure*). Quizá se pueda hablar aquí de «reducción aproximativa», pero en todo caso habría que justificar (tanto en términos generales como ante ejemplos concretos) que hay alguna conexión entre reducción y aproximación interteórica. Esta cuestión la trataremos con detalle en 2.9.

El problema de proporcionar una reconstrucción formal de la aproximación interteórica en general apenas ha sido atacado. El caso concreto de la reducción de las leyes de Kepler a la teoría de la

gravitación de Newton (considero esto como un ejemplo de aproximación interteórica, no intrateórica) ha sido analizado cuidadosamente por E. Scheibe en su artículo de 1973. La reconstrucción que emprende Scheibe de este caso es estimulante. Sin embargo, es dudoso que haya conseguido realmente filtrar un concepto general a partir de este ejemplo particular, como lo pretende en su trabajo posterior de 1975. Existe otra propuesta en esta dirección, a saber, la de Ludwig en la obra antes mencionada. Ludwig primero define un concepto de «*Einbettung*» («encaje») de una teoría dentro de otra, y a continuación establece algunas condiciones para «emborronar» esta noción y definir una «*unscharfe Einbettung*» («encaje aproximativo»). En el artículo que acabamos de mencionar de Stegmüller, éste señala el prometedor programa que se halla detrás de esta noción; pero la elucidación de Ludwig es muy breve y esquemática. En 2.9 se introducirá de manera general y detallada un concepto de aproximación interteórica que guarda cierto parecido con el de Ludwig y se aplicará al caso concreto de Kepler-Newton.

En esta sección nos vamos a restringir a discutir las categorías (b) y (c) de aproximación, es decir, las nociones de aplicación aproximativa de una teoría y de relación aproximativa entre s.t.s.

2. Conjuntos borrosos definidos modelo-teóricamente

Ha llegado el momento de ser más precisos con respecto a la categoría lógica de las entidades que hay que comparar entre sí en una relación aproximativa de los tipos (b) y (c).

Una primera idea que se nos podría ocurrir sería tomar *enunciados* como las entidades que hay que comparar. En un caso del tipo (b) diríamos que el enunciado de una ley empírica (por ejemplo, la ley de la gravitación) se aproxima a un enunciado que describe un sistema físico en términos matemáticos (por ejemplo, las órbitas planetarias descritas en términos de la mecánica de partículas). Esto sería la forma exacta del *explicandum* que hay que elucidar.

Ahora bien, este *explicandum* no parece ser un candidato muy adecuado. ¿Podemos realmente dar un sentido preciso a la aseveración de que un enunciado es la aproximación de otro enunciado? Es muy probable que, al intentarlo, cayéramos en un pantano semántico.

Naturalmente, un modo posible de proporcionar un análisis semántico de una «aproximación entre enunciados» sería en términos de las relaciones entre los *modelos* que satisfacen dichos enunciados. Esto es, de hecho, parecido al camino que efectivamente se propondrá aquí. Ahora bien, en tal caso, si elucidamos la noción de aproximación entre enunciados como relación modelo-teórica, nos podemos olvidar de una vez de la «aproximación entre enunciados», y trabajar solamente con modelos. En lo que sigue, voy a tratar de mostrar que esto es posible.

La idea fundamental es no tomar enunciados, sino modelos definidos conjuntistamente como las entidades mínimas que hay que comparar en una relación aproximativa. Este enfoque de la cuestión concuerda, naturalmente, con la concepción estructural general que hemos expuesto y utilizado en páginas anteriores de este libro. Pero, al menos en sus primeros pasos, la elucidación presente de la aproximación también podría encajar dentro del enfoque más tradicional de Suppes y sus colaboradores.

En efecto, todo lo que necesitamos de momento es admitir que se puede axiomatizar una teoría por definición de un predicado conjuntista, que las entidades que satisfacen ese predicado son los modelos actuales de la teoría, y que conviene hacer una distinción entre éstos y las estructuras conjuntistas de las que sólo exigimos que cumplan los axiomas estructurales, es decir, los llamados modelos potenciales. Recordemos brevemente el sentido de esta última noción. En el caso de una teoría empírica, un modelo potencial es una descripción matemática posible dentro del marco de *esta* teoría de un estado de cosas. Es una descripción *posible* porque, al aplicar el marco conceptual de la teoría a ese estado de cosas, aún no hemos decidido la cuestión de si esa descripción realmente produce un modelo de la teoría o no. Cuando aparecen funciones métricas entre las primitivas de la teoría son incontables las distintas descripciones posibles que se pueden dar del estado de cosas en cuestión. A esta situación ya nos hemos referido en capítulos anteriores. El punto que se trata de subrayar ahora, no obstante, es que entre todas esas descripciones posibles que pueden corresponder a un estado de cosas dado, algunas se considerarán más semejantes entre sí que otras; y el «grado de semejanza» entre distintas descripciones matemáticas posibles de los hechos puede resultar ser una cuestión importante que hay que considerar cuando tratamos de aplicar la teoría a los hechos en cuestión.

Por ejemplo, cuando tratamos de aplicar una teoría dinámica al movimiento de cierto sistema físico S consistente en un conjunto de partículas P consideradas durante cierto intervalo T y mediante una función de posición s, una pregunta importante que hay que hacerse es la de qué grado de semejanza entre dos descripciones posibles del sistema S vamos a requerir para que ambas sean aceptables desde el punto de vista de la teoría. Una respuesta posible a esta pregunta puede darse en los siguientes términos. Si x e y son dos descripciones posibles de S (=dos modelos potenciales), estaremos satisfechos con su semejanza si satisfacen la siguiente condición para ciertos números reales ε y δ previamente dados:

$$[S] \quad P_x = P_y \wedge T_x = T_y \wedge \forall p \in P_x \, \forall t \in T_x: \\ /s_x(p, t) - s_y(p, t)/ < \varepsilon \wedge /D^2 s_x(p, t) - D^2 s_y(p, t)/ < \delta.$$

En tal caso diremos que x es una aproximación de y y a la inversa. Nótese que para enunciar esta condición de aproximación no se necesita ningún axioma propio (ninguna ley empírica) de la teoría. En nuestra terminología esto significa que estamos trabajando solamente al nivel de M_p.

El presente intento hacia una elucidación general de las nociones aproximativas se basa principalmente en la propuesta de tomar la aproximación como una relación entre dos modelos potenciales de una teoría dada, es decir, como una relación diádica sobre M_p. Las características concretas de la teoría son totalmente irrelevantes para la realización de este programa. No es preciso suponer que se trata de una teoría «cuantitativa» (es decir, una teoría con funciones métricas) ni que la aproximación se basa en ciertos métodos especiales de medida (como se presupone a veces cuando se habla sobre la aproximación). El nivel de generalidad de la presente elucidación es tal que igualmente puede aplicarse a teorías muy crudas o cualitativas. El único supuesto que requerimos es que podamos axiomatizar la teoría y distinguir los axiomas estructurales de los propios.

Ahora estamos en posición de enunciar claramente el *explicandum* que buscamos. Es un enunciado metateórico del tipo: «el modelo potencial x es una aproximación del modelo potencial y de la teoría T». ¿Cuál es la vía que conduce al *explicans*? El punto de vista conjuntista nos ayudará en esta cuestión.

Los modelos potenciales de una teoría T son todos, como sabemos, elementos de la misma clase M_p de T. Esto significa que queremos definir una relación de aproximación entre dos elementos (o «puntos») cualquiera de una clase dada. Existe un método bien conocido en el análisis topológico para definir una relación de tal tipo: introducir el concepto de *uniformidad* como especialización de un filtro. Ciertamente, las aplicaciones propuestas de esta noción en topología son conjuntos de números reales. Pero, ¿acaso hay alguna razón convincente para no aplicar esa noción a clases de entidades diferentes, si es conveniente? No veo ninguna razón de este tipo, suponiendo que estemos dispuestos a aceptar que (a) el concepto topológico de estructura uniforme es la elucidación adecuada de la idea intuitiva de aproximación en conjuntos de puntos, y (b) la clase de los modelos potenciales de una teoría axiomatizada por medio de un predicado conjuntista es una entidad bien definida. Me parece que ambos supuestos son perfectamente aceptables. Por lo tanto, reconstruiremos la aproximación modelo-teórica definiendo una estructura uniforme sobre los modelos potenciales[31].

[31] La idea de elucidar la aproximación en las teorías empíricas por medio de uniformidades se retrotrae a Ludwig. Pero las entidades aproximadas en el enfoque de Ludwig no corresponden a nuestros modelos potenciales: son tuplos de términos definidos de varias maneras. Considero que el presente uso de las uniformidades es más simple y homogéneo que el de Ludwig.

Antes de pasar a la definición formal de una uniformidad modelo-teórica discutamos la idea intuitiva que está detrás de ella. Dada una clase cualquiera (en el caso presente, M_p), una uniformidad en ella determina toda una serie de subconjuntos, cada uno de los cuales representa un cierto «grado de aproximación» o «medida de emborronamiento». Vamos a denominar a estos subconjuntos «conjuntos borrosos». Un conjunto borroso consiste de pares de elementos de la clase original (M_p). Si un par $\langle a, b\rangle$ se halla en un conjunto borroso u, esto significa que a y b se aproximan entre sí *por lo menos* en el grado dado por u. También podríamos decir: a y b coinciden al menos en el grado u. O incluso: a y b son u-aproximadamente iguales. En el caso de un conjunto de números reales, en el que tenemos una métrica estándar (el valor absoluto de la diferencia), podemos definir la uniformidad estándar en la que cada conjunto borroso viene determinado por un ε particular:

$$u_\varepsilon = \{\langle a, b\rangle / |a-b| < \varepsilon\}.$$

Pero esto sólo es un caso especial. El concepto de uniformidad no depende del concepto de métrica, es más general. La determinación de conjuntos borrosos empíricamente relevantes puede ser mucho más complicada que en el ejemplo anterior, o quizá ni siquiera expresable en términos métricos.

Un concepto general de esta clase es precisamente lo que necesitamos para elucidar la comparación entre modelos potenciales. Para un tratamiento general de la aproximación no tenemos por qué preocuparnos por los términos específicos que entran en los modelos (principalmente funciones) y que aparecen involucrados en la aproximación. Aparte de su brevedad, el tratamiento abstracto propuesto tiene otras ventajas. Se puede fijar el mismo grado de aproximación, es decir, el mismo conjunto borroso u tomando la aproximación sobre funciones diferentes de la teoría. Por ejemplo, en un estudio gravitatorio del sistema Tierra-Sol podemos enunciar la condición de aproximación considerando $1/m_{\text{Sol}} \approx 0$, o bien, alternativamente, considerando que la fuerza que la Tierra ejerce sobre el Sol es «muy pequeña»; se obtiene el mismo grado de aproximación tanto al considerar masas como al considerar fuerzas. Aún más importante es el hecho de que en un tratamiento general no necesitamos asumir que los conjuntos borrosos usados en ciertas aplicaciones de la teoría están determinados por ninguna lista completamente especificable de condiciones: puede ocurrir que una medida específica de vaguedad dependa del «conocimiento personal» del científico, de sus «intuiciones», de sus necesidades pragmáticas cambiantes, etc. Todo lo que necesitamos saber es que los conjuntos borrosos usados son elementos de una uniformidad asociada al M_p de la teoría. Introducimos cierta notación especial para la definición formal de una uniformidad.

Si C es un conjunto, entonces $\Delta(C)$ es la «diagonal» de C, es decir, el conjunto de todos los pares de elementos idénticos de C.

Si R es una relación diádica, R^{-1} es la inversa, y

$$R \circ R =: R^2 =: \{\langle x, y\rangle / \exists z(\langle x, z\rangle \in R \wedge \langle z, y\rangle \in R)\}.$$

R^n puede definirse análogamente.

Los siguientes axiomas corresponden esencialmente a los establecidos por N. Bourbaki en su *Topologie générale*, p. 131.

*D*1: U es una uniformidad sobre M_p si y sólo si:

(1) $\varnothing \neq U \subseteq \mathscr{P}(M_p \times M_p)$.
(2) $\forall u_1, u_2(u_1 \in U \wedge u_1 \subseteq u_2 \rightarrow u_2 \in U)$.
(3) $\forall u_1, u_2(u_1 \in U \wedge u_2 \in U \rightarrow u_1 \cap u_2 \in U)$.
(4) $\forall u(u \in U \rightarrow \Delta(M_p) \subseteq u)$.
(5) $\forall u(u \in U \rightarrow u^{-1} \in U)$.
(6) $\forall u_1 \exists u_2(u_1 \in U \rightarrow u_2^2 \subseteq u_1 \wedge u_2 \in U)$.

A los elementos de U los llamaremos «conjuntos borrosos (uniformes)». Nótese que un conjunto borroso es una relación diádica sobre M_p.

Los tres primeros axiomas definen esencialmente un filtro (con una pequeña variación en el axioma (1) -cf. Bourbaki, *op. cit.*, 32). Los axiomas subsiguientes (4)-(6) son específicos del concepto de uniformidad.

Vamos a tratar de interpretar esta estructura formal en términos de la aproximación modelo-teórica para ver si aparece como intuitivamente adecuada. En primer lugar, nótese que la razón pragmática para asociar una uniformidad U a la clase M_p de una teoría T es que eso permite liberalizar la aplicación o manejo de T. Si supusiéramos que la teoría se aplica exactamente a su dominio, o que las relaciones entre sus leyes especiales son siempre de tipo exacto, nos tendríamos que enfrentar a una falsación rápidamente. Formalmente, esto significaría que sólo admitimos el uso de la diagonal (como conjunto borroso «degenerado») en el manejo de la teoría. El uso de la diagonal representa exactitud absoluta. Pero los científicos reales son más liberales que eso. Usan conjuntos borrosos mucho mayores en su manejo de la teoría. El tamaño de los conjuntos borrosos usados depende de las diferentes clases de aplicaciones y manipulaciones de la teoría. Cuanto mayor sea el conjunto borroso, tanto más fácil será evitarse problemas. Naturalmente, podemos suponer que si se excede un cierto límite, la teoría es considerada inútil. Pero éste es un punto que ignoraremos de momento; lo discutiremos en el siguiente apartado.

A modo de lema para interpretar rápidamente los axiomas de (*D*1) digamos lo siguiente: asociar una uniformidad a una teoría es una manera de inmunizarla frente a posibles problemas en sus aplicaciones. (Los límites de esta inmunización se discutirán más adelante.) Supóngase que hemos construido la uniformidad U para la teoría T, y que estamos usando en particular cierto $u \in U$ para una aplicación dada de T. Esto significa que hemos «inmunizado» esta aplicación en el grado u. Por ejemplo, la aplicación puede consistir en lo siguiente. (Esto es sólo un caso posible.) Dado un sistema S, hacemos ciertos cálculos teóricos usando T y predecimos que S tendrá la forma matemática descrita por el modelo potencial a. Pero luego, después de ciertas «observaciones empíricas», llegamos a la conclusión de que es mejor describir S mediante el modelo potencial b, donde $a \neq b$. Si $\langle a, b \rangle \in u$, nos quedamos satisfechos, al menos por el momento. Si $\langle a, b \rangle \notin u$, podemos tener dos reacciones posibles: o bien ser aún más liberales y admitir que, después de todo, también quedaríamos satisfechos con un $u' \in U$ tal que $u \subset u'$, $\langle a, b \rangle \in u'$; o bien reconsiderar T y/o la descripción matemática de S. Ambos tipos de reacciones aparecen abundantemente ejemplificadas en la ciencia cotidiana.

Cuantos más pares $\langle a, b \rangle$ admitamos en el u utilizado, tanto más seguros estaremos; y, naturalmente, tanto menos trabajo efectivo podremos llevar a cabo con la teoría.

Prescindiendo del axioma puramente estructural *D*1-(1), «traduzcamos» los demás al «lenguaje de la inmunización». *D*1-(2) afirma que si la teoría está inmunizada por el conjunto borroso u_1, también lo estará por un u_2 mayor que incluya a u_1. *D*1-(3) afirma que si inmunizamos T tanto por medio de u_1 como por medio de u_2, podemos tomar ambos a la vez e inmunizar T mediante su intersección. Este axioma quizá aparezca menos obvio que *D*1-(2), pero es bastante plausible, al menos como idealización. (Nótese que la intersección no puede ser vacía: por *D*1-(4) siempre contiene por lo menos la diagonal.) Si estamos satisfechos con cierto grado de aproximación u en una aplicación dada, sea cual sea ese grado, también estaremos satisfechos con una aplicación exacta en caso de que ésta ocurra casual y afortunadamente; esto es, la exactitud está incluida en cualquier conjunto borroso. Este es el contenido de *D*1-(4)[32]. El axioma *D*1-(5) es obvio en la interpretación presente: el orden de los modelos que integran un par cualquiera de un conjunto borroso carece de relevancia alguna para el «grado de inmunización» de la teoría. *D*1-(6) es un axioma crítico. Afirma que para cualquier conjunto borroso dado u_1 que inmunice T, se puede encontrar otro u_2

[32] No se asume que la diagonal en sí misma sea un conjunto borroso; esto no puede probarse en el presente sistema axiomático. Y creo que así está bien. Sería ciencia-ficción suponer que una teoría pueda ser aplicada alguna vez con absoluta exactitud.

más estricto que siga inmunizando la teoría; u_2 ha de ser al menos el doble de estricto que u_1.

Para una fácil evaluación de este último axioma considérese un ejemplo muy simple. Supóngase que los conjuntos borrosos modelo-teóricos de una descripción geodésica vienen determinados por una condición aproximativa con respecto a la función de posición de las «partículas» o «puntos»:

$$u_\varepsilon = \{\langle x, x'\rangle / \forall p \in P_x \cap P_{x'} (/s_x(p) - s_{x'}(p)/ < \varepsilon)\}.$$

Entonces, si u_ε ha de ser un conjunto borroso de una uniformidad, el axioma $D1$-(6) afirma la existencia de un u_δ del mismo tipo que u_ε, pero el doble de restrictivo, lo cual en este caso especial significa:

$$u_\delta = \{\langle x, x'\rangle / \forall p \in P_x \cap P_{x'} \left(/s_x(p) - s_{x'}(p)/ < \delta \leqslant \frac{\varepsilon}{2}\right)\}.$$

En este ejemplo particular, el axioma $D1$-(6) garantiza la posibilidad formal de determinar la posición de un p dado con exactitud creciente. Este axioma representa ciertamente una idealización de la práctica científica real, porque implica la posibilidad de lograr aplicaciones cada vez más restrictivas de la teoría. (Pero nótese que *no* implica que vayamos a alcanzar nunca Δ, es decir, la exactitud.)

Las propiedades matemáticamente interesantes de las uniformidades en general pueden encontrarse en cualquier libro de texto de topología avanzada. Para el estudio presente sólo necesitamos algunos corolarios simples.

Sea U una uniformidad y u un elemento cualquiera de ella.

$T1$: $\varnothing \notin U$.

Esto se sigue directamente de las condiciones $D1$-(1), (4).

$T2$: $u^n \subseteq u^{n+1}$.

Esto se sigue por inducción a partir de $D1$-(4) y por la definición de u^2 y de u^n en general.

$T3$: $u \in U \rightarrow \forall n(u^n \subseteq U)$.

Esto se sigue de $T2$, $D1$-(2).

3. *Conjuntos borrosos a dos niveles*

Hasta aquí el presente análisis de la aproximación no se ha basado en ningún enfoque particular en filosofía de la ciencia, aparte de los

supuestos muy generales (a) y (b) antes reseñados. Procedemos ahora a desarrollar ulteriormente el análisis de la aproximación dentro de la concepción estructural sneediana. Vamos a ver que este enfoque también es capaz de iluminar aspectos interesantes de la aproximación.

En lo que sigue no haremos uso de todas las complejidades del aparato estructural; por ejemplo, prescindiremos de las nociones de *ligadura* y *especializaciones*. Los constituyentes básicos que emplearemos aquí son: la distinción entre el nivel teórico y el no-teórico dentro de una teoría dada (entre modelos potenciales y modelos parciales)[33]; la noción de *elemento teórico* como par ordenado $\langle K, A \rangle$ (prescindiendo de comunidades científicas e intervalos históricos), donde K es el núcleo y A el dominio de aplicaciones propuestas; la *aserción empírica* de la teoría, resumida en el enunciado «holista» según el cual A puede subsumirse bajo K.

La concepción estructural, en la versión que de ella hemos presentado en capítulos anteriores, no contiene ningún componente aproximativo. Nuestro objetivo aquí es doble: primero, introducir un componente aproximativo en la noción estructural de teoría empírica de tal modo que concuerde con la distinción entre el nivel T-teórico y el T-no-teórico; segundo, proporcionar una reconstrucción formal de la versión aproximativa de la aserción empírica de una teoría. La primera parte del programa se lleva a cabo en esta sección; la segunda, en la siguiente.

Al introducir la distinción entre la clase de modelos potenciales M_p (nivel teórico) y la de modelos parciales M_{pp} (nivel no-teórico) se nos abre una nueva posibilidad para expresar la aproximación. En efecto, los modelos parciales no son más que descripciones matemáticas posibles de sistemas empíricos en un sentido enteramente análogo al de los modelos potenciales. La única diferencia consiste en que la descripción matemática se ofrece en este caso sin términos teóricos. Por tanto, debería estar claro que las mismas razones que nos llevan a hablar de mayor o menor «cercanía» o aproximación entre modelos potenciales, y de descripciones borrosas en vez de exactas, podrían aducirse para el caso de los modelos parciales, es decir, para el caso de descripciones no-teóricas. De hecho, ya hemos visto un ejemplo de aproximación enunciada en términos puramente no-teóricos (cinemáticos) en el caso de la mecánica de partículas (la condición [S], más arriba).

En consecuencia, es razonable expresar formalmente la aproximación no-teórica usando una uniformidad definida sobre M_{pp} de manera enteramente similar a lo que hemos hecho en el caso de M_p.

---

[33] Recordamos que existe una función epiyectiva de la clase de los modelos potenciales M_p a la de los parciales M_{pp}. A esta función la llamamos «*restricción*» y la simbolizamos por r.

Introducimos una clase relacional $V \subseteq \mathscr{P}(M_{pp} \times M_{pp})$ que satisface axiomas análogos a los de $D1$. A sus elementos v_i los llamamos «conjuntos borrosos no-teóricos».

Esto significa que, en principio, podemos asociar dos clases de uniformidades a una teoría empírica: una U para M_p y una V para M_{pp}. Nuestras intuiciones nos llevan a suponer que debería haber alguna conexión entre una uniformidad dada sobre M_p y la uniformidad «correspondiente» sobre M_{pp} o, dicho aún más intuitivamente: la aproximación al nivel M_p se hace en aras de la aproximación al nivel M_{pp}.

Para hacer algo más claras estas ideas intuitivas consideremos una vez más el ejemplo «paradigmático» de la teoría newtoniana de la gravitación y sus aplicaciones.

Sea DGP el predicado «dinámica gravitacional de partículas» que hemos reconstruido en 2.5. Recordemos que los modelos potenciales de esta especialización son tuplos del tipo $\langle P, T, s, m, f \rangle$. Supóngase ahora que queremos aplicar DGP al sistema físico que consiste en p_T (la Tierra) y p_S (el Sol). Construimos un $x_1 = \langle \{p_T, p_S\}, T, s, m, f \rangle$ y afirmamos $DGP(x_1)$. Esta aserción implica, entre otras cosas, que las fuerzas que mutuamente se ejercen la Tierra y el Sol son no-nulas:

$$-\sum_i f(p_T, t, i) = \sum_i f(p_S, t, i) \neq 0.$$

Dado que admitimos que las masas respectivas son finitas, esto implica, a su vez (por el Segundo Principio de Newton, que es una condición definitoria del predicado conjuntista DGP), que las aceleraciones respectivas de la Tierra y el Sol también son no-nulas:

$$D_t^2 s(p_T, t) \neq 0;\ D_t^2 s(p_S, t) \neq 0.$$

Si traducimos este cálculo teórico a una descripción del sistema en términos puramente cinemáticos, obtenemos un modelo parcial $r(x_1) \in M_{pp}$ que matemáticamente corresponde a dos elipses: una elipse grande para la Tierra y una elipse «muy pequeña» para el Sol.

Como la elipse solar es tan «pequeña», los libros de texto de física nos informan que, para ciertas aplicaciones, «es despreciable». Suponemos que el Sol es estacionario, y esto significa que construimos un x_o con la condición $\sum_i f(p_S, t, i) = 0$ en vez de

$$\left| \sum_i f(p_S, t, i) \right| = -G \cdot \frac{m(p_T) \cdot m(p_S)}{|s(p_T, t) - s(p_S, t)|^2}.$$

Por otro lado, continuamos haciendo la suposición de que:

$$\left|\sum_i f(p_T, t, i)\right| = G \cdot \frac{m(p_T) \cdot m(p_S)}{|s(p_T, t) - s(p_S, t)|^2}.$$

Esta «idealización» simplifica considerablemente los cálculos teóricos y la descripción cinemática; y esto a veces es una razón muy buena para admitirla –al menos a fines didácticos–. La traducción cinemática de x_o, $r(x_o) \in M_{pp}$, es entonces una única elipse para la Tierra con el Sol en uno de sus focos.

Naturalmente, el precio que debemos pagar por esta «idealización» (=aproximación) es que el modelo potencial x_o ya no es un modelo actual de *DGP*, sino un modelo de una s.t. que se aproxima, pero no es idéntica a *DGP*. En nuestro marco general, esto significa que la uniformidad que tomamos cuando usamos la teoría de la gravitación contiene un conjunto borroso u determinado por condiciones tales como la de que $\langle x_o, x_1 \rangle \in u$, y que efectivamente usamos u al menos algunas veces con algunos propósitos razonables.

u es un conjunto borroso teórico determinado por una condición acerca de fuerzas (o, alternativamente, acerca de masas). Pero puede ser «traducido» en un conjunto borroso no-teórico v determinado por una condición que establezca una comparación de un sistema de dos elipses con uno de una sola elipse; y entonces tenemos $\langle r(x_o), r(x_1) \rangle \in v$.

Una cuestión adicional que se plantea es la siguiente: ¿corresponde alguna de estas dos descripciones cinemáticas del sistema Tierra-Sol a alguna aplicación propuesta de *DGP*, es decir, a algún sistema «realmente observado»? La respuesta es claramente: no (al menos no en el año de 1980). Los «modelos empíricamente construidos» para el sistema Tierra-Sol no se identifican ni con $r(x_o)$ ni con $r(x_1)$. Las trayectorias reales de la Tierra y el Sol que se derivan de observaciones astronómicas cuidadosas son más complicadas que la descripción en términos de elipse-grande-con-elipse-pequeña, para no hablar ya de la descripción en términos de elipse-única. Sea i el modelo parcial correspondiente a la aplicación propuesta Sol-Tierra en 1980. Tenemos que $i \neq r(x_o)$, $i \neq r(x_1)$. Ello no obstante, a veces se usan x_o y x_1 porque se aproximan a i al menos en cierto grado. Nuevamente, esto significa formalmente que la teoría de la gravitación a veces utiliza conjuntos borrosos no-teóricos v_o y v_1 tales que $\langle i, r(x_o) \rangle \in v_o$ y $\langle i, r(x_1) \rangle \in v_1$, con $v_1 \subset v_o$. También podríamos decir que la teoría de la gravitación a veces utiliza conjuntos borrosos teóricos u_o, u_1, tales que, para cierta extensión teórica de i, o sea $e(i) \in M_p$, ocurre que $\langle e(i), x_o \rangle \in u_o$ y $\langle e(i), x_1 \rangle \in u_1$, con $u_1 \subset u_o$. Si queremos ser más exactos en la descripción teórica de i, podemos construir un modelo potencial x_2 que tome en cuenta las fuerzas ejercidas por los demás planetas sobre la Tierra y el Sol; esto es, tomamos un modelo x_2 de una s.t. que se

aproxime más a *DGP* al considerar las fuerzas planetarias externas que actúan sobre el sistema Tierra-Sol. La descripción cinemática que se deriva de x_2, $r(x_2) \in M_{pp}$, será más complicada que la descripción en dos elipses. A pesar de ello, naturalmente, x_2 no encajará exactamente con i; encajará con i sólo con una mayor aproximación que los modelos previos. Esto significa que escogemos un conjunto borroso u_2 con $\langle e(i), x_2 \rangle \in u_2$ y $u_2 \subset u_1 \subset u_o$; o, alternativamente, un v_2 con $\langle i, r(x_2) \rangle \in v_2$ y $v_2 \subset v_1 \subset v_o$. Y así sucesivamente.

La moraleja esencial que se puede sacar del análisis de este ejemplo sencillo se puede expresar como sigue. La aproximación entre s.t.s. al nivel plenamente mecánico de M_p induce una aproximación entre descripciones cinemáticas al nivel de M_{pp}. Esta última, a su vez, es necesaria para comparar las descripciones cinemáticas calculadas con la aplicación realmente propuesta que hemos obtenido como «modelo construido» a partir de los datos pre-teóricos. Dicho en términos más generales: las aproximaciones al nivel teórico inducen aproximaciones al nivel no-teórico que son necesarias para acercarse sucesivamente a una aplicación propuesta. Estas son, según creo, las características esenciales del mecanismo aproximativo (dentro de *una misma* teoría).

Sin embargo, desde un punto de vista general, no está garantizado que este mecanismo vaya a funcionar siempre. En principio, no está garantizado que una uniformidad dada sobre M_p induzca una uniformidad «correspondiente» en M_{pp}, a pesar de que intuitivamente nos parezca plausible esta idea. De hecho, no hay una prueba general al efecto de que esta inducción sea siempre posible. El axioma crítico *D*1-(6) es el que ocasiona las molestias en este caso: partiendo del concepto antes introducido de uniformidad, no se puede probar en general que una acotación cada vez más estrecha de los conjuntos borrosos teóricos induzca una acotación también cada vez más estrecha de sus correspondientes restricciones no-teóricas. La razón básica de esta imposibilidad radica en el hecho de que *muchos* modelos teóricos pueden corresponder a *un solo* modelo parcial no-teórico. Es fácil, aunque algo tedioso, comprobar que este hecho tiene el efecto mencionado. Dejo al lector interesado la comprobación de este punto.

No obstante, no deberíamos precipitarnos en abandonar nuestras esperanzas de establecer una correlación entre ambos niveles. Puede mostrarse que, si restringimos el concepto original de uniformidad mediante una condición plausible adicional, se obtiene el teorema de inducción deseado. A la uniformidad así modificada la llamaremos una «uniformidad empírica» –«empírica» en el sentido de que sólo parece tener sentido para teorías empíricas (es decir, teorías con una distinción entre conceptos teóricos y no-teóricos). Puede demostrarse entonces que una uniformidad empírica sobre M_p induce una uniformidad correspondiente sobre M_{pp}.

Para enunciar esta propiedad se necesitan algunas definiciones auxiliares. Sea U una uniformidad cualquiera sobre M_p y u_i sus conjuntos borrosos.

D2: $Rs(u_i) =: \{\langle y, y'\rangle / \exists x, x' (y = r(x) \wedge y' = r(x') \wedge \langle x, x'\rangle \in u_i\}$.

Vamos a denominar a $Rs(u_i)$ «la restricción» del conjunto borroso u_i. Nuestro objetivo puede describirse informalmente diciendo que queremos probar que la restricción de un conjunto borroso también es un conjunto borroso. De esta definición se siguen tres corolarios simples:

T4: $Rs(u_1 \cup u_2) = Rs(u_1) \cup Rs(u_2)$.
T5: $Rs(u_1 \cap u_2) \subseteq Rs(u_1) \cap Rs(u_2)$.
T6: $u_1 \subseteq u_2 \rightarrow Rs(u_1) \subseteq Rs(u_2)$.

D3: $V[U] =: \{v / \exists u (u \in U \wedge v = Rs(u))\}$.

Consideremos ahora pares de modelos potenciales que tengan exactamente los mismos componentes *T*-no-teóricos en el sentido estricto de que sus funciones *T*-no-teóricas toman exactamente los mismos valores. Sólo difieren, por tanto, en su parte *T*-teórica. De tales modelos podemos decir, intuitivamente, que son «empíricamente equivalentes», pues contienen o conducen a idénticos resultados empíricos. Puede considerarse que un par de modelos potenciales así describen exactamente el mismo estado de cosas posible; formalmente son distintos, pero en cuanto a su contenido empírico son iguales. ¿Qué podemos decir de tales pares de modelos desde el punto de vista de la aproximación? Si tomamos las estructuras matemáticas que los definen en su totalidad, no podemos decir nada en particular. Pero si restringimos nuestra atención a lo que podríamos llamar «aproximación empírica» (o *T*-no-teórica), es decir, a la parte empírica del proceso de comparación aproximativa, podemos decir intuitivamente que ambos modelos son cuasi-idénticos, o también «tan parecidos como se quiera» (desde el punto de vista empírico). Un mero cambio en la selección de las funciones *T*-teóricas que no influya en la determinación *T*-no-teórica del sistema físico no implica ningún cambio en el grado de aproximación. Dicho aún más precisamente: para todo $u \in U$ y para cualesquiera $x, x' \in M_p$, si $r(x) = r(x')$, entonces $\langle x, x'\rangle \in u$. A una uniformidad que cumpla este requisito la llamaremos una «uniformidad empírica». Para introducir este concepto formalmente, conviene definir previamente el concepto de *pseudo-diagonal* de una uniformidad. Se trata simplemente del conjunto de pares de modelos potenciales que tienen la propiedad susodicha.

*D*4: $\Psi(M_p) =: \{\langle x, x'\rangle / x, x' \in M_p \wedge r(x) = r(x')\}$.

Una uniformidad empírica será consiguientemente una uniformidad en la que cada uno de sus conjuntos borrosos incluye la pseudo-diagonal.

*D*5: U es una *uniformidad empírica* sobre M_p si y sólo si:

(1) U es una uniformidad sobre M_p.
(2) Para todo $u \in U$: $\Psi(M_p) \subseteq u$[34].

De manera enteramente análoga se podría definir una uniformidad empírica sobre M_{pp}. Pero es obvio que los conceptos de uniformidad empírica sobre M_{pp} y de su uniformidad sobre M_{pp} son coextensivos, dado que en este caso $r(y) = y$.

Pasemos ahora al «teorema de inducción».

*T*7: Si U es una uniformidad empírica sobre M_p, entonces $V[U]$ es una uniformidad (empírica) sobre M_{pp}.

Prueba: El supuesto es que U satisface los siete axiomas de una uniformidad empírica. Hay que probar que $V[U]$ también satisface los axiomas de uniformidad con respecto a M_{pp}.

(1) $\varnothing \neq V[U] \subseteq \mathcal{P}(M_{pp} \times M_{pp})$.

En efecto, por *D*2 y *D*3 es obvio que $V[U] \subseteq \mathcal{P}(M_{pp} \times M_{pp})$. Y por *D*1-(1) y *T*1 existe un $u \in U$ tal que $\varnothing \neq u$; de *D*3 se deduce que: $\varnothing \neq Rs(u) \in V[U]$.

(2) $v_1 \in V[U] \wedge v_1 \subseteq v_2 \rightarrow v_2 \in V[U]$.

En efecto, por el supuesto y *D*3, existe un $u_1 \in U$ con $v_1 = Rs(u_1)$. Sea $v_2 = v_1 \cup \{\langle y, y'\rangle, \ldots\}$, donde $y = r(x)$, $y' = r(x')$, para ciertos x, x'. Considérese $u_2 = u_1 \cup \{\langle x, x'\rangle, \ldots\}$.
$v_2 = Rs(u_2)$, por *D*2.
$u_2 \in U$, por *D*1-(2). Luego $v_2 \in V[U]$, por *D*3.

[34] Debo a Ignacio Jané la idea de utilizar la noción de pseudo-diagonal para definir las uniformidades empíricas. Jané me hizo ver que la definición de uniformidad empírica que yo había presentado en mi artículo de 1976 era en parte inadecuada por llevar a consecuencias trivializantes y que usando la noción de pseudo-diagonal captaba la misma idea intuitiva que yo adelanté en aquel artículo, sin arrastrar dichas consecuencias. Jané asimismo desarrolló un aparato formal más potente que le permitió llegar a resultados más elegantes de los que se han alcanzado aquí (cf. Jané [1980]). Sin embargo, por razones de simplicidad expositiva, aquí he conservado el marco conceptual previo. El lector interesado en las generalizaciones más técnicas de Jané puede acudir a su artículo.

(3) $v_1 \in V[U] \wedge v_2 \in V[U] \rightarrow v_1 \cap v_2 \in V[U]$.

En efecto, por el supuesto y $D3$, existen u_1, $u_2 \in U$, tales que $v_1 = Rs(u_1)$, $v_2 = Rs(u_2)$.
$\varnothing \neq u_1 \cap u_2 \in U$, por $D1$-(3), $T1$.
Esto significa que $\varnothing \neq Rs(u_1 \cap u_2)$. Además, por $T5$, $Rs(u_1 \cap u_2) \subseteq Rs(u_1) \cap Rs(u_2) = v_1 \cap v_2$.
Tómese $v_o =: Rs(u_1 \cap u_2)$.
Tenemos $\varnothing \neq v_o \subseteq v_1 \cap v_2$.
Por $D3$, $v_o \in V[U]$. Entonces, $v_1 \cap v_2 \in V[U]$ por la condición anterior, $D1$-(2) para $V[U]$, que ya hemos probado.

(4) $v \in V[U] \rightarrow \Delta(M_{pp}) \subseteq v$.

En efecto, por $D3$, existe algún $u \in U$ con $v = Rs(u)$.
Para todo $x \in M_p$: $\langle x, x \rangle \in u$, por $D1$-(4).
Para todo $x \in M_p$: $\langle r(x), r(x) \rangle \in Rs(u) = v$, por $D2$.
Esto significa precisamente que $\Delta(M_{pp}) \subseteq v$.

(5) $v \in V[U] \rightarrow v^{-1} \in V[U]$.

Esto se sigue trivialmente de $D1$-(5), $D2$, $D3$.

(6) $\forall v_1 \, \exists v_2 (v_1 \in V[U] \rightarrow v_2^2 \subseteq v_1 \wedge v_2 \in V[U])$.

Tómese cualquier $v_1 \in V[U]$.
Por $D3$ hay un $u_1 \in U$ con $Rs(u_1) = v_1$.
Por $D1$-(6) y $T2$ sabemos que habrá un $t \in U$ con $t \subseteq t^2 \subseteq u_1$.
Volviendo a aplicar $D1$-(6) a t, habrá un u_2 con $u_2 \subseteq u_2^2 \subseteq t$.
Obtendremos, pues, $u_2 \subseteq u_2^2 \subseteq t \subseteq t^2 \subseteq u_1$.
Sea $v_2 = Rs(u_2)$. Queremos probar que este v_2 es el que buscamos, es decir, que $v_2^2 \subseteq v_1$.
Sean dos pares cualesquiera conectados en v_2:

$$\langle y_1, y_2 \rangle \in v_2 \wedge \langle y_2, y_3 \rangle \in v_2.$$

Dado que, en tal caso, $\langle y_1, y_3 \rangle \in v_2^2$, queremos probar que $\langle y_1, y_3 \rangle \in v_1$.
Sean x_1, x_2, x_2', x_3, los correlatos T-teóricos de y_1, y_2, y_3, es decir: $y_1 = r(x_1)$, $y_2 = r(x_2) = r(x_2')$, $y_3 = r(x_3)$, con las relaciones $\langle x_1, x_2 \rangle \in u_2 \wedge \langle x_2', x_3 \rangle \in u_2$.
Dado que $\langle x_2, x_2' \rangle \in \Psi(M_p)$ y $\Psi(M_p) \subseteq u_2$ por $D4$-(2), obtendremos $\langle x_1, x_2' \rangle \in u_2^2 \subseteq t$.
Por otro lado, dado que $u_2 \subseteq u_2^2$, también $\langle x_2', x_3 \rangle \in u_2^2 \subseteq t$.
Obtenemos, pues, $\langle x_1, x_2' \rangle \in t \wedge \langle x_2', x_3 \rangle \in t$.
De donde, $\langle x_1, x_3 \rangle \in t^2 \subseteq u_1$.
Tomando otra vez las restricciones y aplicando $T6$,

$$\langle r(x_1), r(x_3) \rangle \in Rs(u_1), \text{ o sea } \langle y_1, y_3 \rangle \in v_1.$$

Con este teorema queda probado que una uniformidad empírica al nivel teórico siempre induce una uniformidad correspondiente al nivel no-teórico de una manera natural. En consecuencia, a fines de aproximación, es conveniente completar el concepto estructural de teoría mediante la noción de uniformidad empírica sobre M_p. Dicho formalmente, a cada elemento teórico representado por un par $\langle K, A\rangle$ hay que añadirle una uniformidad empírica U definida sobre el M_p de K, de modo que transformamos el elemento teórico en el triplo $\langle K, U, A\rangle$.

Si representamos la teoría total en forma de red teórica tal como se ha propuesto en 2.4, con un número más o menos grande (pero en cualquier caso mayor que 1) de elementos teóricos distintos, en principio tendríamos que cada elemento teórico de la red debería contener su propia uniformidad empírica; es decir, las uniformidades empíricas de los diversos elementos teóricos en principio no tendrían por qué coincidir, ni siquiera por qué tener entre sí una relación de orden cualquiera. Sin embargo, al menos en el caso de las redes arbóreas (las únicas que parecen interesantes para las teorías científicas), es plausible admitir que la uniformidad empírica del elemento básico se «transmitirá» al resto de la red, o sea que habrá una sola uniformidad empírica para toda la red (=para toda la teoría). En efecto, el aparato conceptual en este caso es el mismo en todas partes: M_p y M_{pp}. La uniformidad originaria adoptada estará definida sobre M_p y la uniformidad no-teórica inducida según $T7$ estará definida sobre M_{pp}. Por tanto, desde el punto de vista de la estructura conceptual, no hay razón alguna para cambiar de uniformidad al pasar de un elemento teórico a otro. En consecuencia, los únicos factores que podrían determinar un cambio de uniformidad serían, al nivel teórico, las distintas leyes especiales y ligaduras, y al nivel empírico, cierto número de aplicaciones propuestas. Se podría pensar que estas modificaciones en K y/o I efectivamente deberían traer consigo cambios en las uniformidades adoptadas, puesto que, por ejemplo, al añadir leyes especiales o restringir el campo de aplicación, cambiaremos nuestras expectativas sobre el grado de aproximación con que hay que aplicar la teoría al dominio propuesto: nos haremos más exigentes, o menos, esto dependerá de cada caso, sin que sea posible sentar una norma general. Todo esto es cierto. Ahora bien, cambiar nuestras expectativas sobre el grado de aproximación sólo significa que cambiaremos unos conjuntos borrosos por otros, usaremos por ejemplo conjuntos borrosos más pequeños o más grandes; pero ello no implica que cambiemos de uniformidad. La uniformidad, que por principio está definida sobre M_p o M_{pp} sin tomar en cuenta leyes ni aplicaciones, no tiene por qué variar; sólo los elementos de ella que escojamos serán distintos según el caso.

El argumento que se acaba de exponer es sólo de plausibilidad, no pretende ser una prueba de que en toda teoría avanzada necesaria-

mente habrá una sola uniformidad empírica. Quizá investigaciones ulteriores muestren que, en algunos casos concretos de redes, es más conveniente o incluso necesario cambiar de uniformidad al pasar de un elemento teórico a otro. Ahora bien, de momento al menos, no veo ningún síntoma de que esto sea realmente el caso, ni ninguna razón a priori en favor del cambio de uniformidad. Por ello, en lo que sigue se partirá del supuesto de que hay una sola uniformidad empírica para toda la red teórica.

4. La versión aproximativa de la aserción empírica de una teoría

Un aspecto ya analizado del enfoque estructural es la reconstrucción del contenido empírico de una teoría en un período dado en la forma de una única proposición «holista» que afirma que la teoría como totalidad es aplicable a su dominio de aplicaciones propuestas (como totalidad). En 2.4 simbolizamos esta aserción así: $I(R) \in \bar{Q}(R)$[35].

La reconstrucción original toma la aserción empírica de una teoría como una aserción *exacta*. Naturalmente, esto es una fuerte simplificación. En lo que sigue nos proponemos «emborronar» dicha reconstrucción para obtener una concepción más realista de la estructura de la aserción empírica. A este fin usaremos el concepto de aproximación que hemos elucidado hasta aquí.

La idea básica es que los núcleos teóricos de la red teórica R no se aplican exactamente, sino sólo aproximadamente al dominio global I de R, y que esto se hace usando muchos conjuntos borrosos distintos de la uniformidad empírica U asociada a la teoría. Pero, naturalmente, no todo conjunto borroso servirá a este fin. Para empezar, U contiene conjuntos borrosos que son demasiado grandes (recuérdese el efecto que tiene $D1$-(2)) para que puedan emplearse de manera razonable en la aplicación de la teoría. Si se emborrona demasiado la aplicación, nadie se la tomará en serio. Ciertos conjuntos borrosos se considerarán *admisibles*, otros no.

Qué grado de inexactitud se considerará todavía admisible es una cuestión que depende de la aplicación particular que queremos subsumir bajo la teoría. Cuando aplicamos la ley de la gravitación al movimiento planetario no esperamos el mismo grado de aproximación que en el caso en que la aplicamos a la atracción que se ejercen mutuamente dos grandes esferas de plomo sobre la superficie terrestre y en que hacemos las mediciones por medio de una balanza de torsión muy sensible.

Ahora bien, el grado admisible de inexactitud no sólo depende de la aplicación particular de que se trate. También depende del estadio

[35] Es importante hacer notar aquí que la aserción empírica de una teoría es una entidad histórica que varía con el curso del tiempo histórico.

de la evolución científica. Las expectativas de la comunidad científica con respecto a la aproximación cambian con el tiempo, según diversos factores informales, tales como progreso tecnológico en la detección y medidas, progreso matemático en los cálculos teóricos, éxitos previos logrados por teorías rivales, etc. Estos son factores fuertemente pragmáticos. Es difícil imaginar que nunca lleguen a ser completa y formalmente elucidables. Por otro lado, son muy importantes en el proceso de admisión de un conjunto borroso dado en una aplicación particular. Deberemos tenerlos en cuenta de alguna manera en la reconstrucción de la aserción empírica de la teoría.

Aquí nos enfrentamos a un problema parecido al de determinar formalmente el concepto de una aplicación propuesta, que también es fuertemente pragmático. Este problema ha sido ampliamente discutido en capítulos anteriores. Todo lo que puede exigirse del filósofo de la ciencia en tales casos es que trate de encontrar algunas condiciones necesarias (débiles) para la determinación del concepto en cuestión. Esto es lo que se tratará de hacer aquí para la noción: «clase de conjuntos borrosos admisibles (para la aplicación de la teoría en un período dado)». Llamemos a esta clase «α».

Es evidente que esta clase tiene que ser un subconjunto de la uniformidad empírica sobre M_p–, bien, alternativamente, un subconjunto de la uniformidad inducida sobre M_{pp}. Además, como ya hemos señalado, un grado demasiado alto de inexactitud en cualquier aplicación aparecería como inaceptable a la comunidad científica. Así, pues, parece plausible postular un *máximo* para la inexactitud admisible en aplicaciones particulares. (Diferentes clases de aplicaciones admitirán, por lo general, diferentes máximos.) Quizá también sería plausible introducir un mínimo: una suposición realista sería la de que nunca podemos garantizar que teorías reales usadas por científicos reales se apliquen exactamente a un dominio cualquiera –al menos si no son triviales. Sin embargo, aquí no quisiera excluir formalmente la posibilidad de una aplicación exacta al menos en ciertos tipos particulares de dominios de ciertas teorías particulares, por lo que no postularé la exigencia de mínimo en la definición subsiguiente. (El lector que crea que es realmente una condición necesaria para una aproximación admisible puede fácilmente añadir ese requerimiento a los axiomas de dicha definición.) Otro requisito obvio es que la admisibilidad de un conjunto borroso debería ser invariante con respecto a cambios puramente teóricos en sus modelos; esto es, si u coincide plenamente con u' en sus partes no-teóricas, y u es admisible, entonces u' también lo debería ser. Un último requisito que me parece realmente necesario es que cualquier conjunto borroso, para ser efectivamente admisible, deberá tener algún tipo de relevancia no-trivial para al menos alguna aplicación propuesta de la teoría; esto es, cualquier conjunto borroso admisible debería usarse para proporcionar una aproximación no-trivial de al menos alguna aplica-

ción propuesta. (No-trivial significa aquí que se prescinde del caso de identidad al nivel empírico.) Éste es el significado intuitivo del último axioma de la lista.

Dado que no queremos excluir la posibilidad de que se pudieran postular otras condiciones necesarias para el concepto de conjunto borroso admisible, hablaremos de «clase *potencial* de conjuntos borrosos admisibles».

Sea U una uniformidad empírica sobre M_p de la red teórica arbórea R y sea I el dominio global de aplicaciones propuestas de R.

*D*5: α es una clase potencial de conjuntos borrosos admisibles de U si y sólo si:

(1) $\alpha \subseteq U$.

(2) $\exists u_m(u_m \in \alpha \wedge \forall u(u_m \subset u \rightarrow u \notin \alpha))$.

(3) $\forall u, u'(u \in \alpha \wedge u' \in U \wedge Rs(u) = Rs(u') \rightarrow u' \in \alpha)$.

(4) $\forall u \exists x, x'(u \in \alpha \rightarrow r(x) \in I \wedge r(x) \neq r(x') \wedge \langle x, x' \rangle \in u)$.

Hagamos las siguientes observaciones sobre esta definición. En primer lugar, *D*5-(2) permite la existencia de diferentes «máximos» u_m, en la medida en que no estén conectados por la relación de inclusión. En *D*5-(4), la condición $r(x) \neq r(x')$ es necesaria dado que, de lo contrario, el requerimiento quedaría siempre trivialmente satisfecho por los pares idénticos $\langle x, x \rangle$, todos los cuales están en cualquier u.

De manera enteramente similar se podría definir un concepto de conjunto borroso admisible para una uniformidad sobre M_{pp}. El requisito *D*5-(3) se reduciría entonces a una tautología. Podemos plantear entonces una cuestión parecida a la de la sección anterior: ¿está una clase de conjuntos borrosos admisibles al nivel no-teórico relacionada de alguna manera natural con una clase de conjuntos borrosos admisibles al nivel teórico? La respuesta es también en este caso sí –en la medida en que nos restrinjamos a la lista actual de axiomas[36]. Puede probarse que una clase potencial de conjuntos borrosos admisibles sobre M_p induce su correspondiente clase de conjuntos borrosos admisibles sobre M_{pp}.

Sea α una subclase cualquiera de la uniformidad U sobre M_p. Definimos su restricción no-teórica:

*D*6: $\beta[\alpha] =: \{v/\exists u(u \in \alpha \wedge v = Rs(u))\}$.

*T*8: Si α es una clase potencial de conjuntos borrosos admisibles sobre M_p, entonces $\beta[\alpha]$ es una clase potencial de conjuntos borrosos admisibles sobre M_{pp}.

[36] La adición a *D*5 del postulado de «mínimo» no modificaría esta aseveración, como puede comprobarse fácilmente.

Prueba.

(1) $\beta[\alpha] \subseteq V[U]$, por *D*6, *D*5, *D*3.

(2) $\exists v_m(v_m \in \beta[\alpha] \wedge \forall v(v_m \subset v \rightarrow v \notin \beta[\alpha]))$.

En efecto, por *D*5-(2) sabemos que existe al menos un $u_m \in \alpha$ tal que (A) $\forall u(u_m \subset u \rightarrow u \notin \alpha)$.
Tómese $v_m =: Rs(u_m)$.
Supóngase que hubiera un v_o con $v_m \subset v_o \wedge v_o \in \beta[\alpha]$.
Esto implica que existe un $u_o \in \alpha$ con $v_o = Rs(u_o)$.
Dado que $v_o \neq v_m$, también será $u_o \neq u_m$.
Tómese $u_o \cup u_m$. Por *T*4 tendremos

$$Rs(u_o \cup u_m) = Rs(u_o) \cup Rs(u_m) = v_o \cup v_m = v_o = Rs(u_o).$$

Dado que $Rs(u_o \cup u_m) = Rs(u_o)$ y $u_o \in \alpha$, por *D*5-(3) tendremos también $u_o \cup u_m \in \alpha$; además, $u_m \subset u_o \cup u_m$.
Estos dos últimos resultados contradicen (A). Luego el supuesto acerca de v_o es falso.

*D*5-(3) para $\beta[\alpha]$ es trivial, dado que $Rs(v) = v$.
*D*5-(4) para $\beta[\alpha]$ se sigue inmediatamente de *D*5-(4) para α.

Ahora bien, supóngase que disponemos de una clase α de conjuntos borrosos admisibles para aplicar aproximativamente los núcleos de una red R a un dominio I durante un cierto período de la evolución de la teoría. (α misma es una entidad histórica.) ¿Qué significa esto? Intuitivamente esto significa que la aserción exacta $I \in \bar{Q}(R)$ se transforma en una aserción aproximativa –que simbólicamente representamos por $I \tilde{\in} \bar{Q}(R)$– construida por medio de α. ¿En qué forma podemos precisar adecuadamente esta expresión «$I \tilde{\in} \bar{Q}(R)$»?

A este fin, es conveniente introducir una notación especial que conduce a una especie de «lógica de la aproximación», para emborronar relaciones modelo-teóricas.

«$x \sim x'$» significa: $\exists u(u \in \alpha \wedge \langle x, x' \rangle \in u)$[37].

Para cualesquiera predicados modelo-teóricos P, R, ...:

«$P(\tilde{x})$» significa: $\exists x'(x \sim x' \wedge P(x'))$[38].
«$R(\tilde{x}, y)$» significa: $\exists x'(x \sim x' \wedge R(x', y))$.

[37] Dado que la clase admisible α al nivel de M_p induce una clase admisible $\beta[\alpha]$ al nivel de M_{pp} por *T*8, la relación $x \sim x'$ al nivel de M_p siempre podrá traducirse en la relación correspondiente $r(x) \sim r(x')$ al nivel de M_{pp}.

[38] Nótese que, al escribir $P(\tilde{x})$, no queremos indicar que se emborrona x, sino el *lugar* que ocupa x, es decir, su adscripción a P.

Debería estar claro cómo se pueden definir $R(\tilde{x}, \tilde{y})$, $R(\tilde{x}, y, \tilde{z})$, etc., de manera análoga.

Y para clases modelo-teóricas X, Y, ..., tenemos:

«$X \sim Y$» significa: $\forall x\, \exists y (x \in X \rightarrow y \in Y \wedge x \sim y) \wedge$
$\forall y\, \exists x (y \in Y \rightarrow x \in X \wedge x \sim y)$.

«$\bar{P}(\tilde{X})$» significa: $\exists X' (X \sim X' \wedge \bar{P}(X'))$.

«$\bar{R}(\tilde{X}, Y)$» significa: $\exists X' (X \sim X' \wedge \bar{R}(X', Y))$, etc.

Una tarea que podría resultar interesante por sí misma sería la de desarrollar ulteriormente este instrumento simbólico para obtener una especie de «cálculo de la aproximación». Pero no es la tarea aquí propuesta. Lo que queremos es utilizar ese instrumental para reconstruir con precisión la versión aproximativa de «$I \in \bar{Q}(R)$».

Dado que la aserción empírica de la teoría es una relación diádica basada en $\in$ entre un conjunto y una clase de conjuntos, en principio podríamos emborronarla de tres maneras distintas: o bien emborronar el primer término relacional solamente, o bien sólo emborronar el segundo, o finalmente emborronar ambos simultáneamente. Tenemos, pues, las siguientes posibilidades para expresar la aserción aproximativa de la teoría:

(a) $\tilde{I} \in \bar{Q}(R)$, es decir, $\exists Y (Y \sim I \wedge Y \in \bar{Q}(R))$.

(b) $I \in \widetilde{\bar{Q}(R)}$, es decir, $\exists \bar{I} (\bar{I} \sim \bar{Q}(R) \wedge I \in \bar{I})$.

(c) $\tilde{I} \in \widetilde{\bar{Q}(R)}$, es decir, $\exists Y, \bar{I} (Y \sim I \wedge \bar{I} \sim \bar{Q}(R) \wedge Y \in \bar{I})$.

Vale la pena notar que (b) implica (a), pero no al revés; y (c) no tiene ninguna relación lógica con (a) o (b). Cada uno de estos enunciados tiene un significado formalmente preciso de acuerdo con las definiciones anteriores. El símbolo $\sim$ nos permite abreviar extraordinariamente las proposiciones aproximativas respectivas; de lo contrario, las expresiones (a), (b) y (c) serían muy largas y complicadas.

La cuestión que se plantea ahora es la de cuál de las tres versiones posibles, (a), (b) y (c), es la más adecuada para representar la aserción aproximativa de una teoría empírica. (a) significa que sólo tomamos aproximaciones sobre las aplicaciones propuestas; esto no parece plausible, ya que la aproximación de leyes se hace con independencia de las aplicaciones propuestas (como ya se ha visto en el ejemplo de la ley de la gravitación aplicada al sistema Sol-Tierra). Por razones exactamente complementarias, tampoco (b) parece ser un buen candidato, dado que, aunque tomemos una aproximación teórica de una ley, no estaremos justificados en inferir que esta aproximación se

aplica exactamente a alguna aplicación propuesta (recuérdese el mismo ejemplo mencionado). Así, pues, parece que «$\tilde{I} \in \overline{\tilde{Q}(R)}$» es el mejor candidato para expresar la aserción empírica aproximativa de la teoría: se toman conjuntos borrosos tanto para las sistematizaciones teóricas como para las aplicaciones propuestas. Así es como el científico puede tener alguna esperanza en lograr que «la teoría encaje con los hechos».

92 050 000

2.8. RELACIONES INTERTEORICAS

MONADOLÓGICA =
ENSIMISMADO =

Las teorías científicas no son mónadas leibnizianas. No aparecen ensimismadas y satisfechas con su propia existencia. Esto no sólo es así en el sentido obvio de que las teorías se usan para explicar *algo* supuestamente distinto a ellas mismas, ni tampoco en el sentido de que es *alguien* quien las usa para explicar ese algo; en este sentido habría por lo menos dos cosas externas a una teoría con las que ella estaría relacionada: el usuario y el campo de aplicación de la teoría. Ahora bien, en el concepto «refinado» de teoría empírica propuesto en este libro, en el que se incluyen formalmente no sólo las estructuras conceptuales (el núcleo *K*), sino también el dominio de aplicaciones empíricas propuestas (*A*) y el usuario (la «comunidad científica» *CC*)[39], nos podríamos sentir tentados a regresar a una concepción realmente «monadológica» de las teorías científicas: si el concepto de teoría incluye su campo de aplicaciones y su usuario, ¿qué más queda fuera de la teoría que tenga interés para ella? Una teoría científica podría ser concebida entonces, a un nivel más complejo, efectivamente como un ente ensimismado.

Sin embargo, aun admitiendo el concepto pragmáticamente enriquecido de teoría, la visión monadológica no sería una concepción correcta de la relación de las teorías con su entorno. Prescindiendo ahora del campo de aplicaciones y de la comunidad científica, el hecho es que la estructura conceptual de una teoría se relaciona con las estructuras conceptuales de otras teorías. Quizá haya algún contraejemplo a esta afirmación. Quizá haya algún caso de teoría solitaria, que no se relaciona con ninguna otra. Esto no se puede

[39] Cf. capítulo 2.4.

excluir *a priori*, ciertamente; pero parece poco probable que ocurra. Y si ocurre es un caso excepcional. Todas las teorías empíricas que conozco se relacionan de algún modo no-trivial con otras teorías (claramente distintas a ellas mismas). Las teorías científicas no son *self-made men* dispuestas a enfrentar todos sus problemas por su propia cuenta, sino que para resolverlos, al menos en parte, echan mano de la ayuda fraternal que les conceden sus congéneres. Las teorías forman comunidades de individuos simbióticos que se ayudan mutuamente.

En un lenguaje menos metafórico y más preciso, queremos decir lo siguiente. Dado un elemento teórico T de una red R, existen elementos teóricos T', T'', ..., de otras redes R', R'', ..., que parcialmente contribuyen a determinar el modo como la comunidad $CC(R)$ usa el núcleo K de T para aplicarlo al dominio A de T. Esto ocurre a través de relaciones de diversa índole establecidas entre los conceptos de T y los de T', T'', etc. A tales relaciones las llamaremos genéricamente «relaciones interteóricas».

Existe, pues, un objeto de investigación adicional para cualquier metateoría de las teorías empíricas, aparte de las teorías mismas (incluso si por «teoría» se entiende algo tan complejo y pluriforme como lo que aquí llamamos «redes teóricas»); a saber, las *relaciones* entre las teorías (=entre las redes teóricas). El estudio de las relaciones interteóricas pertenece de manera natural a la metaciencia, y en particular al programa estructural aquí adoptado.

Si queda mucho por hacer en la reconstrucción detallada de teorías concretas, aún más falta en el sentido de la identificación y reconstrucción de relaciones interteóricas concretas. Esto no es sólo debido a que las teorías mismas que son los términos de las relaciones no han sido en muchos casos bien reconstruidas. Es debido también a que para detectar relaciones interteóricas en casos particulares hace falta partir de un esquema general que, aunque sea de manera hipotética y sujeta a constante revisión, nos diga qué tipos de relaciones debemos buscar y podemos esperar. En una palabra, nos falta una *tipología* (putativa) de relaciones interteóricas. En este capítulo vamos a dar los primeros pasos en este sentido. En los capítulos subsiguientes se intentará apoyar en resultados concretos algunas de las consideraciones generales de éste.

Aparte del valor que en sí misma tiene la reconstrucción de relaciones interteóricas dentro de un programa metacientífico general, vale la pena mencionar aquí tres importantes resultados adicionales a los que puede contribuir dicha reconstrucción (y la tipología concomitante):

(1) Puede que obtengamos así un concepto preciso de *disciplina científica*, como unidad a nivel superior que trasciende la idea de una yuxtaposición de teorías particulares.

(2) Puede que nos ayude a desarrollar algo así como una «semántica física» (o «empírica»), es decir, una semántica para teorías físicas (o, en general, empíricas) que supere las deficiencias que presenta la semántica formal usual cuando se la aplica a teorías empíricas.

(3) Puede que permita tratar con mayor precisión algunos enojosos problemas del estudio diacrónico de la ciencia, especialmente los asociados al fenómenos de sustitución de teorías (y a la famosa «tesis de la inconmensurabilidad»).

Aunque en este libro no se va a desarrollar y sustentar en detalle la tesis de que el estudio de las relaciones interteóricas es relevante para los tres puntos mencionados, en capítulos subsiguientes se darán indicaciones en este sentido: sobre su relevancia para el punto (1), en 2.10; para el punto (2), en 2.10 y 4.1; para el punto (3), en 2.9.

La tarea que nos proponemos aquí es, pues, proporcionar un primer esbozo de una tipología de relaciones interteóricas. ¿Por dónde empezar?

Restrinjamos de momento nuestra atención a posibles relaciones interteóricas *diádicas*, es decir, relaciones en las que sólo intervienen dos teorías. Naturalmente, podríamos considerar también la posibilidad de relaciones de orden superior: por ejemplo, sucesiones de teorías o marcos de teorías. Pero, de momento, tenemos ya bastante trabajo con desbrozar el paisaje de las relaciones diádicas entre teorías.

Para iniciar la discusión adoptemos primero el punto de vista de la concepción estándar de las teorías. Parece que ella nos llevaría a una clasificación de las relaciones interteóricas en cuatro alternativas básicas. Aunque no he visto tal clasificación propuesta explícitamente en la literatura, creo que gran parte de las discusiones clásicas sobre las relaciones entre teorías dan por supuesto que las alternativas anunciadas son exhaustivas.

Supongamos que disponemos de dos teorías bien identificadas, T y T'. Dentro de la concepción estándar, tales teorías serán «cálculos interpretados», es decir, sistemas de enunciados interpretados sobre un dominio empírico. Entonces las alternativas implícitamente asumidas son como siguen. O bien T y T' no están interpretadas sobre el mismo dominio empírico (por así decir, no tienen empíricamente nada en común), o bien, si lo están, puede ocurrir que: a) se contradigan, o b) sean equivalentes empíricamente, o c) una sea reducible a la otra. (A veces, en lugar de «reducir» se habla de «explicar» una teoría con la otra; sin embargo, dado que el término «explicación» no parece muy adecuado para expresar una relación entre teorías y dado que, en todo caso, se trata de lo mismo que en otros contextos se entiende por «reducción», usaremos este último término.)

Para fijar ideas, supongamos que T y T' han sido axiomatizadas e

identifiquemos cada una de ellas con su respectivo conjunto de axiomas (interpretados). Usando el símbolo «≡» para equivalencia de sistemas de enunciados y «⊨» para la relación de consecuencia entre conjuntos de enunciados, las cuatro alternativas mencionadas serían formalmente:

(1) $T \cap T' = \emptyset$.
(2) $\exists e \in T, \exists e' \in T' : e \rightarrow \neg e'$.
(3) $T \equiv T'$.
(4) $T \models T'$ (o viceversa).

Es curioso que en la literatura se haya discutido poco una quinta alternativa formalmente posible dentro de este esquema, a saber, suponiendo que no se cumplan (2), (3) y (4).

(5) $T \cap T' \neq \emptyset$, $T \neq T'$,

es decir, podría ocurrir que T y T' coincidieran *en parte* de sus axiomas. Puede que los filósofos hayan considerado que este caso es poco interesante, ya sea porque creen que no hay ejemplos reales que cumplan esta relación, o bien porque consideran que los casos realmente interesantes de relaciones interteóricas son aquellos en que hay una relación «de todo a todo» entre ambas teorías, o sea, en que *todos* los axiomas de ambas teorías están involucrados. Sin embargo, ninguno de ambos supuestos parece haber sido justificado ampliamente.

Tampoco parece haber mucho interés en el caso (3), dado que se le suele considerar una «identidad camuflada» entre teorías o un caso especial de (4), en el que la reducción funciona en ambos sentidos.

De hecho, la discusión alrededor de las relaciones interteóricas, especialmente cuando se ha tratado de la relación *diacrónica* entre dos teorías separadas por un proceso histórico de sustitución (o «revolución científica», o «cambio de paradigma») se ha concentrado en la dicotomía: incompatibilidad *versus* reducción. Y por incompatibilidad se ha entendido (1) o (2), o bien incluso un híbrido monstruoso de (1) y (2). Así, la famosa «tesis de la incommensurabilidad» entre teorías separadas diacrónicamente por una sustitución, adelantada por Kuhn, Feyerabend y otros, se interpreta a veces en el sentido de (1), a veces en el de (2), y a veces en un mismo autor encontramos oscilaciones entre (1) y (2): por un lado se dice, por ejemplo, que la mecánica relativista *contradice* la mecánica newtoniana (o sea, ocurriría (2)), pero en otro pasaje, a veces en el mismo contexto, se dice que ambas teorías son inconmensurables «porque sus conceptos básicos están interpretados diferentemente» y, por tanto, aunque los axiomas tengan la misma forma o formas análogas, «hablan de cosas distin-

incomensurabilidad = no traducibilidad

tas». Bajo tal interpretación está claro que nos hallamos en el caso (1), es decir, no hay relación alguna entre las teorías. Tomada literalmente, la tesis de la inconmensurabilidad afirma precisamente esto: que dos teorías separadas por un «cambio revolucionario» no tienen nada que ver entre sí, como sistemas axiomáticos *interpretados* son disjuntas; por tanto, tampoco pueden contradecirse entre sí. Pero está claro que no es esto lo que los defensores de la tesis de la inconmensurabilidad *quieren decir*. Un ejemplo claro de teorías inconmensurables en el sentido preciso de (1) sería, pongamos por caso, el par consistente en la hidrodinámica y la teoría marxiana del valor. Estas dos teorías realmente no tienen nada que ver entre sí, ni desde el punto de vista de su estructura formal ni desde el de sus interpretaciones empíricas. No es plausible suponer que los partidarios de la tesis de la inconmensurabilidad quieran mantener en serio que la relación existente entre la mecánica newtoniana y la relativista es del mismo tipo que la existente entre la hidrodinámica y la teoría del valor. *Qué* es lo que quieren mantener en serio, sin embargo, no está claro. La oscuridad de su tesis es lo que ha permitido que, al cabo de dos décadas de discusiones sin fin, la literatura en torno a ella haya resultado tan voluminosa como infructuosa. Por otro lado, podemos ser benevolentes con la tesis en cuestión y admitir que hay un núcleo intuitivamente interesante en ella, que vale la pena tratar de reconstruir. La esterilidad de las discusiones planteadas hasta la fecha parece indicar por otro lado que no debemos reconstruir ese núcleo interesante ni en el sentido de (1), ni en el de (2), ni, por supuesto, en el de (4). Quizá esto es síntoma de que hay relaciones interteóricas que no se dejan apresar bien dentro del esquema simplista de la concepción enunciativa.

Otro síntoma es el siguiente. Si entendemos la reducción de teorías en el sentido de (4), entonces para mostrar que una teoría es reducible a otra bastaría *deducir* los axiomas de la primera de los de la segunda (quizá con algunas premisas adicionales). Los libros de texto de física y los de filosofía de la ciencia apoyan esta visión de la reducción de teorías: cuando dan ejemplos de reducciones, dicen que los axiomas de una teoría son deducibles de los de la otra. Por ejemplo, dicen que las leyes de Kepler son deducibles de las de Newton, que la mecánica del sólido rígido es deducible de la mecánica de partículas o que los principios de la termodinámica son deducibles de la mecánica estadística. A veces, si son más cuidadosos, observan que la deducción (=reducción) sólo es posible si se admiten ciertas hipótesis adicionales. Ahora bien, cuando examinamos a fondo estos supuestos ejemplos de reducción, nos damos cuenta de que la deducción apetecida no funciona, ni siquiera con hipótesis adicionales. Las razones son varias, pero la más fundamental radica en la diversidad conceptual de las teorías en cuestión, diversidad que no se elimina ni siquiera con hipótesis *ad hoc*. Está claro que si los conceptos básicos de ambas

teorías son distintos o tienen interpretaciones distintas (como es el caso en los ejemplos mencionados), entonces la relación de consecuencia lógica $\models$ no puede aplicarse sin más.

A partir de este hecho los adalides de la tesis de la inconmensurabilidad sacaron la conclusión de que el concepto de reducción no sirve para analizar los casos históricamente interesantes de sustitución de teorías, *porque* en esos casos siempre se da algún tipo de cambio semántico. No obstante, ésta es una conclusión precipitada. Se puede intentar introducir un concepto más elaborado de reducción, que no se limite a la relación $\models$, y que sea intuitivamente adecuado. En tal caso, hay por lo menos alguna posibilidad *a priori* de que obtengamos un concepto de reducción libre de las dificultades lógicas asociadas al fenómeno de los cambios semánticos. Naturalmente, para ver si hay alguna posibilidad razonable en esta dirección antes hay que aclarar nuestra intuición de lo que debe ser una reducción de teorías, y para ello, la primera pregunta es: ¿queremos un concepto de reducción tal que necesariamente implique la invariabilidad semántica de los conceptos básicos de las teorías involucradas? *Mi* intuición y la de muchos de mis colegas es que esto no es una condición necesaria de la reducción. Claro que otros pueden decir que su idea intuitiva de reducción es tal que implica invariabilidad semántica. Pero aun cuando ésa sea su intuición (y todo el mundo es libre de tener las intuiciones que le vengan en gana), lo que no pueden afirmar sin más, so pena de caer en el dogmatismo, es que un concepto de reducción que no implique invariabilidad semántica está necesariamente desprovisto de interés metodológico. No pueden inferir de su intuición que un concepto tal no será formalmente identificable y materialmente fructífero como elucidación de relaciones interteóricas.

"Creo que hay dos ideas intuitivas básicas en las que se piensa cuando se habla de reducción de teorías y que son independientes de la variación o no de significados." La primera es que si T es reducible a T', entonces debería ser el caso que siempre que T sea realmente aplicable, esto pueda justificarse apelando a la aplicabilidad de T', es decir, el uso de la teoría reducida debe ser justificable a partir del uso de la reductora. La segunda es que la teoría reductora debería permitir distinciones más finas que la reducida (a veces esto se expresa diciendo que «T' debe tener más poder explicativo que T»).

La tesis que sostengo ahora es la siguiente. Si, por un lado, se formulan nuestras intuiciones acerca de la reducción del modo que acabamos de ver (o un modo análogo) y, por otro, se abandona una concepción meramente enunciativa de las teorías en favor de una concepción más estructural, entonces resulta claro que no es preciso ni conveniente identificar la relación de reducción con la de deducción, ni tenemos por qué embrollarnos con el problema de los cambios semánticos. Basta con establecer una comparación entre las estructu-

ras correspondientes de ambas teorías que tenga las propiedades intuitivamente requeridas. –

El primer paso en esta dirección, incluso antes de que se empezara a hablar tanto de la «inconmensurabilidad», lo dio E. Adams. En su tratamiento de la reducción de la mecánica del sólido rígido a la de partículas introdujo un concepto modelo-teórico de reducción, el cual, ligeramente simplificado y reformulado, reza así:

«T es reducible a T' si y sólo si existe una relación diádica ρ entre cada estructura x expresada en el lenguaje de T y estructuras x_i' (normalmente más de una) expresadas en el lenguaje de T' tal que:

[R] $(x \rho x_i' \wedge x_i' \in M[T']) \models x \in M[T]$.»

Si analizamos cuidadosamente las condiciones de esta definición veremos que se pueden tomar como una elucidación de los dos requisitos básicos de nuestra intuición. En efecto, que una teoría sea aplicable a un dominio significa formalmente que ese dominio se puede conceptualizar como un modelo de la teoría. Entonces la fórmula [R] dice que, siempre que T es aplicable, esto se puede inferir lógicamente de que T' es aplicable a un dominio ρ-correspondiente (que por tanto será un modelo de T'). En cuanto a la idea de que T' es «más fina» que T, se la puede interpretar modeloteóricamente como diciendo que, siempre que T es aplicable a un dominio, de modo que este último se conceptualiza como un modelo x de T, T' también será aplicable a ese dominio, pero dando lugar normalmente a *varios* modelos x_i' para el mismo; o sea, T' tiene mayores posibilidades de diferenciación conceptual.

En esta versión la noción de reducción aún está conectada, de manera parcial e indirecta, con la relación de consecuencia lógica (aunque es de notar que ahora usamos una noción de consecuencia lógica de orden lógico superior: como relación entre enunciados modelo-teóricos). De todos modos, ambas relaciones no son idénticas, y lo que obtenemos es una relación de reducción más compleja que la inicial: es una correspondencia de cierto tipo entre dos conjuntos de estructuras distintas, que no requiere identificación semántica alguna de los conceptos correspondientes. El modelo x de T y los modelos ρ-correspondientes x_i' de T' pueden estar constituidos por conceptos muy distintos. Lo que se compara ahora no son conceptos aislados, ni siquiera enunciados; son estructuras enteras que representan modelos de las teorías respectivas. El establecer la relación ρ entre los modelos de ambas teorías es indiferente respecto a todos los cambios de significado *dentro* de los modelos que a uno se le puedan ocurrir.

Por medio del aparato metateórico introducido en capítulos anteriores podríamos refinar y desarrollar ulteriormente nuestra

noción de reducción estructural. Por ejemplo, una primera adición consistiría en tener en cuenta el dominio I de aplicaciones propuestas. La condición [R] de la definición anterior sería en tal caso más plausible si nos limitáramos a modelos que corresponden a aplicaciones propuestas:

$$[\mathrm{R}^*] \quad (x \rho x_i' \wedge r(x) \in I(R) \wedge r(x_i') \in I(R') \wedge x_i' \in M[T']) \models x \in M[T].$$

También podríamos tener en cuenta las condiciones de ligadura de las teorías comparadas y la correspondencia entre modelos que cumplen leyes especiales en una y otra teoría, de modo que al final pudiéramos construir una relación ρ entre redes teóricas enteras. Aquí sólo aludimos a estas complejidades adicionales, que han sido tratadas con detalle por Sneed en su [1971], por Stegmüller en su [1973], por D. Mayr en su [1976] y por Balzer y Sneed en su [1977/78]. Aunque todavía hay algunas cuestiones técnicas por resolver en torno al concepto estructural de reducción, estos trabajos muestran que la idea general es plausible y fructífera, y por otro lado es independiente del problema de los cambios semánticos de los conceptos. En principio, pues, existen buenas perspectivas para trabajar con ese concepto de reducción. La siguiente cuestión que se plantea entonces es a qué casos concretos trataremos de aplicarlo.

Adams en su ensayo original, y posteriormente Sneed en su libro, donde presenta una versión más compleja de la reducción, mostraron que este concepto es efectivamente aplicable al caso concreto de la relación entre la mecánica del sólido rígido y la de partículas. Es aún más fácil mostrar que también se aplica al caso más sencillo de la relación entre la mecánica cartesiana del choque y la mecánica newtoniana. No obstante, es dudoso que muchos de los casos interesantes de supuestas reducciones se puedan tratar tan elegantemente como los ejemplos que acabamos de mencionar. La cuestión aparece como especialmente problemática en los «grandes casos» históricos de «superación» de teorías, como el paso de la cinemática planetaria de Kepler a la teoría newtoniana de la gravitación, o de la mecánica clásica a la relativista o cuántica, o de la termodinámica a la mecánica estadística. Dudo que estos casos puedan reducirse, valga la redundancia, a reducciones del tipo Adams-Sneed, aunque probablemente un análisis cuidadoso revele semejanzas de familia con este tipo de relación. La principal razón de disparidad parece ser el hecho de que no se pueden reconstruir por medio de relaciones exactas entre las teorías involucradas, sino sólo mediante algún tipo de aproximación. En particular, está claro que ejemplos como los mencionados no pueden satisfacer la condición [R] (o, en su caso, [R*]), a menos que se construya una ρ inaceptablemente *ad hoc*. Podremos, quizá, reconstruir esos casos mediante una relación ρ natural, *siempre y cuando* se dé una versión aproximativa de [R]. Con otras palabras,

hay que «emborronar» la relación reductiva: hacer de una reducción exacta una inexacta. Esta es la mínima (y ciertamente no-trivial) modificación que tendremos que emprender. Y para ello necesitaremos del aparato aproximativo, especialmente las uniformidades, que hemos introducido en el capítulo anterior. Siguiendo esta estrategia, en el próximo capítulo se introducirá la noción de aproximación interteórica (o «reducción aproximativa») y se aplicará a uno de los casos concretos mencionados.

Resumamos en este punto los tipos de relaciones interteóricas auténticas que hemos detectado hasta aquí. Si admitimos que la equivalencia de teorías es sólo un caso especial de la reducción, y si estamos dispuestos a trabajar con un concepto estructural de reducción que evita los problemas de los cambios semánticos y la inconmensurabilidad, entonces podemos decir que hemos identificado de momento dos clases de relaciones interteóricas, que además, por lo que parece, tienen aplicación a casos concretos de pares de teorías: una noción reelaborada, compleja de reducción exacta y una noción de aproximación interteórica.

¿Es esto todo lo que se puede decir sobre los tipos de relaciones interteóricas? ¿Quedan cubiertos todos los casos por las nociones de reducción y aproximación? La respuesta parece ser no. Para comprender por qué parece que este cuadro de la situación es aún incompleto es conveniente reconsiderar por un momento el modo de funcionamiento de la reducción (exacta).

Cuando se reduce una teoría T a otra T', normalmente no se obtiene efectivamente la reducción poniendo en relación directa los dos núcleos básicos de las teorías. En general, la relación reductiva no opera de manera inmediata del núcleo reductor al reducido, sino que es preciso añadir ciertas condiciones restrictivas a la teoría reductora a fin de poder «traducir» las estructuras de la teoría reductora a las de la reducida. Podríamos decir que una parte de las condiciones (implícitas o explícitas en las exposiciones usuales) que constituyen la relación ρ son «reglas de traducción». Por lo general, tienen un alcance mucho más restringido que los axiomas básicos y a veces incluso aparecen como muy *ad hoc* dentro de la teoría reductora. Aunque estén formuladas en términos de ésta, presuponen en cierto sentido las condiciones bajo las cuales se aplica la teoría reducida. Por ejemplo, cuando reducimos la mecánica cartesiana del choque a la mecánica newtoniana, restringimos la última al caso en que sólo consideramos sistemas de dos partículas móviéndose en la misma dirección y sentidos opuestos con velocidad uniforme –lo cual es sin duda una condición mucho más restrictiva que la ecuación fundamental de Newton. O bien, cuando reducimos la mecánica del sólido rígido a la de partículas, añadimos a la segunda teoría la condición de que las distancias entre las partículas permanezcan fijas en el intervalo de tiempo considerado. La necesidad de condiciones restrictivas de

este tipo para facilitar la reducción parece clara y natural. Pero cuando analizamos más a fondo su carácter lógico-metodológico, la cuestión aparece intrincada. La pregunta básica es: ¿Vamos a decir que tales condiciones pertenecen conceptualmente a la teoría reductora o a la reducida, o a ambas o a ninguna de las dos? Aunque usualmente se formulan en términos de la teoría reductora, sólo adquieren verdadero sentido si se presuponen las condiciones de aplicabilidad de la teoría reducida, es decir, en último término presuponen el concepto mismo de la teoría reducida.

Parece claro, pues, que esas condiciones de reducción poseen conceptualmente un carácter híbrido. Su función es la de ofrecer un puente conceptual entre la teoría reductora y la reducida. Propiamente, no pertenecen ni a la una ni a la otra, aunque contribuyen en alguna medida a establecer la identidad de ambas (por lo menos, la de la teoría reducida).

Lo que acabamos de decir para el caso de la reducción exacta también vale, naturalmente, para la relación de aproximación interteórica. Como veremos en el ejemplo del próximo capítulo, también en ese caso es necesario introducir condiciones especiales en el marco de la teoría aproximadora para obtener la teoría aproximada.

Las consideraciones anteriores nos llevan de manera natural a generalizar nuestra discusión de las relaciones interteóricas. Una vez comprobada la existencia de vínculos de *status* particular en el caso de reducción (exacta o aproximada) de teorías, es plausible suponer que también los podemos encontrar como puentes conceptuales entre teorías que no se hallan en una relación reductiva en ningún sentido intuitivo, sino sólo conectadas por tener algunas aplicaciones comunes o en correspondencia. A tales vínculos en general los llamaremos «*principios-puente*». Son probablemente más frecuentes e importantes de lo que se desprendería del poco interés que han mostrado por ellos hasta ahora los analistas de la ciencia. Puede que parte de la razón por la que han sido poco notados es que no suelen aparecer de manera bien explícita en las exposiciones estándar de las teorías. Pero sin duda juegan un gran papel, incluso en casos en que no se puede hablar de reducción.

Por ejemplo, la termodinámica ciertamente no es reducible a ninguna teoría de la mecánica clásica no-estadística. En particular, no es reducible a la hidrodinámica. Sin embargo, ambas teorías están conectadas a través de varias relaciones. Una de ellas, y muy importante, es la que ya hemos indicado en el capítulo 2.6 como relación básica entre presión, energía interna y volumen:

$$P(z) = -D_V f^U(z).$$

Esta ecuación no es una definición nominal de la presión, como hemos argüido en el capítulo 2.6, ni tampoco tiene el *status* de una ley

fundamental de la hidrodinámica o de la termodinámica. Sin embargo, es esencial a ambas teorías, pues al conectar una magnitud típicamente hidrodinámica, como es la presión, con una termodinámica, como es la energía interna, proporciona la posibilidad de transmitir valiosa información de una teoría a la otra (o, mejor dicho, de la aplicación de una teoría a la aplicación de la otra). En una palabra, se trata de un importante principio-puente entre hidrodinámica y termodinámica debido a la existencia de aplicaciones comunes para ambas teorías. Seguramente podríamos encontrar otros muchos ejemplos parecidos de principios-puente entre teorías no mutuamente reducibles. (Se puede considerar a la ley de Lorentz para cargas en movimiento, que conecta la electrodinámica con la mecánica clásica de partículas, como otro principio-puente de ese estilo.)

Admitamos, pues, que aparecen principios-puente vinculando teorías incluso en casos en los que no se trata de establecer ni una reducción ni una aproximación. La cuestión que se plantea ahora es la siguiente. Si no es para establecer relaciones de reducción o aproximación, ¿existe(n) alguna(s) otra(s) categoría(s) general(es) de relaciones interteóricas que se pueda(n) identificar formalmente y que requiera(n) de principios-puente como los mencionados?

En el presente carezco de una respuesta definitiva a esta pregunta. No obstante, una hipótesis que quisiera proponer al respecto es la de que *algunos* principios-puente, aunque probablemente no todos, se introducen para expresar el hecho de que una teoría T dada *presupone* para su determinación cuantitativa algunos de los conceptos de una teoría previa T', generalmente de estructura más simple. Con otras palabras, los parámetros de la teoría más simple T' se usan para medir algunos parámetros de T, al menos en ciertos modelos; y el modo de hacerlo se expresa en el principio-puente correspondiente. Por ejemplo, la determinación cuantitativa de la termodinámica en algunas de sus aplicaciones presupone ciertos parámetros hidrodinámicos, como P, o geométricos, como V. Y la forma exacta de esta dependencia se expresa mediante la ecuación arriba escrita.

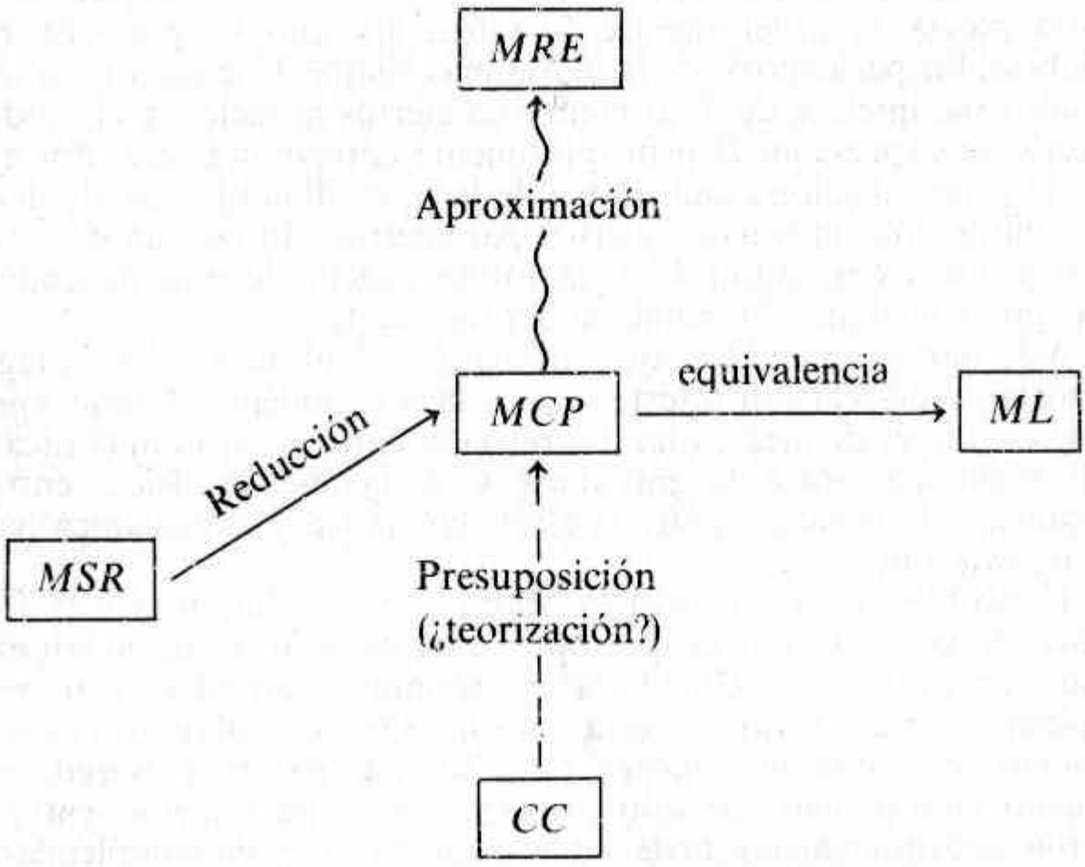
Así, parece plausible que podamos identificar otra categoría importante de relación interteórica, a la que podemos llamar «*presuposición* a fines de medición». La relación entre la termodinámica y la hidrodinámica, entre la mecánica y la geometría física, entre la mecánica y la óptica, y entre la electrodinámica y la mecánica parece ser de este tipo.

El problema que se plantea ahora es el de determinar la forma lógica de esta relación de presuposición, es decir, el de averiguar el modo apropiado de elucidarla en términos formales. Este punto también es una cuestión abierta para mí. Quizá en algunos casos esta relación se pueda identificar como lo que Balzer y Sneed, en el artículo citado, han llamado «*teorización*»: una relación entre dos teorías jerárquicamente ordenadas según su nivel de complejidad, en

la que el aparato conceptual de la teoría más compleja resulta de la extensión modelo-teórica de la menos compleja; es decir, las estructuras de la teoría más compleja se construyen añadiendo parámetros escogidos apropiadamente a las estructuras de la teoría menos compleja. Así, por ejemplo, la mecánica newtoniana sería una teorización de la cinemática galileana en el sentido de que las estructuras de la cinemática galileana, que sólo incluyen parámetros espaciotemporales y derivados, se extienden, con propósitos mecánicos, a estructuras que, adicionalmente, incluyen funciones dinámicas como masa y fuerza.

Sin embargo, es dudoso que los estrictos requerimientos modelo-teóricos que Balzer y Sneed imponen a la relación de teorización realmente se cumplan en todos los casos de lo que llamo «presuposición». En el caso ya discutido de la relación entre la hidrodinámica y la termodinámica no parece ser así. Aunque no dispongamos en este momento de una reconstrucción estructural de la hidrodinámica, es inverosímil suponer que la termodinámica vaya a aparecer como una teorización de la hidrodinámica. Por lo tanto, de momento parece más prudente separar los conceptos de presuposición y teorización. Es posible que su aplicación coincida en algunos casos, pero no tienen por qué identificarse formalmente.

Debería estar claro en este punto que una teoría dada puede hallarse en toda una serie de relaciones interteóricas con toda una serie de teorías diferentes, cada una de las cuales, a su vez, puede



hallarse en diversas relaciones con otras teorías, y así sucesivamente, de tal modo que, si representáramos gráficamente todas estas interconexiones, obtendríamos una enorme y complicada red de teorías (cada una de las cuales, a su vez, sería una red teórica de elementos en el sentido del capítulo 2.4). Así es como habría que imaginar gráficamente una disciplina científica.

A modo de conclusión, podemos dar una ligera idea de cómo aparecería esa red describiendo gráfica y simplificadamente una porción de la misma. Tomemos como punto de partida la mecánica clásica de partículas (*MCP*). Esta teoría se relaciona con al menos las siguientes teorías de las maneras siguientes: es equivalente a la mecánica de Lagrange (*ML*); reduce la mecánica del sólido rígido (*MSR*); presupone la cinemática clásica (*CC*), y es una aproximación de la mecánica relativista especial (*MRE*).

2.9. APROXIMACION INTERTEORICA: EL CASO KEPLER-NEWTON

0. Introducción

En el capítulo anterior se ha esbozado un programa general de estudio de los distintos tipos de relaciones entre teorías, que no se identifican con la mera deducción lógica de los axiomas de una teoría a partir de los de la otra. Uno de los tipos detectados ha sido el de la aproximación interteórica (o reducción aproximativa), como versión aproximativa de la reducción exacta de Adams-Sneed. También se ha señalado que este tipo de relación interteórica puede arrojar alguna luz sobre el fenómeno diacrónico de la sustitución o «superación» de teorías a través de un cambio semántico. En efecto, en muchos de los casos paradigmáticos de pares de teorías separadas por una «revolución científica», a los que se suele asociar el supuesto fenómeno de la inconmensurabilidad, parece claro que hay algún tipo de relación identificable formalmente entre ambas teorías y que esta relación es de tipo aproximativo. Los pares Kepler-Newton, óptica geométrica-óptica ondulatoria, Newton-Einstein, Newton-mecánica cuántica, termodinámica-mecánica estadística, son ejemplos tipicos al respecto. En todos estos casos los físicos suelen decir que la primera teoría del par es una *aproximación* de la segunda; y a pesar de todos los cambios de paradigma e inconmensurabilidades, no veo ninguna razón para dudar de que los físicos *se refieren a algo* realmente cuando hablan de este modo, de que podemos tomarnos en serio su modo de hablar y, finalmente, de que hay alguna esperanza de que podamos proporcionar una reconstrucción formal, filosófica, de eso a lo que se refieren. Es decir, aquí nos encontramos con un *explicandum* interesante;

además, este *explicandum* (teniendo en cuenta los ejemplos mencionados) parece relevante para la discusión de cosas tales como las revoluciones científicas. Nuestra tarea entonces es la de tratar de proporcionar el *explicans* adecuado. Al *explicandum* en cuestión lo denominamos «aproximación interteórica».

La posibilidad de que casos ejemplares de pares inconmensurables de teorías, a pesar de su inconmensurabilidad (o quizá justamente debido a ella), pueden estar en una relación de aproximación (estructural), ha sido señalada independientemente por Stegmüller y Scheibe[40]. Stegmüller habla de «reducción aproximativa» y Scheibe de «explicación aproximativa» de una teoría por la otra. Por mi parte, prefiero hablar de «aproximación interteórica», porque quisiera mantener la elucidación de esta relación tan libre como sea posible de los problemas epistemológicos que afectan los conceptos de reducción y explicación. Naturalmente, éste es sólo un punto terminológico carente de importancia. Concuerdo con la hipótesis metateórica de Stegmüller y Scheibe de que por lo menos muchos casos interesantes de inconmensurabilidad incluirán el tipo de relación que llamo «aproximación interteórica».

Ahora bien, con el fin de evitar confusiones ya desde el principio, es conveniente hacer aquí la siguiente observación sobre la naturaleza de este concepto. Estamos hablando de la «aproximación interteórica» como si estuviéramos pensando en *una sola* clase de relación, esto es, como si solamente fuera posible una forma de elucidación de la aproximación interteórica. Parece que ésta es la premisa más o menos implícita tanto en Stegmüller como en Scheibe. En este punto, sin embargo, mi concepción se aparta de la suya. La «aproximación interteórica» debería entenderse sólo como un «término-racimo» para todo un montón de relaciones interteóricas, todas las cuales tienen «semejanzas de familia» entre sí, pero pertenecen en realidad a tipos distintos. Sus semejanzas de familia provienen del hecho de que todas hacen uso de ciertas estructuras topológicas que tienen algo que ver con nuestra noción intuitiva de aproximación. Pero esto es todo. Como veremos más adelante, no parece posible abstraer de todas estas diferentes clases de relaciones un solo concepto general que sea formalizable con precisión y que las cubra todas. Todo lo que puedo ofrecer, al menos por el momento, es una especie de «algoritmo» para producir todas las clases de aproximaciones interteóricas formalmente posibles. En cierto sentido, al disponer de un algoritmo tal, ya disponemos del concepto general. Pero no he podido encontrar una definición formal general del mismo[41].

---

[40] Stegmüller, en su reciente *Structuralist View of Theories*, (trad. en Alianza) y Scheibe, en su manuscrito, «Eine Fallstudie zur Grenzfallbeziehung in der Quantenmechanik», aún no publicado en el momento de escribir estas líneas.

[41] Estas observaciones aparecerán más claras al final del capítulo.

El primer punto que debemos tener en cuenta en nuestro intento de elucidar la noción o, mejor dicho, las nociones de aproximación interteórica es que se tratará de relaciones «externas» a las teorías comparadas, que no deben ser confundidas con la aproximación «interna» que hemos estudiado en el capítulo 2.7. Por aproximación interteórica entendemos una relación que afecta a pares de redes teóricas enteras, por tanto, a marcos conceptuales distintos. En particular, esta relación no debe confundirse con la aproximación de leyes dentro de *una misma* teoría o, dicho en los términos más precisos de nuestro aparato estructural, con la aproximación entre elementos teóricos de una misma red, todos los cuales coinciden por lo menos en los conjuntos básicos M_p, M_{pp} y M. En los escasos tratamientos sistemáticos que existen hasta la fecha de las nociones aproximativas en teorías empíricas se suele confundir la aproximación «interna» con la «externa»; se trata la aproximación entre leyes dentro de una misma teoría del mismo modo que la aproximación entre teorías distintas (con conceptos básicos distintos). Un buen ejemplo de esta confusión es el tratamiento que se dispensa a la aproximación (o «idealización») en el libro de Krajewski sobre el tema (1978). En contra del supuesto implícito y tomado acríticamente por muchos autores de que no hay diferencias esenciales entre ambos casos, postulamos aquí que la relación que hay, pongamos por caso, entre la ley de los gases ideales y la ley de Van der Waals (ambas constituyen elementos teóricos de la termodinámica del equilibrio) no puede ser la pauta para reconstruir adecuadamente la relación que supuestamente hay entre la termodinámica del equilibrio y la mecánica estadística. O bien, para poner aún otro ejemplo, no podemos asumir que la relación aproximativa que seguramente existe entre la mecánica clásica y la relativista especial será estructuralmente semejante a la que sin duda existe entre la ley de caída libre de Galileo y la ley de gravitación dentro de la mecánica clásica. Naturalmente, *puede* que en ambos tipos de aproximación se puedan detectar ciertos aspectos comunes y quizá *algunas* ideas generales se puedan aplicar a ambos casos. Pero, en rigor, lo razonable es suponer de entrada que las diferencias en naturaleza serán muy significativas. La incapacidad para ver este punto no ha hecho sino reforzar la desconfianza que sienten los defensores de la inconmensurabilidad hacia cualquier intento de reconstruir la relación entre dos teorías conceptualmente distintas por medio de la noción de aproximación. De hecho, si nos constreñimos a usar la aproximación de leyes especiales dentro de una misma teoría como modelo para el tratamiento de la aproximación de teorías, es fácil que en seguida nos encontremos con el mismo tipo de dificultades conceptuales y semánticas que tanto alegran a los partidarios de la inconmensurabilidad.

La aproximación interteórica resulta una noción interesante para comparar teorías con marcos conceptuales irreduciblemente diferen-

tes. Prescindiendo de otros criterios diferenciales para distinguir tipos de aproximación interteórica de los que luego hablaremos, podemos ya establecer una división entre dos grandes casos: un caso es aquel en que los marcos conceptuales de las dos teorías relacionadas son totalmente distintos (los conjuntos correspondientes de conceptos básicos son disjuntos); otro caso es el de una diferencia parcial solamente (los conjuntos correspondientes se traslapan). Si admitimos la distinción a dos niveles entre conceptos *T*-teóricos y *T*-no-teóricos, entonces es probable que el segundo tipo de aproximación interteórica surja cuando disponemos de dos teorías con conceptos *T*-no-teóricos idénticos (o mutuamente interdefinibles), pero distintos conceptos *T*-teóricos. El ejemplo que reconstruiremos en este capítulo pertenece a esta categoría. Esta primera distinción entre dos tipos de aproximación interteórica está relacionada de manera natural con la distinción entre «inconmensurabilidad teórica» e «inconmensurabilidad empírica» que hace Stegmüller en su *Structuralist View of Theories*, § 11. (La primera es «sólo» una inconmensurabilidad parcial, la segunda es total.) Como señala Stegmüller, el segundo tipo de inconmensurabilidad es «particularmente alarmante» para el filósofo clásico de la ciencia y, en este sentido, también es particularmente interesante, caso de que se dé realmente. De ahí que, en la medida en que la aproximación interteórica pueda proporcionarnos alguna clave para comprender la naturaleza y estructura de las inconmensurabilidades, el caso de la aproximación interteórica entre teorías que difieren no sólo en sus estructuras *T*-teóricas, sino también en las *T*-no-teóricas sería más interesante que una aproximación entre teorías que comparten sus respectivas estructuras *T*-no-teóricas. El tratamiento general de la aproximación que se ofrece en este capítulo pretende ser adecuado a ambos tipos. Por otro lado, el ejemplo concreto de aplicación de la noción general que se ha escogido pertenece a la categoría «menos interesante» (= una base *T*-no-teórica común). Puede que esta elección defraude a algún lector, pero dado que el caso menos interesante es más fácil de reconstruir que el más interesante, se ha seguido la estrategia metodológicamente prudente de empezar por un caso más simple, antes de pasar a los más complicados. (Y, después de todo, veremos que el caso «menos interesante» no está tan desprovisto de interés como pueda parecer para contrastar nuestras ideas generales sobre la aproximación interteórica.)

En lo que sigue nos proponemos lo siguiente: desarrollar una noción general, esquemática de aproximación interteórica, que sea aplicable a diferentes casos de relaciones interteóricas aproximativas. La noción general deberá tener un grado suficiente de flexibilidad para que cubra distintas clases específicas de aproximación entre teorías. Esto significa que los requisitos generales formulados para la noción de aproximación sin más no deberán ser demasiado fuertes.

Una vez efectuada esta elucidación general, nos dispondremos a ejemplificar su funcionamiento aplicándola a un caso particular extraído de la historia de la ciencia. El ejemplo escogido es el de la relación entre las teorías respectivas de Kepler y Newton sobre los movimientos planetarios. Este caso particular fue reconstruido ya por Scheibe en 1973, con entera independencia de cualquier metateoría general de las teorías empíricas. La reconstrucción proporcionada por Scheibe es la más detallada que conozco de un caso particular de aproximación interteórica. Por otro lado, no es ni demasiado complicada ni demasiado esotérica. He ahí por qué aparece como muy adecuada para poner a prueba cualquier tratamiento general de la aproximación interteórica.

La estrategia que vamos a seguir para alcanzar nuestro objetivo constará de los siguientes pasos:

1. Introducir la noción de conjunto borroso, que necesitamos para dar cuenta de una versión no-enunciativa de la aproximación, para el caso de la comparación de estructuras conceptuales diferentes.
2. Proporcionar un esquema general adecuado a las relaciones de aproximación interteórica y de aproximación interteórica *estricta*, en ambos casos parcialmente basado en la noción (no-aproximativa) de reducción. (Presuponemos que la aproximación revestirá algún tipo de analogía con la reducción.)
3. Reformular dentro del marco de la metateoría estructural el análisis que hace Scheibe del caso Kepler-Newton. Esto implica, entre otras cosas, que primero hay que identificar claramente las condiciones de aproximación determinadas por Scheibe.
4. Demostrar que las condiciones postuladas por Scheibe para el caso Kepler-Newton se infieren del concepto general de aproximación interteórica estricta al insertar ciertos parámetros concretos en la definición general.

Los pasos 1 y 2 se pretende que sean de validez universal para cualquier estudio de aproximaciones interteóricas. Los pasos 3 y 4 deberían revelar implícitamente los cánones que hay que seguir en el análisis de otros casos particulares de aproximaciones interteóricas.

Antes de empezar propiamente la tarea, introduzcamos ciertas convenciones notacionales.

En lo que sigue, la variable T se referirá siempre a un elemento teórico de alguna red, en el sentido definido en 2.4. *D*1, p. 114. Para simplificar la exposición, y dado que en este contexto no se hacen consideraciones diacrónicas ni aproximativas *internas* a los elementos teóricos, prescindiremos de las nociones de comunidad científica, intervalo histórico y uniformidad interna, que constituyen los elementos teóricos, y admitiremos que cada elemento teórico T_i es un par

$\langle K_i, A_i \rangle$, donde K_i es un núcleo y A_i el dominio de aplicación correspondiente, al que la comunidad científica trata de aplicar el núcleo.

Dado un elemento teórico T, x_T, y_T, etc., serán variables para modelos (actuales, potenciales o parciales) de T. Si x es un modelo (actual, potencial o parcial), entonces t_x será un concepto (conjunto, relación, función) contenido en la estructura que define x. A veces, cuando no haya peligro de confusión, se omitirán los subíndices.

1. Conjuntos borrosos para la comparación interteórica

En el capítulo 2.7 hemos visto cómo se puede trabajar con uniformidades modelo-teóricas para tratar los fenómenos de aproximación dentro de una misma teoría. Allí se definió la uniformidad básica sobre $M_p[T]$. Los conjuntos borrosos destinados a expresar aproximaciones se introdujeron como elementos de un subconjunto U de $\mathcal{P}(M_p \times M_p)$ (cf. 2.7, $D1$, p. 174). Ahora podemos usar la misma idea *mutatis mutandis* para la aproximación interteórica. Aquí se dará, sin embargo, una importante diferencia de contenido. Debemos tener en cuenta el hecho de que ahora debemos ser capaces de comparar no sólo el mismo tipo de estructuras, sino también *tipos distintos* de modelos potenciales: estructuras pertenecientes a $M_p[T]$ con estructuras pertenecientes a $M_p[T']$.

Ahora bien, formalmente, esta ulterior complicación no parece implicar ningún cambio radical en la idea de usar uniformidades para comparar estructuras enteras. Todo lo que se requiere es aceptar el principio metodológico de que una estructura conceptual puede «acercarse» a otra estructura conceptual de distinto tipo del mismo modo como puede «acercarse» a una estructura conceptual de su mismo tipo. Y dentro del marco formal de uniformidades que introducimos, podemos decir exactamente qué significa esto. Ciertamente, una comparación de esta clase no parece muy natural dentro de la concepción enunciativa o desde el punto de vista de un análisis micrológico. Pero aquí le podríamos dar la vuelta al argumento. En efecto, esta situación muestra una vez más cuán restrictivo es el análisis micrológico. Al tomar el punto de partida macrológico, estructural, podemos «jugar» con las estructuras de las teorías de muchas más maneras de lo que podría imaginar el enfoque estándar. Y no veo ningún motivo intuitivo o formal para excluir tales posibilidades.

Si el argumento precedente se acepta, entonces sólo tenemos que hacer una pequeña modificación formal en nuestra noción de uniformidad para poder tratar la comparación de clases diferentes de modelos potenciales. Ahora habrá que definir la uniformidad como un subconjunto de $\mathcal{P}((M_p[T] \cup M_p[T']) \times (M_p[T] \cup M_p[T']))$, en vez de $\mathcal{P}(M_p \times M_p)$. Todos los demás axiomas de $D1$ de 2.7 se pueden

recoger en la misma forma. Incluso al precio de cierta redundancia, parece conveniente volver a formular esos axiomas aquí, teniendo en cuenta las ligeras modificaciones mencionadas.

Para facilitar la lectura, en lo sucesivo escribiremos M_p y M'_p en vez de $M_p[T]$ y de $M_p[T']$. Lo mismo vale para M y M', etc.

*D*1: U es una uniformidad sobre $M_p \cup M'_p$ si y sólo si:

(1) $\varnothing \neq U \subseteq \mathscr{P}((M_p \cup M'_p) \times (M_p \cup M'_p))$.
(2) $\forall u_1, u_2 (u_1 \in U \wedge u_1 \subseteq u_2 \rightarrow u_2 \in U)$.
(3) $\forall u_1, u_2 (u_1 \in U \wedge u_2 \in U \rightarrow u_1 \cap u_2 \in U)$.
(4) $\forall u (u \in U \rightarrow \Delta(M_p \cup M'_p) \subseteq U)$.
(5) $\forall u (u \in U \rightarrow u^{-1} \in U)$.
(6) $\forall u_1 \exists u_2 (u_1 \in U \rightarrow u_2^2 \subseteq u_1 \wedge u_2 \in U)$.

En el capítulo 2.7 ya se dio la justificación intuitiva de cada una de estas condiciones de definición de las uniformidades desde el punto de vista de la aproximación empírica. Esa justificación se puede traspasar al caso de la aproximación interteórica sin dificultad. En este aspecto particular no hay diferencia esencial entre las intuiciones correspondientes a la aproximación «interna» y a la «externa».

Asimismo se pueden traducir al presente caso otras nociones y resultados del capítulo 2.7: la noción de uniformidad *empírica* se puede definir también para $M_p \cup M'_p$ en vez de M_p, y por tanto también se puede probar el correspondiente «teorema de inducción», que permite obtener uniformidades sobre las estructuras no-teóricas a partir de uniformidades empíricas sobre $M_p \cup M'_p$. Finalmente, la introducción del operador modelo-teórico $\sim$, que «emborrona» la aplicación de predicados y relaciones a modelos, y por consiguiente la «lógica de la aproximación» esbozada en 2.7, puede emprenderse aquí del mismo modo que allí.

En el capítulo 2.7 hicimos notar que, en las aproximaciones empíricas, no se usan todos los conjuntos borrosos de una uniformidad U dada, sino sólo los que no son ni demasiado «grandes» ni demasiado «pequeños», y tienen algunas otras propiedades intuitivamente razonables. Por ello introdujimos allí la noción de conjunto borroso *admisible*. Esta noción es aún más importante, si cabe, para tratar las cuestiones de la aproximación interteórica. Es obvio que no vamos a admitir como genuina una aproximación entre teorías basada en conjuntos borrosos formalmente posibles, pero intuitivamente absurdos, por ejemplo, demasiado «grandes». Si no impusiéramos algún tipo de restricción a las uniformidades usadas en aproximaciones interteóricas, resultaría que prácticamente cualquier teoría sería una aproximación de cualquier otra. Por ello, también aquí supondremos que disponemos de la noción de conjuntos borrosos

admisibles, y que cuando hablamos de aproximación, la referimos a conjuntos tales.

Ahora bien, en los casos «ordinarios» de aproximación entre dos modelos (ya sean de la misma teoría o de teorías distintas), establecemos la comparación entre ellos tomando *un* determinado conjunto borroso admisible y decimos que la aproximación es válida dentro de *ese* conjunto borroso particular. Es decir, anteponemos al enunciado modelo-teórico aproximativo un cuantificador existencial sobre conjuntos borrosos admisibles. (Recuérdese el modo como se define el operador $\sim$ de la lógica de la aproximación.) No obstante, parece que en las aproximaciones interteóricas, al menos en las más interesantes, se pretende establecer un enunciado modeloteórico aproximativo mucho más fuerte: se pretende que la aproximación entre un modelo de una teoría y su modelo correspondiente en la otra se dé dentro de *cualquier* conjunto borroso *admisible*. Precisamente el caso Kepler-Newton que analizaremos en la última parte de este capítulo posee esta característica, y es verosímil que ello también ocurra en los casos óptica geométrica-óptica ondulatoria y mecánica clásica-relatividad especial, por lo menos. En tales casos queremos afirmar no sólo que un modelo x se aproxima a otro y dentro del margen dado por cierto u admisible, sino que x se aproxima a y dentro de cualquier u admisible escogido arbitrariamente. Con otras palabras, en vez de cuantificar existencialmente sobre los conjuntos borrosos, cuantificamos universalmente. Esta situación es lo suficientemente interesante y real como para que la distingamos sistemáticamente de la más normal e introduzcamos la noción especial de «aproximación estricta». Simbólicamente usaremos para ella el operador $\approx$ «$x \approx y$» significa:

$$\forall u \in \alpha(\langle x, y \rangle \in u).$$

Análogamente a lo hecho en 2.7 podríamos introducir una serie de usos derivados de este operador aproximativo doble. Por ejemplo, «$P(\widetilde{\widetilde{x}})$» significa:

$$\forall u \in \alpha \exists y(P(y) \wedge \langle x, y \rangle \in u),$$

y así sucesivamente.

2. El esquema general para una aproximación interteórica

Nuestro punto de partida es que la aproximación interteórica es una noción con cierta clase de «parentesco» con la reducción exacta (sneediana). En cierta medida, podría decirse que se trata de una «reducción borrosa». Pero hay que matizar esta expresión, dado que de buenas a primeras no está claro dónde y cómo hay que emborro-

nar la reducción. Además, tendremos que apartarnos un poco del concepto original de Sneed de la reducción exacta y modificarlo para obtener una versión más realista (incluso prescindiendo del contexto aproximativo).

En el capítulo 2.8 ya hemos dado una versión simplificada de la reducción exacta a lo Adams-Sneed. Ahora vamos a ser un poco más precisos y completos. Si prescindimos de las complicaciones debidas a la existencia de ligaduras y especializaciones, los elementos básicos de esa noción de reducción son los siguientes. Una reducción de T a T' es una relación interteórica, llamémosla ρ, entre los modelos potenciales de T y los de T' (es decir, $\rho \subseteq M_p \times M'_p$) tal que:

1) A cada modelo potencial de T le corresponde al menos un modelo potencial de T', es decir, $D_1(\rho) = M_p$.
2) ρ^{-1} (es decir, la relación inversa de ρ) es un función. (A una relación que satisfaga estas dos primeras condiciones la llamaremos una «*cuasi-reducción*».)
3) Los modelos de la teoría reducida se pueden «derivar» de los de la reductora:

$$\langle x, x' \rangle \in \rho \wedge x' \in M' \rightarrow x \in M.$$

En caso de que ρ estuviera definida sobre todo M' (lo que, sin embargo, no parece ser el caso usual), podríamos escribir la condición anterior de la forma

$$\bar{\rho}^{-1}(M') \subseteq M,$$

donde $\bar{\rho}$ es la relación inducida por ρ sobre *clases* de modelos potenciales, es decir, al nivel conjuntista inmediatamente superior. (Esta posibilidad resultará importante para nuestro ejemplo.)

4) Los dominios de aplicaciones propuestas A y A' también están en correspondencia-ρ o, para ser más exactos, en la relación noteórica paralela a la correspondencia-ρ (es decir, la relación que ρ induce al nivel de los modelos parciales), a la cual llamaremos v_ρ.

No voy a ocuparme aquí de la justificación de estas ideas básicas sobre la reducción estructural. En parte se ha dado en el capítulo anterior. Para mayores detalles hay que acudir a *The Logical Structure of Mathematical Physics* de Sneed y *Theorienstrukturen und Theoriendynamik* de Stegmüller. Un concepto estructural aún más elaborado de reducción exacta, que toma en cuenta todas las complejidades de este enfoque, se presenta en el artículo de Balzer y Sneed sobre «redes teóricas» de 1977. Para nuestros propósitos presentes no es necesario tener en cuenta esas complicaciones adicionales. El único resultado de Balzer y Sneed que conviene recordar es el de que una

reducción ρ al nivel teórico de $M_p \times M'_p$ siempre induce una reducción v_ρ al nivel no-teórico de $M_{pp} \times M'_{pp}$.

Introduzcamos ahora el esquema general para una aproximación interteórica emprendida a través de una reducción ρ. Como ya se ha observado al principio del capítulo, no se *definirá* un concepto general que abarque todas las aproximaciones interteóricas posibles. Hay *muchas* clases distintas de relaciones ρ aproximativas; todo lo que puedo hacer aquí es indicar cómo se pueden obtener sistemáticamente a partir del esquema general. Las diversas relaciones ρ aproximativas provienen de los diversos modos como podemos «emborronar» los términos de la relación, y también en la posibilidad de aplicar o bien el operador simple $\sim$ o bien el doble $\approx$ para el emborronamiento.

Consideremos más detalladamente estas diversas posibilidades. En primer lugar, puede emborronarse la teoría reducida o bien la reductora; en segundo lugar, el emborronamiento puede tener lugar al nivel teórico del núcleo en cualquiera de los dos casos o bien al nivel «empírico» del dominio de aplicaciones. Podemos introducir la siguiente notación sistemática para clasificar todos los casos posibles. El índice 1 se refiere a la teoría reducida, mientras que el índice 2 se refiere a la reductora. Además, si se emborrona el núcleo, escribiremos el índice de que se trate como superíndice de ρ; si se emborronan las aplicaciones propuestas, el índice aparecerá como subíndice de ρ. Así obtenemos:

ρ^1: una aproximación interteórica en la que sólo se emborrona la teoría reducida al nivel del núcleo;
ρ_1: una aproximación interteórica en la que sólo se emborrona la teoría reducida al nivel de las aplicaciones;
ρ^2: análogamente para la teoría reductora al nivel del núcleo;
ρ_2: análogamente para la teoría reductora al nivel de las aplicaciones.

Por combinación de estas cuatro posibilidades básicas, obtenemos casos «mixtos» de aproximación. Por ejemplo, ρ_1^1 es un caso en que el operador aproximativo se aplica tanto al núcleo como al dominio de aplicaciones de la teoría reducida, ρ_2^1 es un caso en que la teoría reducida se emborrona a nivel del núcleo y la teoría reductora a nivel de las aplicaciones, y así sucesivamente. También podemos obtener casos «dobles» de uso simultáneo del operador aproximativo a ambas teorías en los mismos niveles: por ejemplo, ρ_{12} es un caso en que tanto la teoría reducida como la reductora aparecen emborronadas a nivel de las aplicaciones propuestas, pero en cambio los núcleos respectivos se relacionan entre sí de manera exacta; $\rho_{1\,2}^{1}$ es un caso en que el nivel empírico de ambas teorías, así como el núcleo de la teoría reducida, se relacionan emborronados, etc.

Posibilidades adicionales se obtienen al sustituir el operador simple $\sim$ por el doble $\approx$ en todos los casos mencionados, o también

combinando el operador doble con el simple (por ejemplo, sería imaginable emprender una aproximación estricta a nivel de las aplicaciones y sólo una aproximación simple a nivel de los núcleos). Está claro que se podría construir una lista sistemática de todas esas nociones posibles de aproximación interteórica, y en un sentido trivial obtendríamos con esa lista «el» concepto general de aproximación interteórica. Sin embargo, no es en este sentido trivial, como mera disyunción de casos, que hablábamos al principio del capítulo del problema no resuelto de hallar un concepto general de aproximación interteórica. Lo que quisiéramos es un concepto definido de una manera simple y elegante, del cual se pudieran desprender, aplicando ciertos criterios generales, las numerosas alternativas de la lista mencionada. Como dijimos, la posibilidad de introducir formalmente un concepto tal es todavía una cuestión abierta. *Otra* cuestión abierta es la siguiente. Considerando la lista de posibilidades de aproximación interteórica que acabamos de esbozar, no está claro que todas esas diversas alternativas no sean meras posibilidades formales, sino que también correspondan a situaciones reales que pueden surgir en el análisis de la ciencia empírica. Después de todo, podría ser que algunas de ellas deban ser excluidas de entrada por consideraciones no-formales, quizá pragmáticas de validez general para la ciencia. Si esto fuera así, también el concepto general de aproximación interteórica que es nuestro *desideratum*, fuera cual fuera su definición, debería ser introducido de tal modo que tales posibilidades formalmente posibles, pero materialmente inapropiadas, quedaran excluidas del dominio de dicho concepto.

Las dificultades generales apuntadas no nos impiden, de todos modos, trabajar con los elementos de la lista esbozada para reconstruir casos concretos de relaciones interteóricas. En lo que sigue, vamos a seleccionar una de las posibilidades de la lista, a definirla formalmente y a comprobar su aplicación en un ejemplo concreto. La aproximación que vamos a introducir es del tipo ρ_2^1. Hay dos razones por las que he escogido este caso. Una es que puede servir de «paradigma» para otros casos de aproximaciones interteóricas, que se podrán definir según el patrón de la introducida aquí emprendiendo simplemente algunas modificaciones en la colocación de los operadores aproximativos. Otra es que probaremos que el caso Kepler-Newton es una instancia de aproximación tipo ρ_2^1.

Para hacer más «legible» la definición utilizaremos las siguientes convenciones y abreviaciones:

a) Usaremos el símbolo «σ» para indicar la relación de especialización entre dos núcleos de elementos teóricos distintos (recuérdese la definición de esta relación en 2.4, p. 114): si H y K son dos núcleos, entonces «$H\sigma K$» significa que H es una especialización de K.

b) Supongamos que tenemos una cuasi-reducción ρ establecida

entre los conjuntos M_p y M'_p de los núcleos H y H'; y sea $H'_o \sigma H'$. Entonces la expresión «$\rho^{-1}(H'_o) \subseteq H$» será una abreviación para la siguiente condición:

$$\forall X, X'(\langle X, X' \rangle \in \tilde{\rho} \wedge X' \in \mathscr{P}(M'_o) \cap C'_o \rightarrow X \in \mathscr{P}(M) \cap C),$$

donde presuponemos que cada $x' \in M'_o$ tiene una imagen-ρ^{-1} en M.

D2: Dados dos elementos teóricos $T = \langle K, A \rangle$ y $T' = \langle K', A' \rangle$, diremos que $\langle T, \rho, T', A'_o \rangle$ es una aproximación tipo ρ_2^1 de T a T' en A'_o si y sólo si:

(1) ρ es una cuasi-reducción de T a T'.
(2) $A'_o \subset A'$.
(3) $\forall H(H \sigma K \rightarrow \exists K'_o, T'_o(T'_o = \langle K'_o, A'_o \rangle \wedge K'_o \sigma K' \wedge \rho^{-1}(K'_o) \subseteq \tilde{H})$.
(4) $\langle A, \tilde{A}'_o \rangle \in \bar{\bar{\nu}}_\rho$.

Las ideas intuitivas que están detrás de esta definición son las siguientes. T es la teoría «menos desarrollada» o «más pobre» que queremos reducir, y T' es la teoría «más desarrollada» o «más rica», destinada a jugar el papel reductor. Se supone que T se aproxima a T', pero sólo dentro de un subconjunto propio A'_o del dominio total de aplicaciones A' de T'. En términos más intuitivos, T' «cubre más» que T. A'_o es el dominio «explicado» tanto por T como por T', pero es justamente sólo un subdominio de A'. Este A'_o se halla en una correspondencia cuasi-reductiva ρ con la *totalidad* del dominio A de T, pero no se trata de una correspondencia reductiva exacta: es sólo una correspondencia *aproximada* del tipo ρ_2^1, tal como se expresa en la condición (4). La condición (3) significa que todas las especializaciones de T tienen su imagen especular aproximada en T', de tal modo que los modelos de cada especialización de T se pueden «derivar» aproximadamente de los de la especialización correspondiente en T', en el sentido explicado más arriba.

Podemos obtener un concepto de aproximación tipo ρ_2^1 *estricta* en vez del de aproximación tipo ρ_2^1 (simple) que acabamos de definir si sustituimos el operador $\sim$ en todas partes por el operador $\approx$, siendo por lo demás la estructura de la definición la misma.

3. La relación entre la teoría de Kepler y el elemento gravitacional de la teoría de Newton

En su trabajo de 1973 Scheibe explicitó las condiciones particulares que determinan la relación entre la teoría planetaria de Kepler y la

teoría newtoniana de la gravitación. Su tratamiento del problema es totalmente concreto en el sentido de que no intenta derivar las condiciones por él encontradas de un esquema interteórico general. La última parte de este capítulo presenta una derivación de las condiciones particulares de la relación Kepler-Newton a partir del concepto general de una aproximación tipo ρ_2^1 estricta. Pero antes es necesario que reformulemos las condiciones de Scheibe en nuestros propios términos para poderlas insertar en el presente marco general.

El primer paso que hay que dar es identificar las estructuras que deben ser comparadas: las correspondientes a la teoría de Kepler y las correspondientes a la teoría de Newton. Contrariamente a lo que se insinúa en algunos libros de texto de mecánica, sostenemos que aquí tenemos un caso en que se ponen en relación dos teorías realmente diferentes, en el sentido fuerte de que no sólo las leyes empíricas incluidas en ellas, sino también las estructuras conceptuales son distintas. Son «teóricamente inconmensurables» en el sentido de Stegmüller (véase más arriba). Las estructuras no-teóricas de ambas teorías son del mismo tipo: descripciones cinemáticas de sistemas de partículas. Pero las «superestructuras» teóricas son diferentes. Para comprender esto hay que tener bien presentes las definiciones de los predicados conjuntistas respectivos, que determinan los dos conjuntos de modelos que hay que comparar. El predicado correspondiente al elemento teórico newtoniano, llamado «dinámica gravitacional de partículas», ya se ha definido en el capítulo 2.5, *D*18, p. 128, por lo que aquí nos ahorramos su formulación explícita. Recordamos solamente que es un predicado que se aplica a estructuras de la forma $\langle P, T, s, m, \bar{G}, \bar{F} \rangle$, donde los términos *T*-teóricos son m, $\bar{G}$ y $\bar{F}$ (= los conceptos «dinámicos»), siendo la ley característica de este elemento teórico (aparte del Segundo Principio de Newton) la ley de gravitación universal. Lo que nos falta ahora es el predicado conjuntista correspondiente a la teoría de Kepler. Por supuesto, que ni la definición que sigue de «sistema kepleriano» ni la que dimos de «dinámica gravitacional» en el capítulo 2.5 pretenden ser históricamente fieles. Las formulaciones de estos elementos teóricos dadas en la segunda mitad del siglo XX tendrán que diferir necesaria y considerablemente de las dadas en el siglo XVII. Pero esta divergencia no es de capital importancia para nuestra discusión, ni siquiera desde el punto de vista diacrónico, dado que lo que interesa dentro de la perspectiva estructural no es la comparación de *enunciados* (fórmulas), los cuales pueden diferir mucho a través de la historia, sino la comparación de *modelos*. Y éstos no parecen haber cambiado sustancialmente desde la época de Kepler y Newton, respectivamente.

La definición que damos a continuación de «sistema kepleriano» se atiene básicamente a la proporcionada por Scheibe en 1973.

*D*3: x es un *sistema kepleriano* si y sólo si existen P, T, s, μ, tales que:

(1) $x=\langle P, T, s, \mu\rangle$.
(2) P es un conjunto finito, no-vacío.
(3) T es un intervalo de números reales.
(4) s es una función de $P\times T$ en $\mathbb{R}^3$ y s es dos veces diferenciable con respecto a su segundo argumento.
(5) μ es una función de P en $\mathbb{R}$.
(6) Existe un $p_o\in P$ tal que, para todo $t\in T$:

(a) $D_t^2 s(p_o, t)=0\wedge\mu(p_o)\neq 0$.
(b) Para todo $p\neq p_o$:
(b_1) $\mu(p)=0$.
(b_2) $D_t^2 s(p, t)=\mu(p_o)\cdot(s(p, t)-s(p_o, t))\cdot/s(p, t)-s(p_o, t)/^{-3}$.
(b_3) $1/2\cdot/D_t s(p,t)-D_t s(p_o,t)/^2-\mu(p_o)\cdot/s(p,t)-s(p_o,t)/^{-1}<0$.

Como se ve, los tres primeros integrantes de un modelo kepleriano, P, T, s, son los mismos que para el caso del predicado «dinámica gravitacional de partículas», y su interpretación es la misma: se trata del triplo que satisface el predicado «cinemática de partículas», ya introducido en el capítulo 2.5, *D*1, p. 117. La magnitud μ, en cambio, es nueva: es la llamada «constante de Kepler». Intuitivamente, tiene algo que ver con lo que podríamos llamar «la masa del sol» en términos newtonianos (en efecto, la partícula seleccionada p_o representa «el sol» de un sistema kepleriano), pero no puede identificarse con ella, ni conceptual ni numéricamente. Es simplemente un concepto específico de la teoría kepleriana y sólo tiene sentido dentro de ella. Es, pues, un concepto Kepler-teórico. Por lo tanto, los modelos parciales de la teoría de Kepler son los mismos que los de la teoría newtoniana, a saber, estructuras que satisfacen el predicado *CP* del capítulo 2.5; y los modelos potenciales se obtienen en un caso al añadir μ a esos modelos parciales y en el otro al añadirles m, $\bar{G}$ y $\bar{F}$.

El conjunto de estructuras $\langle P, T, s, \mu\rangle$ que satisfacen el predicado «sistema kepleriano» es el conjunto de modelos actuales de la teoría de Kepler. La condición (6) representa una formulación moderna de las llamadas «leyes de Kepler»[42]. A este conjunto de modelos lo podríamos designar por «$M[Kep]$», pero para abreviar lo llamaremos simplemente «K». El conjunto de modelos potenciales de esta teoría, $M_p[Kep]$, es el conjunto de estructuras que satisfacen las condiciones (1)-(5) de la definición anterior, y el conjunto de los modelos parciales

[42] La equivalencia entre las ecuaciones y desigualdades de (6) y la formulación clásica de las leyes de Kepler puede encontrarse en libros de texto de mecánica.

es CP. Podríamos añadir una ligadura de identidad para μ, de modo análogo a como se han introducido ligaduras de identidad para los conceptos dinámicos en la mecánica newtoniana. Pero la introducción de ligaduras no añadiría nada sustancial a nuestro argumento sobre la relación entre ambas teorías, por lo que, en aras de la simplicidad, las omitimos. En consecuencia, el núcleo de la teoría de Kepler, al que llamaremos «*Kep*», será: $Kep = \langle M_p[Kep], CP, K\rangle$, y el núcleo de la teoría newtoniana de la gravitación será, usando una convención notacional análoga: $New = \langle M_p[New], CP, N\rangle$.

Llamemos A_K al dominio de aplicaciones propuestas de Kep y A_N al dominio de aplicaciones propuestas de New. En el capítulo 2.4 hemos argüido que, en general, el dominio de aplicaciones propuestas de un elemento teórico deberá ser considerado no como un conjunto «amorfo» de modelos parciales, sino como un conjunto de clases homogéneas de modelos parciales. En el caso actual esto significaría que la determinación conjuntista de los dominios respectivos sería: $A_K \subset \mathcal{P}(CP)$ y $A_N \subset \mathcal{P}(CP)$. En principio, aquí deberíamos atenernos a esta determinación. Sin embargo, el ejemplo que estamos considerando es de tal naturaleza que permite simplificar la determinación conjuntista, lo cual a su vez implica una cierta simplificación de la relación entre los dominios comparados –ya de por sí bastante complicada, como veremos. En efecto, al comparar A_K con A_N no estamos realmente poniendo en relación clases enteras de cinemáticas, sino casos individuales de órbitas planetarias (por ejemplo, la órbita de Marte «vista por Kepler» con la órbita de Marte «vista por Newton»). Es decir, en este caso particular podemos tomar de manera natural el dominio de aplicaciones de cada teoría efectivamente como un conjunto de modelos parciales. Claro que, para no contradecir el principio general de determinación de los dominios de aplicaciones expuesto en 2.4, podríamos tomar en este caso cada dominio como un conjunto de clases unitarias (por ejemplo, $A_K = \{\{y_1\}, \{y_2\}, \ldots, \{y_n\}\}$). No obstante, para simplificar el tratamiento formal es conveniente «rebajar» el nivel conjuntista de los dominios. En lo sucesivo (y sólo para este ejemplo) aceptaremos, pues, que $A_K \subset CP$ y $A_N \subset CP$.

Siguiendo a Scheibe, supondremos que A_K sólo contiene cinemáticas de dos partículas. Esta restricción parece estar históricamente justificada. La teoría de Kepler sólo se aplica por separado a la órbita de cada planeta alrededor del sol, no al sistema planetario como sistema de elementos en interacción. Formulemos este supuesto explícitamente como postulado metateórico:

$P1$: Para todo $y \in A_K$: $/P_y/ = 2$.

(Por lo demás, este supuesto no significa una limitación esencial del alcance de nuestra discusión.)

Otro postulado que vamos a admitir es éste:

$P2$: $A_N \subset \bar{r}(N) =: \{y / y \in \widehat{CP} \wedge \exists x (x \in N \wedge y = r(x)\}$

Esto es, admitimos que todas las aplicaciones propuestas de la teoría newtoniana de la gravitación pueden subsumirse efectivamente bajo el núcleo newtoniano (o sea, realmente satisfacen el Segundo Principio junto con la ley de la gravitación). Dado el estado actual de nuestro conocimiento, este supuesto parece bastante firme.

Obsérvese que la totalidad de la teoría de Kepler consiste únicamente del elemento teórico $T_K = \langle Kep, A_K \rangle$; no tiene ulteriores especializaciones. La red teórica de Kepler se reduce a una red «degenerada» de un solo elemento. Por otro lado, la red de la mecánica clásica es ciertamente una enorme red teórica. Pero aquí sólo estamos interesados en uno de sus elementos, a saber, $T_N = \langle New, A_N \rangle$. En consecuencia, la comparación interteórica se establecerá entre el único elemento teórico que constituye la teoría de Kepler y *un* elemento teórico de la red newtoniana. No es necesario que nos ocupemos de especializaciones para la aproximación.

Finalmente, es conveniente introducir los siguientes conceptos auxiliares. Sea U una uniformidad consistente en conjuntos borrosos definidos sobre $\mathscr{P}(CP \times CP)$ por medio de la siguiente condición:

$$\langle y, y' \rangle \in u_\varepsilon \in U$$

si y sólo si:

$$\forall p \in P_y \cap P_{y'} \forall t \in T_y \cap T_{y'} : |s_y(p, t) - s_{y'}(p, t)| < \varepsilon^{43}.$$

Entonces,

$$(A_N)_\varepsilon =: \{y / \exists y' (y' \in A_N \wedge \langle y, y' \rangle \in u_\varepsilon)\},$$
$$(A_K)_\varepsilon =: \{y / \exists y' (y' \in A_K \wedge \langle y, y' \rangle \in u)\}.$$

De modo enteramente análogo podríamos definir $(R(N))_\varepsilon$ y $(R(K))_\varepsilon$.

Pasemos ahora a la reconstrucción de la relación entre la teoría de Kepler y la de Newton. Reformulando el análisis proporcionado por Scheibe mediante el aparato conceptual que hemos introducido, dicha relación vendría sintetizada en las tres condiciones siguientes[44]:

---

[43] Es decir, cada conjunto borroso u_ε de la uniformidad está determinado por un ε particular.

[44] La formulación presente no es literalmente idéntica a la de Scheibe. En primer lugar, la formulación informal de Scheibe ha tenido que traducirse al marco modelo-teórico. Además, en su artículo Scheibe no hace una distinción explícita entre un nivel T-teórico y uno T-no-teórico de conceptos, aunque sostengo que la distinción está implícita en su tratamiento. De todos modos, éstas son diferencias de detalle en la formulación, que afectan la letra, no el espíritu de la reconstrucción. En una comunicación verbal, Scheibe concordó con esta «traducción» de sus propios términos.

(*S*1) $\forall\varepsilon>0$: $A_K \subseteq (A_N)_\varepsilon$.

(Esto corresponde a la condición (20) del artículo de Scheibe, p. 114.)

(*S*2) $\exists H_i \forall\varepsilon>0$: $H_i \sigma N \wedge R(H_i) \subseteq (R(K))_\varepsilon$.

(Esto corresponde a las condiciones (21) y (22) del artículo de Scheibe, p. 115.)

(*S*3) $\exists\varepsilon>0$: $A_N - (A_K)_\varepsilon \neq \varnothing$.

Esta última condición no corresponde a ninguna condición formalmente establecida en el trabajo de Scheibe. No obstante, corresponde a la intención de sus observaciones informales sobre la «superiodidad» de la teoría de Newton sobre la de Kepler. En un pasaje hacia el final del artículo dice:

> «Para las explicaciones [=reducciones] aproximativas tenemos en cambio la asimetría de que la inmensa mayoría de entornos-ε de casos newtonianos no contiene ningún caso kepleriano, mientras que, a la inversa, en cada entorno-ε de un caso kepleriano se halla un caso newtoniano» (Scheibe, *op. cit.*, p. 116).

Hay que notar, sin embargo, que el contenido de (*S*3) es más débil que la formulación de Scheibe que acabamos de citar, según la cual «la inmensa mayoría» de aplicaciones propuestas de la teoría de Newton no son, ni siquiera ε-aproximadamente, aplicaciones propuestas de la teoría de Kepler. A pesar de ello, la formulación más débil dada en (*S*3) es suficiente para garantizar la superioridad de la teoría newtoniana, y por otra parte, simplifica la discusión. Probablemente podríamos reconstruir la versión más fuerte esbozada por Scheibe utilizando un aparato topológico más elaborado, como Scheibe mismo ha señalado[45], pero no creo que esta complicación adicional tenga ninguna relevancia particular para el argumento. Es suficiente establecer la presente condición más débil para mostrar la «superioridad de Newton con respecto a Kepler», y esto es todo lo que necesitamos.

4. La relación Kepler-Newton como un caso de aproximación tipo ρ_2^1 estricta

Hemos llegado al punto en que podemos mostrar la aplicabilidad de nuestro esquema general de aproximación interteórica a un caso concreto: vamos a derivar las condiciones (*S*1)-(*S*3) de la relación

[45] En una comunicación verbal al autor.

Kepler-Newton a partir del concepto general de aproximación tipo ρ_2^1. El postulado interteórico, a partir del cual hay que deducir (*S*1)-(*S*3) es el siguiente:

*P*3: $\langle T_K, \rho_{KN}, T_N, A_o^N \rangle$ es una aproximación tipo ρ_2^1 estricta de T_K a T_N en A_o^N, donde ρ_{KN} y A_o^N se determinan como sigue:

(a) ρ_{KN} es una cuasi-reducción definida sobre $M_p[Kep] \times M_p[New]$ que satisface la condición (18a) dada por Scheibe con respecto a las magnitudes teóricas (Scheibe, *op. cit.*, p. 114):

$$m(p_o) \cdot \left(1 + \frac{m(p_1)}{m(p_o)}\right)^{-2} = \mu(p_o),$$

y con respecto a los conceptos no-teóricos respectivos vamos a exigir simplemente la identidad de los mismos bajo ciertas circunstancias especificadas. La cuasi-reducción no-teórica inducida es:

$\langle y_K, y_N \rangle \in v\rho_{KN}$ si y sólo si:

$$\exists x_K, x_N: x_K \in K \wedge x_N \in N \wedge y_K = r(x_K) \wedge y_N = r(x_N) \wedge \langle x_K, x_N \rangle \in \rho_{KN} \wedge y_K = y_N.$$

Debería estar claro en este punto que tanto ρ_{KN} como $v_{\rho KN}$ satisfacen las condiciones de una cuasi-reducción.

(b) A_o^N queda determinado como $A_o^N = \bar{r}(H_o^N)$, donde H_o^N es una especialización particular de N, a saber, la determinada por los axiomas adicionales que Scheibe formula como sus condiciones (23) y (24) (Scheibe, *op. cit.*, p. 115). Scheibe ha mostrado que este H_o^N es aplicable a todos los casos de cinemáticas de dos partículas. La forma particular de esta especialización, no obstante, no tiene por qué preocuparnos aquí. Basta con que exista[46].

Ahora queremos probar que las condiciones (*S*1)-(*S*3) se deducen de *P*3, es decir, de las condiciones de la aproximación tipo ρ_2^1 estricta que postulamos para el caso Kepler-Newton. Primero necesitamos algunos lemas que se siguen fácilmente del postulado y de la definición de aproximación tipo ρ_2^1.

Lema 1: $A_o^N \subset A_N$.

Esto se sigue de *P*3 y *D*2-(2).

[46] Por el postulado *P* 1, debería estar claro por qué escogemos este A_o^N: los sistemas de dos partículas que lo integran son el «reflejo newtoniano» de las aplicaciones de la teoría de Kepler.

Lema 2: $\exists H_i^N (H_i^N \subseteq N \wedge A_o^N \subseteq \bar{r}(H_i^N) \wedge \rho_{KN}^{-1}(H_i^N) \in \tilde{K})$.

Esto se sigue de *P*3, *D*2-(3), y del hecho de que no consideramos ligaduras y de que la única especialización de K que hay que tomar en cuenta es K mismo.

Lema 3: $\langle A_K, A_o^N \rangle \in \bar{v}\rho_{KN}$.

Esto se sigue de *P*3 y *D*2-(4).

Teorema 1: *D*2 y *P*3 implican (*S*1).

Prueba: Por *Lema* 3, tenemos $\langle A_K, A_0^N \rangle \in \bar{v}_{\rho KN}$.

Por definición de $\bar{v}_{\rho KN}$ esto significa que:

$$\forall y_K \in A_K \exists y_N \in A_o^N : \langle y_K, y'_N \rangle \in v_{\rho KN}.$$

Por definición de aproximación estricta, esto implica que

$$\forall y_K \in A_K \exists y_N \in A_o^N \forall \varepsilon > 0 \exists y'_N \in A_N:$$
$$\langle y_K, y'_N \rangle \in v_{\rho KN} \wedge \langle y_N, y'_N \rangle \in u_\varepsilon.$$

Dado que $v_{\rho KN}$ se ha construido como la identidad, obtenemos:

$$\forall y_K \in A_K \exists y_N \in A_o^N \forall \varepsilon > 0: \langle y_K, y_N \rangle \in u_\varepsilon.$$

Por definición de $(A_N)_k$ se sigue que:

$\forall y_K \in A_K \forall \varepsilon > 0$: $y_K \in (A_N)_k$, esto es, que $\forall \varepsilon > 0$: $A_K \subseteq (A_N)_k$.

Teorema 2: *D*2 y *P*3 implican (*S*2).

Prueba: Por *Lema* 2, existe un H_o^N tal que:

$$H_o^N \subseteq N \wedge A_o^N \subseteq \bar{r}(H_o^N) \wedge \rho_{KN}^{-1}(H_o^N) \subseteq \tilde{K}.$$

El único miembro de esta conyunción que necesitamos aquí es el último. Por el resultado antes mencionado de Balzer y Sneed en su articulo conjunto de 1977 sabemos que la cuasi-reducción «teórica» $\bar{\rho}_{KN}^{-1}$ implica la cuasi-reducción «no-teórica» derivada $\bar{v}\rho_{KN}$, la cual «refleja» al nivel no-teórico las correspondencias establecidas por la primera. En este caso esto significa que:

$$v_{\rho KN}^{-1}(\bar{r}(H_o^N)) \subseteq \bar{r}(\tilde{K}).$$

Dado que, por construcción, $v_{\rho KN}$ es una identidad, $\bar{r}(H_o^N) \subseteq \bar{r}(\breve{K})$. Aplicando a esta inclusión la definición de aproximación estricta,

$$\forall y_N \in \bar{r}(H_o^N) \exists y_K \in \bar{r}(K) \forall \varepsilon > 0 : \langle y_N, y_K \rangle \in u_\varepsilon .$$

Por definición de $(\bar{r}(K))_\varepsilon$, lo anterior equivale a:

$$\forall y_N \in \bar{r}(H_o^N) \forall \varepsilon > 0 : y_N \in (\bar{r}(K))_\varepsilon .$$

o sea,

$$\forall \varepsilon > 0 : \bar{r}(H_o^N) \subseteq (\bar{r}(K))_\varepsilon .$$

De ahí obtenemos (*S*2) por cuantificación existencial sobre H_o^N.

Teorema 3: *D*2 y *P*3 implican (*S*3).

Prueba: La validez de (*S*3) quedará garantizada si logramos encontrar un $z_N \in A_N$ tal que $z_N \notin (A_K)_\varepsilon$ para algún ε.

Escojamos un sistema de dos partículas que sea una aplicación propuesta de *New* y del que sepamos que es una parte física de una aplicación de *New* «más grande»; por ejemplo, podemos tomar el sistema Tierra-Luna, que es parte del sistema Tierra-Luna-Sol. (Naturalmente, es a través de las presentaciones estándar de T_N que sabemos que un sistema tal existe.) Prescindiendo de las funciones teóricas, este sistema será una cinemática de partículas, llamémosla y_N. Dado que la cardinalidad de P_{y_N} es 2, tendremos que $y_N \in A_o^N$. Por hipótesis, también sabemos que existe un $z_N \in A_N$ con $P_{z_N} \supset P_{y_N}$; por lo tanto, la cardinalidad de P_{z_N} es mayor que 2. (Por ejemplo z_N podría ser la cinemática correspondiente al sistema Tierra-Luna-Sol.)

Dado que $z_N \in A_N \subset \bar{r}(N)$ (por *P*2), podemos aplicar la ley de gravitación en particular a z_N. Ahora bien, es un ejercicio bien conocido de mecánica clásica usar la ley de gravitación para derivar las órbitas de un sistema de dos partículas y las órbitas de las mismas partículas cuando ese sistema se extiende a un sistema de tres partículas. Y sabemos que resultarán diferencias no negligibles en las órbitas de las partículas calculadas en ambos casos. De ahí que obtengamos para el sistema cinemático escogido:

$$(*) \quad \forall p \in P_{y_N} \forall t \in T_{z_N} \cap T_{y_N} : /s_{z_N}(p, t) - s_{y_N}(p, t)/ \geqslant k,$$

para cierto $k > 0$.

Tomemos ahora un ε_o con $\varepsilon_o < k/2$. Y tomemos y_K con $y_K = y_N$, es decir, $y_K \in \bar{r}(K)$. Si $y_K \notin (A_K)_{\varepsilon_o}$, entonces el teorema ya queda probado tomando $y_N \in A_N$ en vez de z_N. Supongamos, sin embargo, que se da $y_K \in (A_K)_{\varepsilon_o}$.

Esto significaría:

$$\exists y'_K \in A_K : \langle y_K, y'_K \rangle \in u_{\varepsilon_o}.$$

Supongamos ahora que también fuera el caso que $z_N \in (A_K)_{\varepsilon_o}$. Entonces, dado que $P_{y_K} \subset P_{z_N}$, también debería ser el caso que, en particular,

$$\langle z_N, y'_K \rangle \in u_{\varepsilon_o}.$$

Entonces tendríamos:

$$\langle y_K, y'_K \rangle \in u_{\varepsilon_o} \wedge \langle z_N, y'_K \rangle \in u_{\varepsilon_o}.$$

Esto implica:

$$\forall p \in P_{z_N} \cap P_{y_K} \forall t \in T_{z_N} \cap T_{y_K} :$$
$$/s_{y_K}(p, t) - s_{y'_K}(p, t)/ < \varepsilon_o \wedge /s_{z_N}(p, t) - s_{y'_K}(p, t)/ < \varepsilon_o.$$

Por aritmética elemental:

$$/s_{y_K}(p, t) - s_{z_N}(p, t)/ < 2\varepsilon_o < k.$$

Esta última desigualdad contradice la desigualdad en (*). Por lo tanto, $z_N \notin (A_K)_{\varepsilon_o}$, mientras que $z_N \in A_N$, tal como había que mostrar.

De la conyunción de estos tres teoremas aquí probados se desprende que la relación entre la teoría de Kepler y la de Newton es un caso de aproximación tipo $\rho\frac{1}{2}$, tal como se ha postulado en $P3$. En consecuencia, podemos afirmar que este concepto de aproximación interteórica es adecuado al menos a este caso.

2.10. MARCOS TEORICOS: EL EJEMPLO DE LA TERMODINAMICA

0. Introducción

Como ya se ha expuesto en capítulos anteriores, para la concepción estructural las teorias científicas no son meros conjuntos de enunciados o leyes. En una primera aproximación, las teorías se podrían interpretar como lo que hemos llamado «elementos teóricos». Pero también se ha dicho ya que la elucidación que parece más apropiada para nuestras intuiciones presistemáticas acerca de lo que son las teorías es en términos de redes teóricas enteras, aunque a veces las redes teóricas pueden ser muy pequeñas o hasta «degeneradas». En este libro se han examinado ya tres ejemplos de teorías en este sentido: la teoría de la mecánica clásica de partículas, reconstruida como enorme red en el capítulo 2.5; la teoría de la termodinámica simple reconstruida como red en el capítulo 2.6 y la teoría planetaria de Kepler tratada en el capítulo anterior como «red» de un solo elemento. En lo sucesivo, cuando hablemos de teorías, nos referiremos siempre a redes teóricas.

Ahora bien, no siempre el uso popular del término «teoría» corresponde a redes teóricas y menos aún a elementos teóricos. A veces se habla de la *mecánica clásica* (y no de la mecánica clásica *de partículas*) y de la *termodinámica* (y no de la termodinámica *simple*) sin más como de «teorías». Con ello se apela a la intuición de que en el interior de una disciplina, la física por ejemplo, hay unidades intermedias situadas entre el nivel de cada una de las teorías (en nuestro sentido: como redes teóricas) y el de la disciplina misma en su

totalidad; unidades, cada una de las cuales engloba *varias* teorías que poseen entre sí ciertas «semejanzas de familia». Los físicos usan ambiguamente el término «teoría» tanto para nuestras redes teóricas como para entidades conceptuales que forman un grupo o familia de redes teóricas «semejantes». Pero esta ambigüedad en el lenguaje popular de los físicos no debe confundirnos. Se trata de dos niveles lógicamente distintos: el de las redes teóricas y el de clases (aunque no son meras clases) de redes teóricas. Para estas unidades de nivel superior no utilizaremos el rótulo «teorías», sino «marcos teóricos». Así, por ejemplo, no diremos que la mecánica clásica es una teoría, sino que es un marco teórico que engloba diversas teorías: la mecánica clásica de partículas, la mecánica del sólido rígido, la hidrodinámica, etc. Lo mismo se aplica a la mecánica cuántica o a la relativista. Y también hay que considerar lo que normalmente se llama «*la* termodinámica (reversible)» como un marco para varias teorías distintas, entre ellas la termodinámica de sistemas simples que hemos analizado en 2.6.

Naturalmente, la cuestión que se plantea en seguida es la de si esta intuición acerca de la existencia de marcos teóricos como unidades de nivel intermedio entre las teorías propiamente dichas y las disciplinas nos puede servir de base para elucidar un concepto metateórico que sea formalmente preciso y metodológicamente fructífero. Este capítulo intenta dar una respuesta positiva a esta cuestión, y para ello, siguiendo la estrategia general de este libro, consideraremos con detalle un ejemplo concreto. Nuestro punto de partida será la aceptación intuitiva de la existencia de algo así como «el marco teórico de la termodinámica (reversible)», y el resultado final será una definición precisa de esta noción y la comprobación de su relevancia metateórica. Con ello se habrá proporcionado un instrumento más de análisis a la metateoría estructural, un elemento que tanto esta metateoría como otras alternativas habían descuidado por completo hasta la fecha.

Aparte de esta finalidad básica, existe otro motivo para tratar el caso del marco termodinámico del modo como lo haremos aquí. En el transcurso del proceso de identificación del marco en cuestión comprobaremos la relevancia de análisis operacionalistas anteriores para la actual filosofía de la ciencia. Como es fácil que a este respecto surjan malentendidos sobre la posición «ideológica» del autor, conviene aclarar este punto desde el principio.

El presente análisis del caso de la termodinámica pretende apoyar una visión general de la relación existente entre las operaciones cualitativas y los conceptos métricos «abstractos» en las ciencias empíricas. Y esto, a su vez, puede servir para evaluar el «operacionalismo», tanto en sus versiones clásicas, como en las más recientes. No es un hecho casual que, históricamente, muchos «operacionalistas», desde Mach hasta Giles, pasando por Bridgman, han mostrado un

profundo interés en los fundamentos de la termodinámica y han tratado de basar su concepción filosófica general en sus análisis de esta disciplina. Las presentaciones usuales de la termodinámica en los libros de texto, incluso al nivel más teórico, se ocupan más explícitamente del papel que juegan las operaciones de laboratorio para aclarar la teoría de lo que es el caso en otras disciplinas físicas. Creo que ésta es la razón principal por la que la termodinámica ha atraído a tantos filósofos con una tendencia operacionalista y ha repelido o ha sido ignorada por filósofos con una concepción anti-operacionalista.

El análisis de la termodinámica que presentamos aquí, sin embargo, no es ni «operacionalista» ni «anti-operacionalista» –al menos en la medida en que estos rótulos tengan un significado razonablemente claro. Ciertamente, los operacionalistas tuvieron razón al señalar una característica importante de las ciencias empíricas, a saber, la relevancia de las operaciones cualitativas de laboratorio para un uso adecuado de los conceptos científicos; no obstante, el presente análisis no aspira en absoluto a justificar la aseveración mucho más fuerte de los operacionalistas en el sentido de que el significado empírico de los conceptos científicos se puede reducir completamente a operaciones cualitativas, ni siquiera a apoyar la forma más liberal del operacionalismo reciente, según la cual los conceptos teóricos métricos están siempre estrechamente ligados a las nociones operacionales cualitativas por medio de los llamados «teoremas de representación». (Este es el enfoque de Suppes y su grupo.)

Antes de entrar propiamente en la reconstrucción del marco teórico de la termodinámica, conviene hacer dos observaciones metodológicas generales sobre la forma y el alcance del procedimiento seguido en dicha reconstrucción, observaciones que podrían aplicarse igualmente a casos análogos.

La primera es la siguiente. El método axiomático de definir predicados conjuntistas se ha usado hasta ahora para sistematizar y aclarar la estructura de *teorías* singulares. Ahora bien, no veo ninguna razón para restringir su aplicación a este caso. También se puede usar para elucidar estructuras *metateóricas*. Para ver este punto más claramente vamos a introducir algunas distinciones metodológicas.

Como ya explicamos en el capítulo 1.2, la actividad que llamamos «teorización» ocurre a distintos niveles. Vamos a desarrollar un poco más esta idea. En primer lugar está la teorización directa de los fenómenos que conduce a la formulación de conceptos y leyes empíricas, que a su vez permiten la explicación y predicción de esos mismos fenómenos. Llamaremos a este estadio de la actividad teorizadora «Nivel Conceptual 1». (El estadio que consiste en reunir datos experimentales y observacionales podría denominarse «Nivel Conceptual 0».) Los conceptos y leyes del Nivel Conceptual 1 están

conectados flojamente entre sí y se suelen expresar en términos bastante vagos (incluso si aparecen matematizados). El científico profesional trabaja normalmente a este nivel.

Algunas personas –principalmente lógicos, matemáticos y científicos con tendencias filosóficas– a veces se sienten insatisfechas con el modo como la ciencia aparece al Nivel Conceptual 1 y tratan de contribuir con un mayor grado de claridad y orden en la disciplina de que se trate. Los objetos de estudio de esa gente no son entonces los fenómenos mismos, sino más bien los conceptos y las leyes que se refieren a esos fenómenos. Vamos a decir que esta clase de estudio ocurre al Nivel Conceptual 2. Una tarea típica de este nivel es la de axiomatizar una teoría empírica; esto se puede hacer introduciendo predicados conjuntistas. Ejemplos de ello los hemos visto en los capítulos 2.5, 2.6 y 2.9.

Sin embargo, nuestro impulso teorizador no tiene por qué detenerse al Nivel Conceptual 2. Podemos tratar de desarrollar conceptos y teorías generales que den cuenta de los conceptos empíricos y de las teorías empíricas que ya se han sistematizado al Nivel 2, especialmente por medio del método axiomático. Nuestros objetos de estudio serán entonces grupos enteros de teorías ya axiomatizadas. Esto es una tarea típicamente filosófica. A este estadio lo llamaremos «Nivel Conceptual 3». Los procedimientos metodológicos usados en el Nivel 3 pueden ser tan rigurosos y exactos como los del Nivel 2. En particular, también podemos utilizar el método de axiomatizar mediante un predicado conjuntista para aclarar ciertas nociones que pertenecen al Nivel 3. Ciertamente, las entidades que satisfagan los predicados conjuntistas del Nivel 3 no serán modelos empíricos; se tratará más bien de teorías enteras o entidades de tipo similar. Pero las pautas de procedimiento serán las mismas. Con otras palabras, el modo como axiomatizamos metateorías filosóficas será similar al modo como axiomatizamos teorías empíricas.

La tarea emprendida en este capítulo corresponde al Nivel Conceptual 3. El objetivo es desarrollar una metateoría de las teorías termodinámicas (que eventualmente puede servir también de pauta para metateorías similares en otras ramas de la ciencia); el método empleado a este fin es la axiomatización conjuntista.

La otra observación metodológica general que queríamos hacer se refiere al alcance de la tarea emprendida. La metateoría de las teorías termodinámicas que se propondrá a continuación no trata de todos los aspectos de una teoría termodinámica que uno debería considerar para obtener una visión completa de este tipo de teorías. Sólo se tomarán en cuenta aquellas características y estructuras que parecen ser elucidables de manera razonablemente rigurosa en el presente estadio de la filosofía de la ciencia en general. En especial, aquí no trataremos de todas las cuestiones que corresponden a lo que podríamos llamar una «semántica física» de las teorías termodinámi-

cas[47]. La principal razón de esta negligencia es que hay todavía muchas cuestiones de semántica física que no pueden ser reconstruidas con una precisión mínimamente satisfactoria. En consecuencia, aquí no se pretende ofrecer una metateoría completa de la termodinámica, aunque ciertamente el propósito es reconstruir al menos algunos aspectos esenciales comunes a todas las teorías termodinámicas. Esto significa que los predicados metateóricos introducidos en este capítulo están, en cierto sentido, incompletamente determinados, es decir, determinados sólo por algunas condiciones necesarias (esenciales). Por ejemplo, el predicado «es una teoría termodinámica» debería leerse realmente «es una *posible* teoría termodinámica». Lo mismo vale para los demás predicados metateóricos centrales que se introducen. O bien, alternativamente, uno podría decir que tales predicados están completamente determinados, pero sólo en la medida en que se refieren a las características formalmente describibles de las teorías termodinámicas. Hay que tener en cuenta esta advertencia al considerar las definiciones formales que se presentan a continuación.

1. El concepto de una teoría termodinámica

Las entidades que encajan en el marco teórico de la termodinámica (reversible), al cual en lo sucesivo abreviaremos por «M.T.T.», serán teorías de una clase particular. A ellas las llamaremos «teorías termodinámicas». Así, pues, el primer paso que hay que dar es definir el predicado conjuntista, «es una teoría termodinámica», que abreviaremos por «*TT*». Este será el predicado que ellas deben satisfacer. Esto es una condición necesaria, aunque no suficiente, para que una teoría se incluya en el M.T.T. Se requiere algo más para completar el marco: la relación de tales teorías con ciertas estructuras operacionales. A este segundo punto volveremos en las secciones subsiguientes. De momento vamos a definir el predicado «*TT*».

En el capítulo 2.6 ya hemos reconstruido estructuralmente un ejemplo de teoría termodinámica: la termodinámica de los sistemas simples. Se trata de la teoría termodinámica completa más simple que podamos imaginar. Aparte de la energía interna y la entropía, las únicas funciones primitivas consideradas en esa teoría son volumen, presión y número de moles. Las demás teorías termodinámicas son más complejas. Incluyen otros parámetros tales como la deformación elástica, la carga eléctrica o el momento magnético. No obstante, es plausible partir de la termodinámica simple como paradigma para las demás teorías, más complejas; es un paradigma en el sentido de que la estructura esencial al marco será la misma en todas partes. Si esta idea

[47] La idea de una «semántica física» y sus problemas se discuten con cierto detalle en el artículo conjunto de Moulines y Sneed de 1978 sobre Suppes.

de partida es correcta (y admitimos que lo es), entonces podemos abstraer los rasgos esenciales de la termodinámica de sistemas simples como base para definir el predicado general *TT*.

Los rasgos esenciales de la termodinámica simple que resultaron de la reconstrucción lógica emprendida en el capítulo 2.6 y que son presumiblemente relevantes para las demás teorías del marco termodinámico son los siguientes:

(a) El dominio de cualquier modelo de la teoría es un conjunto de estados isomorfo a una sucesión continua de números reales, de tal modo que se pueden definir funciones diferenciables reales sobre dicho conjunto.

(b) Un subconjunto particular del conjunto mencionado en (a) es el conjunto de todos los llamados «estados en equilibrio».

(c) Cualquier modelo de la teoría es una estructura constituida por ciertas funciones métricas, entre ellas, energía y entropía, junto con el dominio de estados y el subconjunto de estados en equilibrio.

(d) Los únicos conceptos termodinámicos-teóricos son los de energía, entropía y la propiedad monádica «es un estado en equilibrio». Todas las demás funciones provienen de teorías no-termodinámicas (presupuestas), por ejemplo, las incluidas en la geometría física o la mecánica clásica.

(e) La ley fundamental de la teoría es una relación funcional general entre la entropía, la energía y otras funciones métricas (en el caso de la termodinámica de sistemas simples, estas funciones adicionales son: volumen y número de moles). Este principio-guía sólo es válido para los estados en equilibrio y se expresa por medio de una ecuación del tipo:

$$S(z)=f^S(U(z), \ldots),$$

o bien, alternativamente,

$$U(z)=f^U(S(z), \ldots),$$

donde f^S, respectivamente f^U, son funcionales no especificados que satisfacen ciertas propiedades formales y que determinan la relación entre S, respectivamente U, por un lado, y las demás funciones, cuyos valores constituyen el dominio de f^S, respectivamente f^U[48]. Vamos a llamar a todas las funciones que quedan interrelacionadas por el principio-guía de f^S «*parámetros fundamentales*». Más adelante, cuando pasemos a discutir el marco teórico, veremos que todos estos

[48] Dado que la formulación del principio-guía por medio de f^S es equivalente a su formulación por medio de f^U, *en lo sucesivo sólo hablaremos de* f^S («representación entrópica»).

parámetros fundamentales son extensivos, es decir, tienen la propiedad de aditividad simple. Pero de este punto no nos vamos a ocupar de momento.

(f) Diferentes clases de modelos (y, por tanto, de aplicaciones) de la teoría resultan de diferentes modos de especificar f^s, que expresan otras tantas «leyes especiales» de la teoría.

(g) Sabemos por capítulos anteriores que una reconstrucción completa de la teoría presupone la introducción de condiciones de ligadura. También sabemos que hay dos clases de ligaduras: las generales, relevantes para todos los modelos de la teoría, y las especiales, asociadas a cada una de las especializaciones de la teoría. Aquí sólo estamos interesados en las ligaduras generales. En el capítulo 2.6 se postularon ocho condiciones generales de ligadura para las nociones de equilibrio, energía y entropía. (Las demás funciones se suponen sometidas a condiciones de ligadura de teorías subyacentes.) Ahora bien, dentro de este grupo de ligaduras generales se puede establecer una distinción adicional que es esencial para la reconstrucción presente: en algunos casos la formulación de las ligaduras sólo requiere de conceptos que aparecen en las estructuras que definen los modelos de la teoría; pero en otros casos es preciso hacer referencia a ciertas operaciones cualitativas, que no son componentes de los modelos mismos: un ejemplo bien conocido ya es la noción de concatenación de estados. Esta distinción no se había notado en aplicaciones anteriores del enfoque estructural, dado que en este enfoque no parecía haber lugar para nociones operacionales. Aquí, sin embargo, podemos y debemos subrayar la distinción. Diremos que la primera clase de ligaduras pertenece a la noción de teoría termodinámica (entidad en la que no aparecen involucrados conceptos operacionales). La segunda clase de ligaduras, por el contrario, son parte de las condiciones que determinan la estructura más general que denominamos «marco teórico termodinámico». Las únicas ligaduras de la primera clase relevantes para la termodinámica son las llamadas «ligaduras de identidad». Las ligaduras bastante más interesantes y complicadas que hacen referencia a nociones tales como concatenación, equilibrio parcial y transición pertenecen a la segunda clase.

Una de las principales hipótesis metateóricas del presente capítulo es que las características (a)-(g) tienen la misma forma para todas las teorías termodinámicas. Esta hipótesis implica, entre otras cosas, que los conceptos de equilibrio, entropía y energía son comunes a todas ellas, ésta es la razón por la que suponemos que las ligaduras asociadas a esos conceptos son las mismas a través de todas las teorías termodinámicas), y que cualquier teoría de este tipo incluye, además de entropía y energía, *al menos un* parámetro fundamental. Este parámetro adicional siempre será termodinámico-no-teórico (por la característica (d)). Por ejemplo, el estudio termodinámico de

sistemas electrostáticos y magnéticos requiere que se añada como parámetro fundamental la carga eléctrica y el flujo magnético, respectivamente; para estudiar termodinámicamente cuerpos sólidos introducimos el tensor de deformación elástica; los sistemas en movimiento requieren de las nociones de momento y momento angular como parámetros adicionales[49]. Todos estos parámetros provienen de teorías que no presuponen la termodinámica: la electrostática, la magnetostática o la mecánica del sólido rígido, en los ejemplos mencionados.

Los modelos de una teoría termodinámica vienen determinados por funciones S y U específicas y por otros parámetros fundamentales que se usan para describir los sistemas termodinámicos (=conjuntos de estados) de los que se ocupa la teoría (sistemas simples, sistemas eléctricos, sólidos, etc.). Con otras palabras, estos parámetros, incluyendo naturalmente a S y U, son conceptos primitivos en la axiomatización de la teoría. Sin embargo, ellos no son necesariamente las *únicas* funciones primitivas (en el sentido de no-definidas). El término «fundamental» aplicado a esos parámetros no se usa aquí en el sentido de «no-definido». Ahora bien, aparte de los parámetros fundamentales antes enumerados, hay otras funciones que se usan en las teorías termodinámicas. Ejemplos de tales funciones adicionales son: temperatura, presión, potencial químico y potencial electrostático. Estas funciones están relacionadas con pares de parámetros fundamentales por medio de derivadas parciales. Vamos a llamarlas, por esta razón, «*funciones derivadas*». Por ejemplo, sabemos que la temperatura es la derivada parcial de la energía con respecto a la entropía; la presión (negativa) es la derivada parcial de la energía con respecto al volumen, etc. Algunas de estas correlaciones (como en el caso de la temperatura y el potencial químico) pueden considerarse como *definiciones* en sentido estricto, de modo que los conceptos derivados introducidos (temperatura, potencial químico) también serán «derivados» en el sentido axiomático de no ser primitivos. Pero esto no tiene por qué ser el caso de todos los parámetros derivados de la teoría. En otros casos (por ejemplo, el caso de la presión, como ya se ha argüido en el capítulo 2.6), la correlación entre el parámetro derivado y el par correspondiente de parámetros fundamentales debería tomarse como un enunciado de hecho, es decir, como una ley empírica que correlaciona tres conceptos primitivos distintos. En consecuencia, *al menos en algunas* teorías termodinámicas, algunas de las funciones derivadas serán conceptos primitivos, de modo que los modelos de la teoría incluirán en la estructura que los define ciertos parámetros no-fundamentales, aparte de los fundamentales. En lo que sigue se entenderá por parámetro derivado específicamente una

[49] Cf. Giles, *Mathematical Foundations of Thermodynamics*, Cap. 13 y 14, y Callen, *Thermodynamics*, Cap. 13.

función métrica de estado que está relacionada (ya sea definicional o empíricamente) con un par de funciones fundamentales a través de derivadas parciales.

Otro concepto primitivo que puede introducirse para describir los modelos de una teoría termodinámica es un conjunto de índices numéricos. Algunas funciones termodinámicas se pueden descomponer en varios «componentes». Por ejemplo, la función número de moles de la termodinámica simple puede descomponerse en diversas funciones «parciales» para los números de moles de cada una de las sustancias químicas que componen el sistema total. Lo mismo se aplica a los componentes del tensor de tensión elástica en el caso de sistemas termodinámicos con deformaciones elásticas[50]. En todos estos casos, como ya hemos visto en el capítulo 2.6 para el número de moles, es más adecuado, en vez de descomponer formalmente la función primitiva (número de moles, deformación elástica), introducir un conjunto de índices numéricos, cada uno de los cuales se referirá a un componente particular.

Tomando en cuenta las observaciones anteriores, vamos a definir el concepto de teoría termodinámica por medio de un predicado conjuntista (metateórico). En lo sucesivo, θ será una variable para «teoría».

*D*1: θ es una *TT* si y sólo si para cada $x \in M[\theta]$ existen tres conjuntos Z, Z^e, I y funciones S, U, y un conjunto de funciones $\{f^1, ..., f^n\}$ con $n \geqslant 1$ y un conjunto (posiblemente vacío) $\{g^1, ..., g^m\}$ de funciones adicionales tales que:

(1) $x = \langle Z, Z^e, I, S, U, f^1, ..., f^n, g^1, ..., g^m \rangle$.

(2) Z es una variedad diferenciable acotada unidimensional.

(3) Z^e es un subconjunto finito no-vacío de $\mathscr{P}(Z)$ y, para todo $e \in Z^e$, e es un abierto no-vacío.

(4) I es un subconjunto finito no-vacío de los números naturales.

(5) S, U y todas las f^i son funciones métricas de Z en $\mathbb{R}^m$, donde $m \geqslant 1$. Por lo menos S y U son dos veces diferenciables.

(6) $\{g^1, ..., g^m\}$ es un conjunto (posiblemente vacío) de funciones derivadas con respecto al conjunto de funciones $\{S, U, f^1, ..., f^n\}$.

(7) Existe una función de $\mathbb{R}^{n+|I|}$ en $\mathbb{R}$ tal que:

(a) f^S es al menos dos veces diferenciable.

(b) f^S es una función estrictamente monótona con respecto a cada uno de sus argumentos.

[50] Cf. Callen, *Thermodynamics*, Cap. 13.

(c) S es débilmente monótona con respecto a la composición funcional de f^S con U y f^1, ..., f^n.

(d) $\forall z(z \in \cup Z^e \leftrightarrow S(z) = f^S(U(z), f^1(z), ..., f^n(z)))$.

(8) El número n es el mismo para todo $x \in M[\theta]$.

(9) Sólo Z^e, S y U son predicados θ-teóricos.

(10) $\forall x, x' \in M[\theta] \forall z \in Z_x (z \in \cup Z^e_x \wedge z \in Z_{x'} \rightarrow z \in \cup Z^e_{x'})$.

(11) $\forall x, x' \in M[\theta] \forall z \in Z_x \cap Z_{x'} : U_x(z) = U_{x'}(z)$.

(12) $\forall x, x' \in M[\theta] \forall z \in Z_x \cap Z_{x'} : S_x(z) = S_{x'}(z)$.

(13) Distintas clases de aplicaciones (modelos) de θ resultan de diversas especificaciones de f^S (diversas leyes especiales de θ)[51].

Sobre la energía y la entropía hay que hacer notar aquí lo siguiente. Hemos dicho que estas dos funciones son comunes a todas las teorías termodinámicas, pero la definición que acabamos de ofrecer de *TT* no incluye una caracterización muy estricta de energía y entropía. Las condiciones que se requieren aquí para ambas funciones son demasiado generales y débiles para que las podamos distinguir de manera precisa del resto de las funciones que intervienen en una *TT*. Por lo tanto, se podría objetar que la aseveración de que la entropía y la energía son comunes a todas las teorías termodinámicas tiene escaso contenido. Esto es, en cierta medida, correcto para esta fase de la reconstrucción del marco teórico. *D*1 no nos dice gran cosa sobre lo que podríamos llamar la semántica física de entropía y energía aparte de ciertas propiedades formales generales y la aserción de que son θ-teóricas en cada $\theta \in TT$. Pero no hay que olvidar dos puntos sobre esta cuestión. Primero, ya se ha dicho que aquí no pretendemos construir una semántica completa de la termodinámica, y esto es particularmente válido para los conceptos de energía y entropía; pero incluso prescindiendo de esta salvedad la afirmación de que la energía y la entropía son comunes a todas las teorías termodinámicas no carece de contenido: esta afirmación resultará crucial para caracterizar formalmente el marco teórico de la termodinámica, aun cuando debamos aguardar investigaciones futuras para proporcionar una semántica completa de energía y entropía. En segundo lugar, en la última sección de este capítulo se hará el intento de proporcionar una caracterización más estricta de *parte* de la semántica de entropía y energía, al conectar estas funciones métricas con algunas operaciones empíricas especiales. Así, pues, la caracterización de entropía y energía que se dará en este capítulo no resultará ser, después de todo, tan vaga como podría sospecharse a primera vista.

---

[51] El sentido y la justificación de estas condiciones de definición debería quedar claro recordando las definiciones más específicas, pero análogas, de los predicados conjuntistas del Cap. 2.6.

En el capítulo 2.6, § 2, p. 150, introdujimos el símbolo especial $\mathfrak{Z}$ para referirnos a la clase de todos los estados de todos los sistemas de que se ocupa la termodinámica de sistemas simples. Podemos generalizar ahora esta notación para el caso de cualquier teoría termodinámica θ: $\mathfrak{Z}[\theta]$ será el conjunto de estados de todos los sistemas tratados por θ.

2. Bases operacionales para la termodinámica

Las bases operacionales son esenciales para la tarea de caracterizar el marco teórico de la termodinámica. Una teoría dada no se considerará como perteneciente al marco teórico termodinámico a menos que vaya acompañada de un conjunto de bases operacionales correspondientes a cada una de las funciones primitivas de la teoría. Las bases operacionales determinan el significado operacional de una teoría, esto es, determinan el modo como la teoría va a ser aplicada operacionalmente. Después de introducir las bases operacionales correspondientes sabremos, para cada función, qué clase de operaciones de laboratorio van asociadas a ella.

El párrafo anterior, y en general la introducción de bases operacionales, puede sonar a operacionalismo. Como no es mi intención defender una filosofía operacionalista de la ciencia, es preciso que haga algunas observaciones sobre este punto. En primer lugar, la introducción de la presente noción de una base operacional no significa ningún compromiso con la aceptación del operacionalismo como doctrina filosófica. Es cierto que la mayoría de autores que han estudiado las propiedades de lo que llamo bases operacionales (Falk, Jung, Giles, Duistermaat, Hornix) adoptaron un punto de vista operacionalista, al menos tendencialmente. Pero esto no es mi problema. Se pueden adoptar algunas ideas y resultados de determinados autores sin necesidad de comprometerse con la totalidad de su concepción filosófica.

En segundo lugar, decir que las bases operacionales proporcionan el significado operacional de los conceptos de una teoría no implica que *todo* el significado de esos conceptos sea elucidable en términos operacionales, ni tampoco que el significado operacional es lo único que realmente importa. La semántica de los conceptos físicos es una cuestión muy complicada, que está lejos de haber sido resuelta. Por lo que alcanzo a ver, en ella están involucradas muchas más cosas que meras operaciones de laboratorio. No obstante, en cualquier caso, *también* las operaciones de laboratorio intervienen en la semántica física. Las entidades operacionales son un componente importante en este contexto. En lo que sigue trataremos de mostrar por qué esto es así.

En tercer lugar, decir que las funciones métricas primitivas de la

teoría van acompañadas de bases operacionales no significa que estas últimas las *definen*, ni tampoco significa que las bases operacionales son epistémicamente primarias a los conceptos métricos correspondientes. No se presupone aquí que las bases operacionales son «puramente observacionales»; también están «cargadas de teoría» en el sentido de que su introducción va de la mano de la introducción de los conceptos métricos primitivos de la teoría. Y su significado varía de teoría a teoría.

Después de estas observaciones debería estar claro que el presente enfoque no es operacionalista (al menos no en el sentido usual del término «operacionalismo»), si bien tomo prestadas algunas ideas de un operacionalista tal como Giles. Todo lo que sostengo aquí es que ciertas operaciones de laboratorio tienen algo que ver con el uso empírico de ciertos conceptos de la teoría; y dado que sabemos, al menos desde Wittgenstein, que el uso de un término tiene algo que ver con el significado del mismo, podemos concluir que las operaciones de laboratorio correspondientes a ciertas funciones métricas tienen algo que ver con su significado. Si a esta posición todavía se la quiere considerar como operacionalismo, se trata de un tipo de operacionalismo bastante aguado. Y, a fin de cuentas, tales rótulos filosóficos son lo de menos.

Lo que se puede decir positivamente sobre las bases operacionales es que los conceptos que ellas contienen son no-métricos y cualitativos; están relacionados con situaciones describibles cualitativamente, que se pueden encontrar o producir en el laboratorio. Se refieren a cosas tales como distintas clases de estados de la materia cualitativamente identificables o transiciones de un estado a otro.

Una base operacional es una estructura compuesta de un conjunto de estados y ciertas relaciones definidas sobre este conjunto. En el presente caso de la termodinámica estas relaciones son: dos relaciones diádicas, llamadas, respectivamente, «transición (bajo aislamiento)» y «equilibrio», y una operación (es decir, una relación triádica) llamada «concatenación». Bases operacionales diferentes contendrán, en general, diferentes clases de transición, equilibrio y concatenación; es decir, el conjunto de axiomas que definen implícitamente tales relaciones admite muchos modelos empíricos diferentes. Intuitivamente, esto significa que las diferentes bases operacionales expresan diversas posibilidades de transición, concatenación y situaciones de equilibrio. Algunas de ellas serán interesantes para la termodinámica, otras no. En suma, los cuatro conceptos primitivos de las bases operacionales de la termodinámica que escogemos aquí son: *estado*, *transición*, *equilibrio* y *concatenación*. Estos son los conceptos operacionales básicos del marco termodinámico.

En este punto puede surgir la cuestión de la selección de estos conceptos. La justificación es, en términos generales, la siguiente: todo lo que puede afirmarse en termodinámica sobre los aspectos *operacio-*

nales de sus teorías puede reducirse a enunciados sobre ciertas clases de concatenaciones de estados, procesos de transición y estados de equilibrio. Esta es una hipótesis metateórica sobre el aspecto operacional del marco termodinámico, y es difícil ver cómo podría probarse o refutarse formalmente. Es una metahipótesis general sobre el modo como las teorías termodinámicas se relacionan con ciertas operaciones empíricas. La base de apoyo para esta hipótesis sólo puede consistir en presentaciones concretas de la termodinámica y en informes de laboratorio. Creo que existe tal apoyo. En cualquier caso, ésa es la hipótesis de partida para lo que sigue.

La mayoría de estudios de fundamentos de la termodinámica (los de Falk y Jung, Hornix, Giles) también toman los conceptos antes mencionados como básicos, aunque por razones posiblemente erradas (operacionalismo estricto). Giles y Duistermaat introducen explícitamente «estado», «transición» y «concatenación» como los conceptos primitivos de sus reconstrucciones. Nótese, sin embargo, que su interpretación de estos conceptos, especialmente del de transición, es mucho más restrictiva que la adoptada aquí. (De hecho, Giles sólo considera la transición adiabática, mientras que Duistermaat sólo se interesa por las transiciones adiabática y energética.) Giles intenta obtener el concepto de equilibrio termodinámico como concepto definido, pero su interpretación de este concepto es demasiado limitada para nuestros propósitos, dado que sólo se refiere al equilibrio (total) completo. Aquí, en cambio, tomaremos como concepto primitivo la noción general de equilibrio (no especificado); en ello seguimos las observaciones intuitivas de Tisza sobre este concepto (recuérdese su cita al respecto en el capítulo 2.6) y las reconstrucciones más sistemáticas de Falk y Jung y Hornix.

La presente interpretación de estos conceptos operacionales es más cercana a la de Falk y Jung que a la de Giles, si bien no coincide tampoco con la primera. En general, la caracterización presente intenta ser más amplia y más «subdeterminada» que cualquiera de las anteriores, por la simple razón de que no estamos comprometidos con el reduccionismo operacionalista. No obstante, el punto digno de mención aquí es que todos estos distintos enfoques, a pesar de sus diferencias de detalle, coinciden ampliamente en la selección de las operaciones fundamentales de la termodinámica; creo que esto es así porque todos apuntan al mismo aspecto del marco termodinámico: sus «bases operacionales».

Dejando aparte el concepto de *estado*, del cual ya hemos hablado en la parte introductoria del capítulo 2.6, vamos a hablar del significado intuitivo de las tres restantes relaciones cualitativas que constituyen una base operacional de la termodinámica.

Las *transiciones* son relaciones entre los estados. Simbolizaremos la transición de un estado z a otro z' así: $z \Rightarrow z'$. Hay muchas clases distintas de transiciones posibles. En general, el concepto de transi-

ción está vinculado al concepto de *proceso* físico. Siempre que analizamos un sistema físico en términos de estados, podemos decir que un proceso dado que tiene lugar en dicho sistema es una sucesión de transiciones de un estado a otro. Ahora bien, la termodinámica no se interesa por todas las clases de transiciones posibles. Las transiciones termodinámicamente interesantes son aquellas que pueden efectuarse bajo algún tipo de *aislamiento*. Esta es la razón por la que algunos autores (Hornix, por ejemplo) también llaman «aislamiento» al tipo de relación que aquí llamamos «transición». De hecho, el rótulo más apropiado para expresar la idea intuitiva detrás de este concepto sería: «transición-bajo-aislamiento-del-tipo-*k*», donde *k* es una variable; pero éste es un término demasiado largo, por lo que nos limitaremos a hablar de «transición (termodinámica)». Hornix ha catalogado cuatro tipos principales de transiciones termodinámicamente interesantes (o «aislamientos»): transición bajo aislamiento adiabático; transición bajo aislamiento energético; transición bajo aislamiento dimensional de algún tipo (es decir, aislando el sistema en alguna de sus dimensiones espaciales); transición bajo aislamiento químico (es decir, aislando la masa de algún compuesto químico del sistema).

Cada una de estas transiciones está asociada conceptualmente a alguna función métrica termodinámica de una manera específica. En la sección siguiente trataremos de esta conexión. De momento, nuestra tarea es analizar la estructura lógica general de tales relaciones de transición independientemente de cualquier posible conexión con funciones métricas.

Intuitivamente, el concepto de transición se suele asociar a la idea de una sucesión *temporal*. Si un «proceso» lleva de z a z', como se dice, pensamos de manera natural que z' «tiene lugar» *después* de z –esto es, pensamos que z y z' están separados por cierto lapso de tiempo. No obstante, no deberíamos suponer que esta visualización intuitiva es realmente necesaria para el formalismo de las teorías termodinámicas. En particular, *no* deberíamos presuponer que cuando hablamos de transiciones termodinámicas estamos introduciendo un parámetro métrico temporal t entre los conceptos básicos de la teoría. El tiempo (métrico) no es esencial a la termodinámica (reversible). Esto a veces se expresa diciendo que no deberíamos hablar de «termodinámica», sino de «termostática», si bien en el capítulo 2.6 ya dimos razones por las que el término «termostática» tampoco es del todo apropiado.

El concepto presente de transición es muy amplio: incluye no sólo transiciones «naturales» o «espontáneas» (como las consideradas por Giles y Duistermaat), sino también cualquier clase de cambios sucesivos «manipulados» o «forzados» en el sistema. Normalmente, las transiciones se dan a consecuencia de algún tipo de *interacción* del sistema con su «entorno». La noción intuitiva de interacción es el vínculo conceptual entre las transiciones y otro concepto operacional

básico: las concatenaciones de estados. Las concatenaciones interesantes son las que normalmente (aunque no siempre) producen transiciones.

Ya se dio una primera caracterización intuitiva de la concatenación termodinámica en el capítulo 2.6 al introducir las condiciones de ligadura. Ahora vamos a extendernos más sobre sus propiedades. Formalmente, una concatenación es una operación en el sentido algebraico sobre el conjunto de estados. Si z es el estado que (unívocamente) resulta de la concatenación de los estados z' y z'', escribiremos: $z = z' \circ z''$. Es sabido que no sólo en la termodinámica, sino en prácticamente todas las ramas de la ciencia, aparecen operaciones de concatenación. La termodinámica sólo está interesada en ciertos tipos de concatenaciones: nuevamente aquellas que están relacionadas con parámetros métricos termodinámicos. Ejemplos de tales operaciones son: la concatenación de dos estados cuyas realizaciones materiales correspondientes están separadas por una pared móvil, o bien una mezcla de gases, etc.

Ante un examen cuidadoso de la literatura sobre fundamentos de la termodinámica comprobamos que no hay acuerdo general entre los autores acerca de la interpretación intuitivamente adecuada de la concatenación termodinámica. En un extremo encontramos a Giles considerándola como una cuestión puramente conceptual:

«Hay que subrayar que el proceso de formar la unión [=concatenación] $A \circ B$ es un proceso puramente conceptual; en particular no implica ninguna interacción entre A y B (Giles, *op. cit.*, p. 22, nota).

En el otro extremo observa Duistermaat:

«El acto de unir [=concatenar] estados de sistemas físicos y la observación del comportamiento del sistema acoplado es tan común en la física experimental que se puede esperar una estructura de adición [=concatenación] en casi cualquier teoría física. Si bien en muchos casos no hay peligro de confusión, a veces se necesita ser más cuidadoso en la interpretación física del concepto de adición. En particular, la adición de estados no se considera como un mero "acto mental", sino como una operación física real.

Por ejemplo, la descripción del estado 'dos partículas cargadas juntas en el espacio' no debe considerarse como la mera adición de las descripciones de ambas partículas (aisladas). La descripción de este estado es incompleta si no hace mención del equipo auxiliar y del esfuerzo requerido para acercar las dos partículas cargadas» (Duistermaat, «Energy and Entropy as Real Morphisms for Addition and Order», pp. 330-331).

En cuanto a Falk y Jung parece que adoptan una posición intermedia:

«Por lo que respecta al concepto de combinación [=concatenación]... en muchos casos es apropiada la imagen de juntar sistemas e introducir una interacción a través de una "pared" adecuada (Falk y Jung, *op. cit.*, pp. 159-160).

Esta última caracterización parece que concuerde mejor con la de Duistermaat que con la de Giles. Nótese, sin embargo, el matiz poco comprometido introducido por la frase «en muchos casos». Parece indicar que, si bien la interacción física es la situación normal cuando hablamos de concatenación, la unión «puramente conceptual» de estados también podría incluirse bajo el concepto general de concatenación.

En cualquier caso, ésta será la interpretación intuitiva de la operación «$\circ$» que se propone aquí: cualquier tipo de unión de estados será subsumido bajo el rótulo «concatenación». No importa si esta unión se interpreta como un acto «puramente conceptual» o bien como una manipulación física, siempre y cuando sea termodinámicamente relevante, es decir, siempre y cuando la interpretación particular de «$\circ$» admitida sea relevante para la determinación de alguna función termodinámica particular en la que estamos interesados. Cuando pasemos a discutir las propiedades formales del marco teórico termodinámico veremos más claramente lo que significa que una operación «$\circ$» sea relevante para una función termodinámica.

La única restricción general que introducimos aquí sobre el concepto «$\circ$» es que sea una operación algebráicamente *cerrada*. Esto significa que si dos estados cualesquiera z, z', se concatenan dentro del contexto de una teoría θ, entonces $z \circ z'$ (sea cual sea la interpretación particular de «$\circ$») también será un estado en esta teoría. Dicho formalmente, si z, $z' \in \mathfrak{Z}[\theta]$, entonces también $z \circ z' \in \mathfrak{Z}[\theta]$. Si una concatenación lógicamente posible de estados no satisface esta condición (un ejemplo extremo: explosiones instantáneas debidas a la combinación de gases), entonces la excluiremos del dominio de posibles interpretaciones de «$\circ$».

El *equilibrio* termodinámico es una relación diádica entre estados. Escribiremos «$z \sim z'$» para «z está en equilibrio con z'». Esta relación no es reflexiva, dado que algunos estados son estados de inequilibrio (o sea, no están en equilibrio consigo mismos). Pero tampoco es irreflexiva, pues algunos estados sin duda están en equilibrio consigo mismos: son estados de equilibrio.

También aquí, como en el caso de las transiciones y concatenaciones, hay muchas clases diversas de relaciones de equilibrio termodinámicamente relevantes: equilibrio térmico, químico y de presión pueden citarse entre las formas más usuales. Esto significa que interpretamos esta relación como de equilibrio *parcial*: equilibrio con respecto a alguna propiedad termodinámica específica. Por ejemplo, dos estados pueden hallarse en equilibrio mutuo químico, pero no térmico. El tipo específico de equilibrio que se dé entre dos estados diferentes depende del tipo específico de «paredes» que los separen –es decir, del tipo específico de interacción de laboratorio que actúa en su concatenación. Así, pues, también existe una relación operacional entre los conceptos de equilibrio y concatenación.

El concepto de equilibrio completo, distinto al concepto de equilibrio solamente parcial, es una noción que introduce cada teoría termodinámica para formular sus leyes fundamentales adecuadamente. Ya sabemos que se trata de un concepto típicamente θ-teórico para cada teoría θ de la termodinámica; no es un concepto operacional. Sin embargo, existe ciertamente un vínculo que conecta (¡no *define*!) el equilibrio completo con el equilibrio parcial. En términos generales, el vínculo es éste: dada una teoría θ, admitiremos que un estado dado z en θ se halla en un estado de equilibrio (completo) si se halla en equilibrio parcial consigo mismo para todas las relaciones de equilibrio (parcial) introducidas por las bases operacionales subyacentes a θ. La formulación de esta conexión se verá más claramente en la sección siguiente.

El concepto de equilibrio (parcial) es la contrapartida operacional de lo que se ha denominado una «función derivada» en termodinámica (temperatura, presión, etc.). También juega un papel esencial en la formulación de procesos «cuasi-dinámicos» –en conexión con el principio de Gibbs y supuestos relacionados.

Después de las observaciones informales que anteceden, estamos en posición de introducir formalmente el concepto general de base operacional de la termodinámica.

D2: B es una *base operacional* (para una teoría termodinámica) si y sólo si existen $\mathscr{S}, \Rightarrow, \circ, \sim$, tales que:

(1) $B = \langle \mathscr{S}, \Rightarrow, \circ, \sim \rangle$.
(2) $\mathscr{S}$ es un conjunto no-vacío.
(3) $\Rightarrow$ es una relación diádica sobre $\mathscr{S}$, y $\Rightarrow$ es transitiva.
(4) $\circ$ es una función diádica de $\mathscr{S} \times \mathscr{S}$ en $\mathscr{S}$, y $\circ$ es no-idempotente, conmutativa y asociativa.
(5) $\sim$ es una relación diádica sobre $\mathscr{S}$ y $\sim$ es simétrica y transitiva.
(6) $\forall z, z', z'' \in \mathscr{S}: z \circ z'' \Rightarrow z' \circ z'' \rightarrow z \Rightarrow z'$.
(7) $\forall z, z', z'' \in \mathscr{S}: z \circ z' \sim z'' \leftrightarrow (z \sim z'' \wedge z' \sim z'')$.

Las condiciones *D2*-(2) a *D2*-(5) son los «axiomas estructurales» para la noción de base operacional; determinan las propiedades formales de sus cuatro componentes. A los elementos de $\mathscr{S}$ se les llama «*estados*». $\mathscr{S}$ es cualquier conjunto de entidades operacionalmente diferenciables del tipo de lo que llamaríamos «estados de un sistema». No se especifica nada sobre la cardinalidad de $\mathscr{S}$. Esta cuestión, que resulta importante cuando consideramos los modelos de las teorías termodinámicas, se deja abierta aquí. La razón es que estamos tratando de los aspectos *operacionales* de la termodinámica;

la cuestión de la cardinalidad es un asunto *teórico* (en el sentido de no-operacional).

La única propiedad formal que requerimos de la transición es que sea transitiva; esto es, si el sistema puede pasar de z a z' de cierto modo, y de z' a z'' del mismo modo, entonces debería ser posible decir que puede pasar de z a z'' de *ese* modo, sin mencionar z'. Nótese que no exigimos, pero tampoco excluimos que la transición sea reflexiva. La relación entre las transiciones y los parámetros fundamentales que se postulará en la siguiente sección permitirá comprender por qué tiene sentido intuitivamente admitir ambas posibilidades: que el sistema «pase» de un estado al mismo (cuando el parámetro fundamental correspondiente permanece constante) o bien que pase a un estado distinto (cuando el parámetro correspondiente está cambiando de magnitud).

Esta noción general de transición nos permite definir una noción general de reversibilidad relativa. Un proceso $z \Rightarrow z'$ es reversible relativamente a esta transición determinada si también es el caso que $z' \Rightarrow z$. Formalmente:

*D*3: Rever (z, z') si y sólo si: $z \Rightarrow z' \wedge z' \Rightarrow z$.

Esta es una noción de reversibilidad mucho más amplia que la noción corriente (reversibilidad *adiabática*), la cual es, naturalmente, la más interesante. De la presente idea de reversibilidad en general no se puede hacer un gran uso sistemático. Sin embargo, para expresar formalmente algunas situaciones posibles puede ser conveniente tener a mano dicha noción.

Por lo que respecta a la concatenación, el principal requerimiento es que $\langle \mathcal{S}, \circ \rangle$ sea un semigrupo conmutativo –que es lo que naturalmente esperaríamos de cualquier operación de concatenación fisicamente interesante (cf. Duistermaat, *op. cit.*, p. 331).

Postulamos que el equilibrio entre estados sea una relación simétrica y transitiva, pero no exigimos reflexividad. Con respecto a algunos estados (a saber, los estados de inequilibrio), tiene sentido decir que no están en equilibrio consigo mismos, y, en consecuencia, está claro que tampoco estarán en equilibrio con ningún otro estado. Así, pues, el equilibrio no es una relación de equivalencia en sentido estricto.

Los últimos axiomas establecen algunas conexiones generales entre los conceptos fundamentales de una base operacional. El primero conecta la transición con la concatenación; el último considera el caso en que un estado producto de la concatenación de otros dos se halla en equilibrio con un tercer estado.

De los axiomas postulados aquí se pueden derivar algunos corolarios simples e intuitivamente plausibles.

$T1$: $\forall z, z'(z \sim z' \rightarrow z \sim z \wedge z' \sim z')$.

Esto es, si dos estados se hallan en equilibrio es que son estados de equilibrio. Esto se sigue directamente de $D2$-(5).

$T2$: $\forall z, z'(z \circ z' \sim z'' \rightarrow z \sim z')$.

Esto es, si un sistema compuesto está en equilibrio con otro sistema, los componentes del compuesto también estarán en equilibrio entre sí. Esto resulta directamente de $D2$-(5) y $D2$-(7).

T 3: $\forall z, z'(z \sim z' \leftrightarrow z \circ z' \sim z \circ z')$.

Esto es, si dos estados están en equilibrio entre sí, el sistema compuesto de ambos se halla en equilibrio.

Prueba: Si $z \sim z'$, entonces, por $T1$, $z' \sim z'$; por tanto, tenemos $z \sim z' \wedge z' \sim z'$; luego, por $D2$-(7), $z \circ z' \sim z'$; y, aplicando nuevamente $T1$, obtenemos $z \circ z' \sim z \circ z'$. Recíprocamente, si $z \circ z' \sim z \circ z'$, por aplicación sucesiva de $D2$-(7), obtenemos primero $z \sim z \circ z'$, y finalmente $z \sim z'$.

$T4$: $\forall z(z \sim z \rightarrow z \circ^{n)} \ldots \circ z \sim z \circ^{n)} \ldots \circ z)$.

Esto se sigue de T 2 por iteración en el caso $z \sim z$. Este corolario puede que tenga algún interés para sistemas compuestos que se suponen construidos a partir de n réplicas del mismo sistema en un estado dado.

En resumidas cuentas, no se pueden sacar muchas conclusiones interesantes de los axiomas de una base operacional. Esto se debe al carácter muy general de los mismos. Considero que esta situación es intuitivamente apropiada; los conceptos puramente operacionales no deberían considerarse como muy potentes a menos que aparezcan conectados con conceptos métricos. El contenido esencialmente físico de la termodinámica se obtiene cuando ambos componentes, el teórico y el operacional, quedan vinculados entre sí.

3. El marco teórico de la termodinámica

Vamos a elucidar finalmente la noción de marco teórico de la termodinámica como estructura global que cubre todas las teorías termodinámicas. En este punto ya disponemos de los dos componentes básicos: por un lado, la noción de teoría termodinámica, la cual es, por así decir, el aspecto teórico del marco; por otro, la noción de base

operacional de la termodinámica, la cual representa su aspecto operacional. El marco es la estructura que resulta de la interconexión de ambos componentes.

La idea intuitiva general que se halla detrás de la vinculación que vamos a establecer es: a cada función fundamental de una teoría termodinámica se le asociará una base operacional que fija el «significado operacional» de dicha función. Así, pues, cada parámetro termodinámico fundamental resulta conectado con formas particulares de transición, concatenación y equilibrio.

Dado que, por lo que hemos visto en § 1, cada teoría termodinámica incluye al menos tres parámetros fundamentales distintos, ella estará vinculada a por lo menos tres bases operacionales.

Energía y entropía son los dos parámetros fundamentales comunes a todas las teorías termodinámicas. En consecuencia, sus correspondientes bases operacionales serán «universales» dentro del marco termodinámico. En particular, la energía y la entropía van asociadas a dos tipos bien conocidos de transiciones de laboratorio: transición bajo aislamiento *energético* y transición bajo aislamiento *adiabático* (abreviadamente: transición energética y transición adiabática). Estas dos transiciones tienen propiedades formales especiales. La transición energética siempre es reversible. La transición adiabática es condicionalmente conexa[52]. Además, la transición energética y la adiabática se hallan en cierta relación lógica: la primera es un caso especial de la segunda. Su conexión con las funciones métricas energía y entropía es la siguiente: en una transición energética la función energía permanece constante; en una transición adiabática la función entropía no puede disminuir de valor. Finalmente, se hallan en una conexión particular con cualquier tipo de equilibrio (termodinámicamente interesante). En efecto, si un estado dado no se halla en cierto tipo de equilibrio (parcial), entonces existe otro estado que puede alcanzarse partiendo del primero a través de una transición adiabática (a veces incluso a través de una transición energética) y que sí se halla en ese equilibrio. La termodinámica (reversible) no se interesa por situaciones de inequilibrio en las que no ocurra esto.

Todas las demás transiciones consideradas están asociadas con las funciones métricas correspondientes del mismo modo como lo está la transición energética con la función energía.

Los llamados parámetros intensivos que resultan de la derivación parcial de f^S con respecto a uno de sus argumentos permanecen constantes cuando se da la situación de equilibrio correspondiente.

Todos los parámetros fundamentales son extensivos con respecto a *cualquier* clase de concatenación.

---

[52] Falk y Jung fueron los primeros en descubrir este hecho acerca de la transición adiabática. El término «relación condicionalmente conexa» lo introdujeron después Roberts y Luce para el caso de la termodinámica.

La función energía satisface el «principio de conservación de la energía», es decir, siempre que la energía no permanece constante en un estado dado, esto significa que hay otro estado (posiblemente uno compuesto) que está concatenado con el primero de tal modo que la energía permanece constante para el sistema compuesto de ambos estados.

La entropía y la energía satisfacen los principios de Gibbs de máximo y mínimo, respectivamente: siempre que la energía de un sistema compuesto permanece constante en el entorno de cierta concatenación de estados, el sistema compuesto se halla en el equilibrio (parcial) correspondiente si y sólo si la entropía alcanza un máximo (relativo). Análogamente ocurre para el caso de que la energía alcance un mínimo.

Finalmente, la noción *T-teórica* de equilibrio (total) se conecta con los diversos tipos *operacionales* de equilibrio a través del postulado de que un estado se halla en equilibrio (total) cuando se halla en todas las clases operacionales de equilibrio.

Todas estas conexiones entre el componente teórico y el operacional del marco se sintetizan en la definición siguiente de «teoría termodinámica con bases operacionales» (*TTBO*).

D4: $\langle \theta, \mathscr{L} \rangle \in \widehat{TTBO}$ si y sólo si:

(1) $\theta \in \widehat{TT}$

(2) $\mathscr{L}$ es un conjunto de bases operacionales y $\|\mathscr{L}\| \geqslant 3$.

(3) $\forall B_i \in \mathscr{L}$: $\mathscr{S}_i \subseteq \mathscr{S}_\theta$.

(4) Existen dos bases operacionales distintas, B_A, $B_E \in \mathscr{L}$, tales que:

(a) $\Rightarrow_A$ es condicionalmente conexa, es decir:

$$\forall z, z', z''(z \underset{A}{\Rightarrow} z'' \wedge z' \underset{A}{\Rightarrow} z'' \rightarrow z \underset{A}{\Rightarrow} z' \vee z' \underset{A}{\Rightarrow} z).$$

(b) $\forall z, z'(z \underset{E}{\Rightarrow} z' \rightarrow z' \underset{E}{\Rightarrow} z)$.

(c) $\forall z, z'(z \underset{E}{\Rightarrow} z' \rightarrow z \underset{A}{\Rightarrow} z')$.

(d) $\forall z, z'(z \underset{A}{\Rightarrow} z' \rightarrow S(z) \leqslant S(z'))$.

(e) $\forall z, z'(z \underset{E}{\Rightarrow} z' \rightarrow U(z) = U(z'))$.

(f) $\forall z(\neg z \underset{i}{\sim} z \rightarrow \exists z_i(z \underset{A}{\Rightarrow} z_i \wedge z_i \underset{i}{\sim} z_i)$.

(5) Para cualquier parámetro fundamental f_i de θ tal que $f_i \neq S$, existe una $B_i \in \mathscr{L}$, con $B_i \neq B_A$, tal que:

(a) $\forall z, z'(z \underset{i}{\Rightarrow} z' \rightarrow f_i(z) = f_i(z'))$.

(b) $\forall z, z'(D_{f_i} f^S(z) = D_{f_i} f^S(z') \leftrightarrow z \underset{i}{\sim} z')$.

(6) Para cualquier parámetro fundamental f_i de θ y para cualquier $B_j \in \mathscr{L}$: $\forall z, z'(f_i(z \circ_j z') = f_i(z) + f_i(z'))$.

(7) $\forall z(DU(z) \neq 0 \rightarrow \exists z' \exists i(DU(z \circ_i z') = 0))$.

(8) Para todo z, z', f_i:

(a) $\Delta U(z \circ_i z') = 0 \wedge \forall f_j(f_j \neq f_i \wedge f_j \neq U \wedge f_j \neq S \rightarrow \Delta f_j(z) = \Delta f_j(z') = 0) \rightarrow (z \succsim_i z' \leftrightarrow \text{relmax}(S, z \circ_i z'))$.

(b) $\Delta S(z \circ_i z') = 0 \wedge \forall f_j(f_j \neq f_i \wedge f_j \neq U \wedge f_j \neq S \rightarrow \Delta f_j(z) = \Delta f_j(z') = 0) \rightarrow (z \succsim_i z' \leftrightarrow \text{relmin}(U, z \circ_i z'))$.

(9) $\forall z(z \in \cup Z^e \leftrightarrow \forall B_i \in \mathscr{L}(z \succsim_i z))$.

D5: θ pertenece al marco teórico de la termodinámica si y sólo si existe un $\mathscr{L}$ tal que $\langle \theta, \mathscr{L} \rangle \in TTBO$.

A partir de las condiciones enunciadas en *D*4 se podrían derivar ciertas importantes propiedades generales que son comunes a todas las teorías pertenecientes al marco termodinámico; por ejemplo, la positividad de la temperatura, así como los llamados principios de Kelvin y de Clausius. Las pruebas serían, en cada teoría, enteramente análogas a las ya proporcionadas en el caso particular de la termodinámica de sistemas simples, en el capítulo 2.6. Estas pruebas no dependen de las propiedades específicas de la teoría particular reconstruida allí; se pueden generalizar a cualquier teoría termodinámica. Con ello se ha reconstruido lo esencial del marco teórico de la termodinámica.

Para concluir este capítulo, quisiera hacer unas observaciones generales sobre la relevancia potencial de la reconstrucción de marcos teóricos de las diversas disciplinas para el estudio de los aspectos diacrónicos de la ciencia. Ciertamente, la reconstrucción del marco termodinámico que aquí se ha ofrecido tiene un carácter puramente sincrónico. Sin embargo, las categorías conceptuales que hemos identificado a través de esta reconstrucción pueden ser útiles también para un análisis diacrónico de la ciencia. La razón principal de esto es la siguiente. Si se admite la categoría de marco teórico como algo distinto de las teorías singulares, se abre una perspectiva más amplia en el estudio diacrónico de la ciencia, en particular, en el estudio de lo que se puede llamar «la evolución científica». En efecto, se puede constatar entonces que el término «evolución de las teorías científicas» es ambiguo: puede referirse ya sea a la evolución de teorías singulares dentro de un marco teórico dado, o bien a la evolución del marco teórico entero como estructura global. Con otras palabras, deberíamos considerar la evolución «normal» de la ciencia al menos a dos niveles distintos: el de teorías y el de marcos. Y el examen de la

evolución de un marco teórico debería incluir no sólo la evolución de teorías singulares y de sus interconexiones, sino también la evolución de lo que hemos llamado «aspectos operacionales». Existe la posibilidad de que la evolución del segundo aspecto ofrezca un cuadro diferente del ofrecido por la evolución de teorías. Aunque todo esto todavía pueda sonar a especulación (y no trataremos más de esta cuestión en este libro), sin embargo, es una dirección que debería ser explorada en el futuro.

3. GENESIS Y DESARROLLO DE PROGRAMAS CIENTIFICOS

FORMA Y CONTENIDO DE LAS REVOLUCIONES CIENTIFICAS: EL CASO DE LA MECANICA NEWTONIANA

A Isaac Newton le debe la historia de la ciencia tres grandes revoluciones científicas: una en mecánica y astronomía, otra en óptica y una tercera en matemáticas. Las tres son independientes entre sí. Aquí sólo vamos a analizar la primera.

La revolución mecánica de Newton puede considerarse como un paradigma de revolución científica en el sentido de Kuhn. Por ello creo que es particularmente relevante para la filosofía de la ciencia analizarla con cierto detalle.

La estrategia que seguiré en este análisis será partir de las ideas generales de Kuhn acerca de las revoluciones científicas y tratar de reconstruir lógicamente la revolución newtoniana en sus términos, lo cual no sólo permitirá esclarecer y precisar la estructura de dicha revolución, sino también los propios conceptos kuhnianos, cuya adecuación se va a poner a prueba.

Para empezar, parto de la hipótesis metacientífica de que *hay* revoluciones científicas y de que la aparición de la mecánica newtoniana es una de ellas. Quizá haya todavía filósofos e historiadores que se adscriban a una visión acumulativista y continuista de la historia de la ciencia y, por tanto, nieguen la realidad de las revoluciones científicas. Creo que Kuhn y otros han mostrado convincentemente que tal visión es históricamente inadecuada y metodológicamente perniciosa. Algunos críticos de la noción de revolución científica hacen notar que el que un determinado cambio científico sea o no revolucionario es cuestión de grado. Pero esta observación, aun siendo válida, no prueba nada. Pues del hecho de que un concepto sea gradual no se desprende que sea ficticio o vacío; la mayoría de los conceptos

científicos avanzados son graduales, y no por ello dejan de ser útiles para establecer distinciones. Tal es el caso del concepto de revolución científica.

Asumo, pues, que hay revoluciones científicas y, en particular, que una de las más profundas en la historia de la ciencia ha sido la mecánica newtoniana. De ver en qué consiste, y no de negarla, es de lo que aquí se trata.

Las revoluciones científicas, por ser la ciencia una parte de los fenómenos culturales y sociales en general, tienen sin duda antecedentes y consecuentes socioculturales. Pero no es este aspecto externo de las revoluciones científicas el que se va a discutir aquí. Aquí se trata de analizar su estructura interna. Creo que ésta se puede estudiar con bastante independencia de los factores socioculturales externos. Esto es debido a que, a diferencia de otro tipo de revoluciones, las científicas no atañen a objetos sociológicos masivos, como puedan ser las instituciones políticas, las relaciones económicas o las actitudes mentales, sino que tienen lugar al nivel «abstracto» de los conceptos técnicos empleados por pequeños grupos de profesionales. Esto permite estudiar las revoluciones científicas más como cambios en las relaciones conceptuales que como cambios en las relaciones humanas.

Naturalmente, no todos son cambios puramente conceptuales en las revoluciones científicas. Como ha señalado Kuhn, los cambios conceptuales suelen ir acompañados o incluso posibilitados por ciertos factores sociopsicológicos *internos* a la comunidad científica. Aunque no voy a entrar en los detalles de este tipo de factores, sí es conveniente recordarlos brevemente y cotejarlos con el caso particular que nos ocupa.

Desde el punto de vista sociopsicológico interno a la ciencia, cabe mencionar los siguientes aspectos esenciales:

— a) Las revoluciones científicas se producen en una época de sentimiento generalizado de crisis dentro de la comunidad científica.
— b) Las revoluciones científicas las inician un grupo reducido de investigadores, muchas veces uno solo, generalmente jóvenes.
— c) Las revoluciones científicas afloran repentinamente como cambio de *Gestalt* (*Gestalt-switch*) en la mente de algún investigador.
d) Las nuevas propuestas de los revolucionarios científicos chocan con gran resistencia por parte del resto de la comunidad científica y provocan enconadas polémicas.
e) El grupo de contrarrevolucionarios científicos no desaparece de golpe, sino que va agonizando lentamente, a lo largo de períodos a veces bastante largos.

Todos estos aspectos sociopsicológicos se encuentran perfectamente ejemplificados en el caso de la revolución mecánica newtoniana. En efecto:

a) En el último tercio del siglo XVII la astronomía entró en una crisis provocada por varios problemas profundos; entre ellos pueden mencionarse los siguientes:

1) La imposibilidad de dar cuenta de las leyes de Kepler para las órbitas planetarias dentro de los esquemas teóricos existentes.
2) La imposibilidad de hallar leyes adecuadas para las órbitas de los cometas.
3) Las leyes de Kepler no se aplicaban más que de forma muy burda a la órbita de la Luna alrededor de la Tierra.
4) No se hallaba explicación satisfactoria alguna para el fenómeno de las mareas.

b) La revolución la inició Newton solo alrededor de 1666, es decir, cuando tenía unos veinticinco años de edad. Sin embargo, no fue sino hasta más de veinte años más tarde, con la publicación de los *Principia*, cuando la revolución newtoniana trascendió a la comunidad científica. Que esta revolución fue obra de Newton solo apenas puede dudarse. Es cierto que por los años 1670 Hooke también propuso una fuerza central inversamente proporcional al cuadro de la distancia para derivar las leyes de Kepler. Sin embargo, no por ello puede considerarse que Hooke iniciara la revolución mecánica. En primer lugar, Hooke nunca elaboró matemáticamente su idea ni realmente demostró que las leyes de Kepler se derivaran de ella. Y, en segundo lugar, lo que es mucho más importante, la ley de gravitación no es el aspecto esencial de la revolución mecánica, como a veces se cree. En realidad, la revolución consistió en la propuesta de un nuevo esquema dinámico fundamental, del cual la ley de gravitación es sólo un aspecto derivado. El mérito de este nuevo esquema general le cabe exclusivamente a Newton.

c) Sobre el proceso mental que condujo al joven Newton a su nueva concepción mecánica no tenemos datos suficientes. Pero todo parece indicar que tuvo realmente lugar como cambio de *Gestalt*. La famosa anécdota de la manzana, cuya veracidad histórica parece comprobada, sería un síntoma que corroboraría esta hipótesis.

d) Fuera del estrecho grupo de sus colaboradores y amigos (Halley, Cotes, Clarke), Newton chocó con enconada resistencia y duros ataques a sus ideas, y hasta poco antes de su muerte, o incluso más tarde, no puede decirse que convenciera al grueso de la comunidad científica. No sólo los científicos continentales, con Huygens y Leibniz a la cabeza, atacaron el esquema newtoniano, sino que incluso en Inglaterra tuvo dificultades con dos de los físicos más eminentes: Hooke y Boyle.

e) Las propuestas alternativas a la mecánica newtoniana persistieron por largo tiempo después de la publicación de los *Principia*, especialmente en el continente. Puede decirse que no fue sino hasta

mediados del siglo XVIII, con Euler y D'Alembert sobre todo, que la mecánica newtoniana se impuso definitivamente.

Podemos admitir, pues, que en sus aspectos sociopsicológicos la revolución newtoniana sigue las pautas señaladas por Kuhn, al menos en sus rasgos generales. Pero no son estos aspectos los esenciales a una revolución científica, sino más bien lo que hemos llamado las relaciones conceptuales. Y a ello dirijo la atención ahora.

Dicho en términos generales, una revolución científica es un corte más o menos profundo a nivel conceptual en la evolución de una disciplina científica. Este corte separa dos períodos de lo que Kuhn llama «ciencia normal»: la revolución acaba con un determinado período de ciencia normal para dar lugar a uno nuevo.

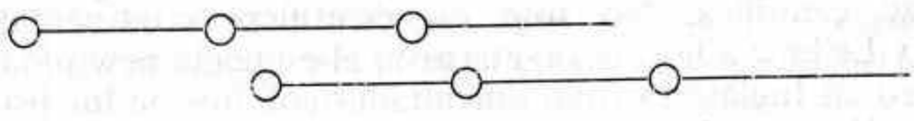
Exceptuando las llamadas revoluciones *iniciales*, que construyen una disciplina por primera vez, las demás revoluciones científicas son procesos conceptuales de *sustitución*. Una revolución científica no consiste únicamente en derruir un viejo edificio, sino además en construir uno nuevo. Obviamente, la revolución newtoniana no fue una revolución inicial, puesto que a fines del siglo XVII se había acumulado ya una larga tradición en mecánica y astronomía. Por tanto, deberemos tratar de averiguar qué tipo de sustitución se operó concretamente en la revolución newtoniana. Y para ello deberíamos tener previamente las ideas un poco más claras sobre el fenómeno de la sustitución en general.

En el análisis de la sustitución en general se plantean en seguida dos problemas de principio. Podemos decir que el primero se refiere a la *forma* de la sustitución y el segundo a los *contenidos* de sus términos.

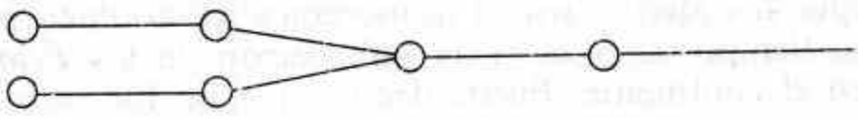
Acerca del primero nos hemos de preguntar si las sustituciones que se producen a lo largo de la historia de una disciplina dada constituyen, dicho en términos matemáticos, un orden lineal o bien simplemente un orden no-lineal. (Es evidente que por lo menos constituyen un orden.) Si representamos un período de ciencia normal por un segmento y una revolución por un punto o «nudo» que separa distintos segmentos, podemos ilustrar las distintas posibilidades del siguiente modo:

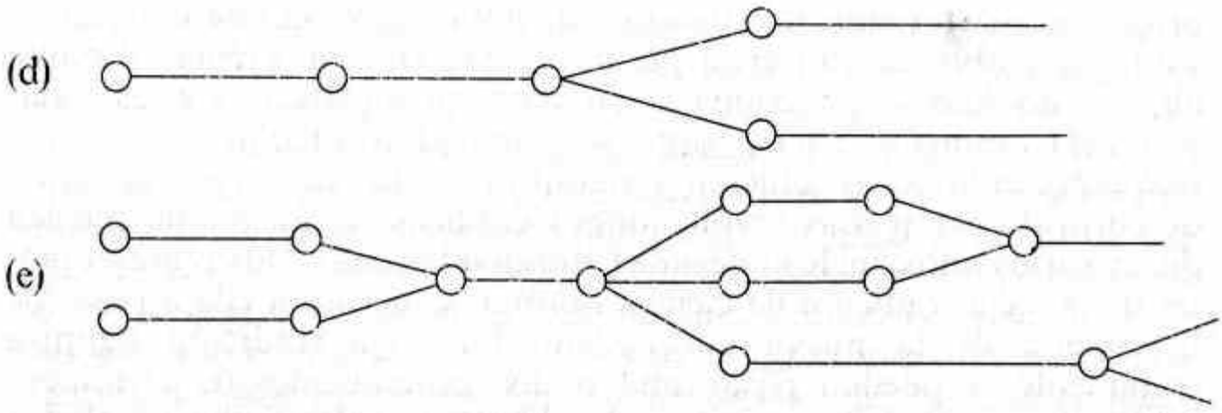
(a)

(b)



(c)





Sólo en el caso (a) tenemos un orden lineal o conexo. Los demás casos son distintos ejemplos de órdenes no-lineales. Lógicamente, todas estas estructuras son posibles para representar la evolución de una disciplina científica. El problema estriba en averiguar cuál de esas estructuras lógicamente posibles es también históricamente adecuada. La hipótesis implícita en la obra de Kuhn es que el orden de las sustituciones (dentro de una disciplina dada) es un orden lineal, es decir, un grafo del tipo (a). En tal caso, en cada período histórico existe *una sola* tradición dominante de ciencia normal y ella es sustituida después de una revolución por otra *única* tradición de ciencia normal, en el período posrevolucionario.

Lakatos, en cambio, en *La falsación y la metodología de los programas de investigación científica*, sostiene la adecuación de un grafo del tipo (b). Los «programas de investigación» de Lakatos corresponden a las «tradiciones de ciencia normal» de Kuhn, y Lakatos sostiene explícitamente que en un mismo período histórico suelen coexistir dos o más programas paralelos en competencia. Ni Kuhn ni Lakatos, sin embargo, parecen haber pensado en la posibilidad de estructuras más complejas, como (c) y (d), y menos en un retículo complicado del tipo (e). No es éste el lugar de discutir en detalle estas otras posibilidades, pero deberían ser tenidas en cuenta en una metateoría general de la evolución científica.

La pregunta ahora es: ¿en cuál de estas distintas posibilidades lógicas encaja realmente la revolución newtoniana? ¿Es un caso de sustitución lineal como sugiere Kuhn, o bien de surgimiento de un programa de investigación paralelo como postularía Lakatos, o bien tiene alguna de las estructuras alternativas? Seguramente no es posible dar una respuesta tajante a estas cuestiones, a menos que se supersimplifique la situación histórica real. Sin embargo, parece que, al menos en una primera aproximación, la estructura kuhniana, es decir, un orden lineal, es la más adecuada para el caso que nos ocupa: la tradición newtoniana sustituyó unívoca y completamente a la tradición cartesiana.

Cabría imaginar dos alternativas en principio. Primera, un partidario del paralelismo *a lo* Lakatos podría argüir que, de hecho, el

programa newtoniano no sustituyó de golpe al programa cartesiano, sino que ambos discurrieron paralelamente en competencia durante algunas décadas: el programa newtoniano en Inglaterra y el cartesiano en el Continente. Sin embargo, sería injusto con Kuhn suponer que éste representa las revoluciones científicas como cambios puntuales, de corta duración. Las revoluciones científicas, según Kuhn, pueden durar varios años, incluso décadas, durante las cuales los representantes de la vieja tradición de ciencia normal se aferran a ella a pesar de la crecida de la nueva concepción. La vieja tradición agoniza lentamente, y pueden pasar una o dos generaciones de científicos hasta que muera definitivamente. La diferencia entre Kuhn y Lakatos no estriba en que para el segundo puedan coexistir programas rivales durante cierto tiempo, mientras que para el primero no, sino que radica en la forma en que uno y otro imaginan esa coexistencia. Para Lakatos, dos programas igualmente potentes y en verdadera competencia pueden coexistir durante un largo período. Para Kuhn, en cambio, la coexistencia sólo se da dentro de un lapso relativamente breve entre un programa nuevo en ímpetu avasallador y uno viejo en lenta agonía, que cada vez convence menos. Creo que, para el caso que nos ocupa, el esquema kuhniano es más aproximado que el de Lakatos. Históricamente está claro que, lo más tarde después de la muerte de Huygens en 1695, el programa cartesiano ya no tuvo nada nuevo que ofrecer. Y mucho antes de la muerte de Huygens dicho programa había entrado ya en una crisis profunda.

→ Otra cuestión es la de determinar si Kuhn tiene razón en suponer que lo sustituido está unívocamente determinado. Es decir, podría ocurrir que una revolución científica sustituyera a más de una tradición de ciencia normal a la vez. Su estructura vendría visualizada entonces por el grafo (c). Se podría argüir con cierta plausibilidad que éste es justamente el caso de la revolución newtoniana: ella desbancó no sólo a la tradición cartesiana, sino también a la aristotélica. Incluso puede documentarse que así lo sintieron los propios revolucionarios newtonianos. El Prólogo de Cotes a la segunda edición de los *Principia* va dirigido claramente contra cartesianos y aristotélicos. También pueden interpretarse en este sentido los comentarios polémicos de Newton en sus *Regulae Philosophandi* y en el *Scholium Generale*.

Sin embargo, esta interpretación pasa por alto la diferencia que existe entre comentarios de tipo retórico dirigidos a un público general y discusiones técnicas de problemas específicos. Fuera cual fuera la influencia que el aristotelismo pudiera ejercer todavía sobre el público educado en general en la época de Newton, lo cierto es que ningún científico profesional se tomaba ya en serio la física aristotélica. Esta había sido enterrada definitivamente por lo menos medio siglo antes. Ya no representaba ningún problema en las discusiones de los físicos profesionales. (Otra cosa distinta es que en las universidades

dominadas por la Iglesia católica siguiera impartiéndose física aristotélica; pero éste ya no es un problema de teoría de la ciencia, sino de teoría de la ideología.) En la juventud de Newton, los trabajos científicos sobre mecánica se hacían dentro de la tradición cartesiana. Incluso físicos muy críticos como Huygens trabajaban dentro de esta tradición. Hasta principios del siglo XVIII, el libro de texto de física que se estudiaba en todas las universidades serias tanto del continente como de Inglaterra era el de Rohault, un cartesiano ortodoxo[1]. Esta y no otra fue la tradición científica con la que rompió Newton. Basta fijarse en el lujo de detalles físico-matemáticos que Newton emplea en el Segundo Libro de los *Principia* para refutar definitivamente la hipótesis cartesiana de los vórtices. Nada parecido en contra de la física aristotélica se encuentra en su obra. En principio, pues, podemos admitir que, en cuanto a su estructura formal, el proceso de sustitución que se dio en la revolución newtoniana puede representarse según el esquema kuhniano de orden lineal. Esto, naturalmente, no prueba, pero sí corrobora su validez general.

Determinado el esquema formal de la revolución newtoniana, pasemos ahora al aspecto más complejo de su contenido. Hemos dicho que una revolución científica es un fenómeno de sustitución. Pero hay que aclarar cuáles son los términos de esta relación sustitutiva. Hasta aquí hemos usado términos muy vagos como los de «tradición de ciencia normal» o «programa de investigación». Es el momento de precisarlos.

En la primera edición de su libro sobre las revoluciones científicas, Kuhn utilizó el término «paradigma» para caracterizar la estructura que determina un período de ciencia normal. Una revolución científica consistiría así en la sustitución de un paradigma por otro. No obstante, las discusiones ulteriores mostraron que se trataba de un término multívoco y confuso. El propio Kuhn reconoció en su *Postcript* de 1969 que fue un error escoger ese término. Sólo una parte de lo que él denominó «paradigma» merecía realmente ese nombre. En el *Postscript*, Kuhn propone un nuevo término: «matriz disciplinaria». Y trata de ser un poco más preciso respecto a sus componentes. Estos son cuatro tipos de cosas, según Kuhn:

1) «Generalizaciones simbólicas»; es decir, las leyes fundamentales de la teoría generalmente expresadas en lenguaje matemático y que contienen los conceptos básicos de la matriz disciplinaria.

2) «Modelos» ontológicos o heurísticos destinados a clarificar la naturaleza de los referentes admisibles de las generalizaciones simbólicas.

[1] Cf. el apéndice de Cajori a su edición de *Sir Isaac Newton's Mathematical Principles*.

3) Valores metodológicos que constituyen el aspecto normativo de la matriz disciplinaria.

4) Ejemplares paradigmáticos que permiten la aplicación sucesiva de las generalizaciones simbólicas a campos nuevos a través de ciertas relaciones de semejanza. Según reconoce Kuhn, ésta es la porción de la matriz disciplinaria que le indujo a usar el término «paradigma».

La estructura general de los componentes primero y último de la matriz disciplinaria ha sido precisada formalmente por Sneed, Stegmüller y sus colaboradores mediante el aparato metateórico de la concepción estructural. Aquí no es necesario entrar en esta precisión formal. Baste notar que el primer componente (las «generalizaciones simbólicas») es más complejo de lo que Kuhh imagina: incluye varios elementos además de las leyes fundamentales, entre otros, la distinción entre el nivel teórico y el empírico de la matriz. Este complejo componente ha sido bautizado con el nombre de «núcleo estructural». En lo que sigue utilizaré esta denominación.

El contenido de una revolución científica consistirá, pues, en la sustitución de una matriz disciplinaria por otra.

Esto implica que, en una revolución científica, uno o varios de los componentes de la matriz serán profundamente alterados para dar paso a una nueva matriz. Generalmente quedará alterado el grueso de los componentes mencionados, dado que están interrelacionados entre sí. Sin embargo, no es necesario que se alteren *todos*. Algunos se pueden conservar después de la revolución. Basta que se altere una gran parte de los componentes de la matriz para que sintamos que nos hallamos ante una genuina revolución científica.

La pregunta ahora es: ¿cuáles de los componentes de la matriz disciplinaria newtoniana permiten caracterizar el advenimiento de ésta como una verdadera revolución?

Empecemos con el núcleo estructural, el elemento más exactamente determinable. La ley fundamental de la mecánica newtoniana es la ecuación dinámica fundamental: $f=m\cdot a$. Podemos añadir también al núcleo de la mecánica newtoniana (aunque no de la mecánica clásica en general) el principio de acción y reacción. Estas dos leyes son, tanto en su forma como en su contenido, completamente nuevas. Nada análogo a ellas se encuentra en la mecánica cartesiana, cuyas leyes fundamentales son las leyes del choque.

Ahora bien, lo que hace a estas dos leyes particularmente revolucionarias es que están formuladas mediante un concepto básico, el de *fuerza*, que considero como el más revolucionario dentro de la revolución newtoniana. Este concepto permite la constitución por primera vez de una genuina dinámica. Claro que el término «fuerza» (*vis* o *actio*) existía ya mucho antes de Newton. Pero lo que importa es que el uso sistemático y matematizado que hace de él Newton es

totalmente novedoso. El concepto de fuerza en la revolución newtoniana es sumamente desconcertante. Se ha visto como sospechoso desde que surgió hasta nuestros días. Huygens y Leibniz veían en él un retorno a las cualidades ocultas. D'Alembert más tarde lo caracterizó como concepto metafísico. Mach lo consideraba superfluo y lo «definió» cinemáticamente. Hertz quiso hacer una mecánica sin fuerzas. Y, en nuestro siglo, Hermes y Simon trataron de axiomatizar la mecánica newtoniana sin que la fuerza apareciera como concepto primitivo. A pesar de todo ello, los físicos siguen usando fructíferamente ese concepto, y todos nos hemos acostumbrado desde el bachillerato a ver el mundo como un mundo de fuerzas newtonianas en acción.

Uno de los aspectos más desconcertantes en el uso que hace Newton del concepto de fuerza, y el que más críticas ha provocado, es su ambigüedad o, mejor dicho, su multivocidad. En efecto, bajo el rótulo general de *vis* o *actio*, Newton subsume cosas aparentemente tan dispares como percusiones, ímpetus (*vis impressa*), fuerzas elásticas y, lo más grave de todo, acciones a distancia. Esta multivocidad y la imposibilidad de determinar de una vez por todas qué es la fuerza a partir de nociones más elementales ha sido piedra de escándalo para muchos críticos de Newton. Sin embargo, desde el punto de vista metodológico actual, esto sólo indica que el concepto newtoniano de fuerza es un concepto fuertemente teórico, por oposición a los conceptos observacionales o semiobservacionales (básicamente cinemáticos) usados en la física anterior. Como es sabido, los conceptos fuertemente teóricos de la ciencia son conceptos abiertos, «porosos», que admiten múltiples interpretaciones, y esto es lo que los hace tan potentes. Una de las características esenciales de la revolución newtoniana fue el haber introducido el concepto más fuertemente teórico manejado hasta entonces.

Pasemos ahora a otro de los componentes de la matriz newtoniana que también intervinieron en la revolución: los ejemplares paradigmáticos. Recuérdese que, según Kuhn, éstos son las primeras aplicaciones exitosas de una matriz disciplinaria, las que justifican, por consideraciones de analogía, la aplicación de la matriz a otros campos. Está claro que los ejemplares paradigmáticos cambiaron radicalmente en el paso de la mecánica cartesiana a la newtoniana. Los paradigmas cartesianos eran los choques elásticos entre esferas perfectas. Su ilustración material idónea era el juego de billar. Podrían añadirse entre los paradigmas cartesianos el péndulo simple y el compuesto, aunque esto quizá fuera cuestión a discutir en detalle. En cualquier caso, ni las bolas de billar ni los péndulos constituyeron los ejemplares paradigmáticos de la mecánica newtoniana que convencieron a la comunidad científica. Estos fueron las órbitas de los planetas, satélites y cometas del sistema solar (las grandes anomalías del sistema cartesiano precisamente); quizá habría que añadir la caída

libre sobre la Tierra, ya que aunque ésta había sido estudiada por Galileo y los cartesianos, fue realmente Newton el primero en incluirla en un marco teórico general y coherente. Obviamente, las «semejanzas de familia» que pueden desprenderse del estudio de las órbitas celestes son muy distintas de las correspondientes a los choques elásticos. Ello explica el gran desarrollo que tuvo la mecánica newtoniana sobre todo como mecánica celeste en los siglos XVIII y XIX.

Consideremos finalmente los modelos ontológicos y los valores metodológicos de la revolución newtoniana. También ellos suponen cambios profundos. La mecánica cartesiana era una mecánica del *plenum*. En ella, materia y espacio se confundían. No se admitía el vacío absoluto, sino sólo distintos tipos de materia más o menos densa. Es cierto que algunos cartesianos heterodoxos como Huygens admitieron la posibilidad del vacío. Pero era más bien una posibilidad remota e inútil para su trabajo teórico. Para explicar los movimientos de las distintas clases de materia se acudía a la hipótesis de los vórtices que Huygens empleó consistentemente para dar cuenta de la gravitación.

Nada de esto se encuentra en Newton. En la ontología newtoniana, el vacío absoluto cobra una importancia preeminente. El estado normal del espacio, por así decir, es el vacío. La materia es la excepción. Y para explicar el movimiento de las partículas materiales se introducen entidades no-materiales: las acciones a distancia. Desaparecen totalmente los vórtices cartesianos.

Espacio y tiempo también sufren una transformación radical. En la ontología cartesiana eran relativos. En la newtoniana son absolutos, entidades subsistentes por sí mismas. Los sistemas de coordenadas son todos intercambiables en la mecánica cartesiana, mientras sean inerciales. En la mecánica de Newton hay un sistema de coordenadas privilegiado puesto por Dios en el centro del sistema solar («Hypothesis II» del Tercer Libro de los *Principia*). Desde el punto de vista matemático también hay cambios en el concepto formal de espacio. El espacio cartesiano es simplemente el espacio euclídeo. Newton le añade a éste un elemento esencialmente nuevo: vectores. Es decir, introduce una estructura afín en el espacio euclídeo normal, obteniendo un espacio vectorial de tres dimensiones[2].

En cuanto a los valores metodológicos de Newton enfrentados a los de la tradición anterior, nos extenderemos sobre ellos en el capítulo siguiente. Me limito aquí a señalar que éste es el punto en que más claramente se ve el choque entre las dos matrices. Newton y sus discípulos polemizaron enérgicamente contra la metodología cartesiana. Tanto el Prólogo de Cotes a los *Principia* como buena parte de las *Regulae Philosophandi* y del *Scholium Generale* van dirigidos contra la metodología cartesiana. Esta les parecía a los

[2] *Cf. S. Bochner, The Role of Mathematics in the Rise on Science*, pp. 191-192.

newtonianos apriorística y metafísica, por oposición a la suya, que sería inductivista y empírica. En suma, a los «vain dreams of our own devising» forjados por los cartesianos, Newton oponía la formulación de leyes empíricas firmemente establecidas por la inducción a partir de los fenómenos. Hoy día matizaríamos toda esta discusión con interpretaciones metodológicas más cuidadosas. Pero ésta es otra cuestión. El punto es que la nueva metodología newtoniana fue sentida por sus contemporáneos también como una gran revolución.

3.2. LOS FUNDAMENTOS METODOLOGICOS DE LA FILOSOFIA NATURAL DE ISAAC NEWTON

0. Introducción

Suele considerarse a Newton, más aún que a Galileo, como el padre de la física moderna. Sus *Philosophiae Naturalis Principia Mathematica*, por su aplicación sistemática de las matemáticas al estudio de los fenómenos naturales y por su apoyo riguroso en la experimentación y la observación, valen como paradigma del método científico tal como se le entiende actualmente.

Tan asombrosos fueron los resultados alcanzados por Newton para sus contemporáneos y sucesores que nadie se preocupó mucho por las concepciones metodológicas y filosóficas generales que habían sido su premisa histórica. Para las generaciones de científicos inmediatamente posteriores a Newton sólo cabía la admiración irrestricta ante su obra, no el análisis crítico de sus premisas. Y los físicos y metodólogos más recientes, que han asimilado las teorías de Newton como parte del bagaje científico moderno, han tendido a presuponer que también la metodología de Newton concuerda con los cánones actuales. El razonamiento es claro, aunque falaz: «Si los resultados concretos de Newton siguen siendo en principio válidos para nosotros, también lo han de ser sus presupuestos metodológicos.»

La dificultad de lectura de los *Principia* para el científico moderno contribuyó también a que, con el tiempo, Newton y su obra se convirtieran en un mito intocable e inanalizable.

Sólo en décadas recientes los esfuerzos de historiadores de la ciencia como Alexandre Koyré, I. Bernhard Cohen, Marie Boas, A. R. Hall, Gerd Buchdahl y otros nos han proporcionado una imagen

mucho más realista de la labor científica de Newton y de sus presupuestos históricos. Se han desenterrado muchos de los aspectos de Newton que habían sido engullidos por el «memory hole» de la historiografía científica oficial, tales como la influencia que ejercieron en la concepción general newtoniana los escritos de los neoplatónicos británicos de la época o el hecho de que Newton dedicara tanto tiempo y energía a sus investigaciones de alquimia como al resto de su obra científica[3].

No obstante, a pesar de lo mucho escrito sobre Newton en tiempos recientes, todavía falta una apreciación cabal de sus concepciones metodológicas generales que sea relevante para el filósofo de la ciencia.

No me voy a referir aquí a temas concretos de la filosofía newtoniana, como podrían ser su concepción del espacio absoluto o su noción de fuerza, sino más bien a temas más generales, de los que la filosofía de la ciencia se ha ocupado intensamente, como la inducción, el carácter de las hipótesis científicas y las explicaciones teleológicas. A través del estudio de lo que Newton opinaba sobre estos puntos, trataremos de comprender su idea general acerca del funcionamiento y el objetivo de la empresa científica; esa idea, como veremos, puede que no sea muy del gusto de los metodólogos actuales, pero también es lo suficientemente coherente como para que pueda hablarse de una metodología newtoniana característica. Tratemos de ver cuál es ésta.

1. Las «Regulae philosophandi»

El acceso natural a la metodología de Newton lo constituyen sin duda las reglas del pensamiento científico que él propone explícita y enfáticamente al comienzo del tercer libro de los *Principia*. Estas reglas se pueden considerar, y así se han considerado usualmente, como una profesión de fe *inductivista*. Por ello se ha visto en Newton a un paladín del método inductivo. Sin embargo, ya aquí debemos hacer dos cualificaciones: en primer lugar, que el inductivismo de Newton muestra importantes puntos de divergencia con respecto a lo que actualmente se entiende por método inductivo; en segundo lugar, que las concepciones metodológicas de Newton no quedan agotadas, ni mucho menos, con las «Regulae philosophandi» y que hay que tener en cuenta el contexto polémico en que ellas surgieron.

[3] Todavía no se ha podido calibrar exactamente la significación que la alquimia tuvo en la labor científica de Newton, pero las investigaciones de estos últimos años parecen señalar que fue enorme y plenamente insertada en la tradición alquimista medieval y renacentista. El hecho de que la historiografía científica haya tardado casi tres siglos en redescubrir un Newton alquimista es digno de ser notado.

Pero antes de adentrarnos en las interpretaciones veamos el texto de las reglas, con algunos de los comentarios y especificaciones más reveladoras que Newton les añade (cf. *Newton's Principia*, páginas 378-400)[4].

REGLA I

No debemos admitir más causas de las cosas naturales que aquellas que son a la vez verdaderas y suficientes para explicar sus apariencias.

A este fin, los filósofos dicen que la Naturaleza no hace nada en vano, y más es en vano cuando menos hace el servicio; pues la Naturaleza se complace con la simplicidad y no afecta la pompa de las causas superfluas.

REGLA II

Por lo tanto, a los mismos efectos naturales les debemos asignar, en la medida de lo posible, las mismas causas.

Estas dos primeras reglas podrían resumirse, en realidad, en una sola: la segunda regla es una especificación de la primera; el propio Newton lo indica al anteponer «por lo tanto» al texto de la segunda regla. Regla I junto con Regla II constituyen lo que hoy denominaríamos un principio de economía metodológica, una forma de la famosa «navaja de Occam»: no se deben admitir hipótesis superfluas para explicar los fenómenos naturales; los fundamentos de una disciplina científica deben ser lo más simples posible.

La cuestión que se plantea de inmediato, naturalmente, es: ¿cómo se llega a ese fundamento sólidamente simple? La respuesta de Newton se halla contenida (al menos en primera aproximación) en las reglas tercera y cuarta: el conocimiento de las causas suficientes se obtiene por inducción a partir de los fenómenos.

REGLA III

Las cualidades de los cuerpos que no admiten intensificación ni disminución de grado, y que se encuentra que pertenecen a todos los cuerpos dentro del alcance de nuestros experimentos, han de considerarse cualidades universales de todos los cuerpos, sean cuales sean.

[Comentario] ... Ciertamente no debemos abandonar la evidencia de los experimentos en aras de sueños y vanas ficciones de nuestra propia invención; ni tampoco hemos de apartarnos de la analogía de la naturaleza, que acostumbra a ser simple y siempre en consonancia consigo misma.

---

[4] Dado que no existe traducción castellana de los *Principia* vierto al castellano los pasajes pertinentes de la traducción inglesa de Florian Cajori, revisión, a su vez, de la que hizo Andrew Motte en 1729, bajo la supervisión del propio Newton.

Son también iluminadores (e inaceptables para muchos filósofos actuales) los ejemplos con que Newton aclara esta regla:

> Que todos los cuerpos son impenetrables lo aprehendemos no por la razón, sino por la sensación. Encontramos que los cuerpos que manejamos son impenetrables, y de ahí concluimos que la impenetrabilidad es una propiedad universal de todos los cuerpos, sean cuales sean. Que todos los cuerpos son móviles y provistos de ciertas capacidades (que llamamos inercia) de perseverar en su movimiento o en su reposo, lo inferimos solamente de las propiedades análogas observadas en los cuerpos que hemos visto... Finalmente, si aparece universalmente, por experimentos y observaciones astronómicas, que todos los cuerpos sobre la Tierra gravitan hacia la Tierra, y esto proporcionalmente a la cantidad de materia que contenga cada uno de ellos; que asimismo la Luna, de acuerdo con la cantidad de su materia, gravita hacia la Tierra; que, por otro lado, nuestro mar gravita hacia la Luna; y todos los planetas uno hacia el otro; y los cometas de igual manera hacia el Sol; como consecuencia de esta regla hemos de admitir que todos los cuerpos, sean cuales sean, están provistos de un principio de mutua gravitación.

REGLA IV

> En la filosofía experimental debemos buscar proposiciones que se infieran por inducción general a partir de los fenómenos de manera exactamente verdadera o muy cercana a la verdad, a pesar de cualesquiera hipótesis contrarias que puedan imaginarse, hasta que llegue el momento en que ocurran otros fenómenos, a través de los cuales puedan hacerse o bien más exactas o bien sujetas a excepciones.
>
> [Comentario]. Debemos seguir esta regla, de modo que el argumento por inducción no pueda evadirse con hipótesis.

En la interpretación de estas reglas hay que investigar cuidadosamente lo que Newton concebía como «inducción» y «fenómeno». Su concepto de inducción es sin duda más amplio de lo que los actuales filósofos de la ciencia entienden por «inducción». La inducción newtoniana no atañe sólo al caso en que del hecho de que todos los objetos *observados* de un dominio empírico poseen una determinada propiedad se infiere que *todos* los objetos del dominio poseen esa propiedad; se refiere también (y principalmente) al caso distinto en que de las observaciones hechas en un dominio se hace una inferencia acerca de otro dominio, distinto del primero, pero *análogo* en algún sentido. Por ejemplo, de la observación de cuerpos en caída libre sobre la superficie de la Tierra se hacen inferencias sobre los movimientos planetarios. Dicho brevemente, la «inducción» newtoniana incluye el razonamiento analógico además de la inducción «normal».

Las inferencias por analogía juegan un papel de primer orden tanto en los *Principia* como en la *Optica* de Newton. Fueron consideraciones de analogía las que llevaron a Newton a postular la

ley de gravitación universal. Mediante la inducción en sentido estricto sólo podía inferir, por ejemplo, que la Tierra ejerce una atracción sobre todos los cuerpos cercanos a su superficie, incluidos los no-observados, y, estirando mucho la inducción, que también ejerce una atracción sobre la Luna[5]. Pero por mucho que se liberalicen las reglas de la inducción en sentido usual, Newton no podía inferir mediante ellas que todos los cuerpos del universo se atraen mutuamente. Para llegar a esta última proposición, Newton tuvo que recurrir al razonamiento por analogía: considerar que el predicado «cuerpo cercano a la superficie terrestre (incluyendo la Luna)» es lo suficientemente análogo al predicado «cuerpo girando alrededor del Sol», por lo que infirió que el Sol ejerce una atracción sobre los planetas del mismo tipo que la que ejerce la Tierra sobre la Luna y las manzanas. Y de ahí pasó, nuevamente apoyado en la analogía, a la proposición general sobre la atracción universal. Todo esto entendía Newton por «inducción», bastante más de lo que entienden los actuales filósofos de la ciencia.

Pero el concepto newtoniano de inducción presenta aún más complejidades que las reseñadas. El uso implícito de la inducción que hace Newton contiene la idea de que las inferencias inductivas ocurren por lo menos a dos niveles, que en nuestra terminología actual quizá se denominarían el nivel de las «regularidades empíricas» y el nivel de las «leyes teóricas». En un primer paso inductivo, un número finito de observaciones sobre un dominio dado nos permite establecer una regularidad empírica o «ley de bajo nivel» sobre ese dominio. Así, por ejemplo, de un número finito de observaciones astronómicas sobre las posiciones de los planetas se infieren las leyes de Kepler. Esto es lo que Newton denomina en el tercer libro de los *Principia* un *fenómeno* (no una «ley»). Inmediatamente después de las «Regulae Philosophandi», Newton establece seis «Phaenomena» en ese sentido, entre ellos las leyes originarias de Kepler. Partiendo de los teoremas del primer libro deducibles de las leyes generales del movimiento, pudo «explicar» a su vez esos «fenómenos» mediante el supuesto de una fuerza centrípeta inversamente proporcional al cuadrado de la distancia. Y de ahí infirió que todos los astros están sometidos al mismo tipo de fuerza, la gravitación. Es decir, las leyes de Kepler y las leyes análogas para los satélites le sirvieron de apoyo «inductivo» para el establecimiento de la ley de la gravitación universal.

La ley así establecida debe ser admitida «a pesar de cualesquiera hipótesis contrarias que puedan imaginarse», como reza la Regla IV. Esta observación es naturalmente de carácter polémico. Es una defensa del método «inductivo» (como lo entendía Newton) frente al

[5] Que esta última inferencia se considere inductivamente válida o no, depende de si el predicado «cuerpo cercano a la Tierra» se «estira» lo bastante como para que incluya la Luna.

método «deductivo» (como lo entendían sus adversarios), lo mismo que el comentario a la Regla III:

«Ciertamente no debemos abandonar la evidencia de los experimentos en aras de sueños y vanas ficciones de nuestra propia invención.»

Tales comentarios críticos y tal defensa del «inductivismo» iban dirigidos contra aquellos que especulaban sobre la Naturaleza sin hacer experimentos ni observaciones. Se trataba, para Newton, de unos adversarios muy concretos (y muy poderosos entonces): los cartesianos.

2. «Nonne hypotheses finxisti?»

El rechazo de la especulación sobre la Naturaleza sin comprobación empírica está evidentemente relacionado con la famosa sentencia de Newton contra las «hipótesis» en el *Scholium Generale* de los *Principia:* «Hypotheses non fingo»[6].

La lectura ingenua de esa frase, con negligencia de su contexto histórico, hizo surgir una imagen ahistórica de la metodología newtoniana, aceptada generalmente hasta hace poco. Era la imagen de un Newton que sólo se preocupaba por los hechos experimentales y que rechazaba estrictamente la formulación de hipótesis generales. Se vio en Newton al prototipo del inductivista riguroso (en el sentido del inductivismo del siglo XIX) o se le consideró incluso como el predecesor de la metodología positivista. De esta interpretación ahistórica de Newton encontramos un ejemplo interesante por su contundencia en la *Mecánica* de Ernst Mach:

«La repetida declaración explícita de Newton de que no le interesaban las especulaciones sobre las causas ocultas de los fenómenos, sino la investigación y constatación de los *hechos*, la línea de pensamiento que se expresa clara y concisamente en sus palabras «*hypotheses* non fingo», le caracteriza como a un *filósofo* de rango *eminente*» (subrayado de Mach).

Y por «filósofo de rango eminente» está claro que entendía Mach alguien que concordaba con su propia concepción positivista.

Como suele ocurrir con tales interpretaciones simplistas hechas por personas inteligentes, la interpretación machiana no es *completa-*

[6] El párrafo en que se encuentra esta frase es el siguiente (en la traducción de la versión inglesa de Motte-Cajori): «Pero hasta ahora no he podido descubrir la causa de esas propiedades de la gravedad a partir de los fenómenos, y no invento hipótesis; pues todo lo que no se deduce de los fenómenos debe llamarse una hipótesis; y las hipótesis, ya sean metafísicas o físicas, de cualidades ocultas o mecánicas, no tienen lugar en la filosofía experimental.» [*Newton's Principia*, p. 547].

mente falsa, pero sí producto de una grave deformación histórica. El rechazo que hace Newton en el *Scholium Generale* de toda especulación o «hipótesis» no debe sacarse del contexto histórico (y, en particular, polémico) en que se manifestó. Pero es asimismo una deformación histórica el intento de reinterpretar (o «redimir») a Newton, como han hecho incluso autores recientes, presuponiendo que él entendía el término «hipótesis» en el mismo sentido en que lo entendemos hoy. No tiene mucho interés empezar a discutir si Newton realmente inventaba o no hipótesis en el sentido actual (¡claro que lo hacía!). Lo que es interesante para la historia de la ciencia es averiguar por qué y con qué intención Newton, en un momento dado, en un contexto dado, hizo su tajante declaración: «Hypotheses non fingo», y qué entendía entonces por «hypotheses».

Lo primero que hay que observar es que el propio Newton en sus escritos no publicados anteriores a los *Principia*, como por ejemplo en el importante ensayo *De Motu Corporum*, 1684[7], no tuvo inconveniente alguno en usar el término «hipótesis» como sinónimo de «ley general».

Asimismo, y para vergüenza de intérpretes apresurados, en *todas* las versiones publicadas de los *Principia* aparecen varias «*hypotheses*» (denominadas así explícitamente), coexistiendo pacíficamente con el ataque del *Scholium Generale* a las hipótesis en general. En el texto de la segunda y tercera edición de los *Principia* (las ediciones generalmente conocidas) pueden leerse las siguientes «Hypotheses»:

— En el *Segundo Libro*, Sección IX, se formula una hipótesis sobre la expresión matemática de «la resistencia de los fluidos debida a la carencia de lubricidad» [*Newton's Principia*, p. 385].

— En el *Tercer Libro*, en la sección sobre Teoría de la Luna, leemos una «hipótesis» sobre la «precesión de los equinocios» [*Newton's Principia*, p. 489].

— Finalmente, al comienzo del *Tercer Libro*, encontramos una «hipótesis» que alguien como Ernst Mach debiera haber calificado de «insensata especulación»: «El centro del sistema universal es inmóvil» [*Newton's Principia*, p. 419].

Pero más aún. Alexandre Koyré e I. Bernhard Cohen fueron los primeros en hacer notar que en la primera edición de los *Principia* (1687) las dos primeras *Regulae Philosophandi* y los *Phaenomena* subsiguientes aparecen formulados como «*Hypotheses*» (!). Claro que en esta primera edición no se encuentran todavía ni las polemizantes

[7] Cf. Hall y Hall, *Unpublished Scientific Papers...*, manuscritos B y C, p. 243. *De Motu Corporum* puede considerarse como el «ensayo general» de los *Principia*.

reglas tercera y cuarta, ni el *Scholium Generale* con el anatema a las hipótesis[8].

Hasta la época de la primera edición de los *Principia*, e incluso más tarde, Newton no tuvo ningún prejuicio en contra de las hipótesis. ¿Cuál fue la piedra de escándalo que motivó su posterior condena? No podemos determinarlo con entera seguridad, pero aparece como muy probable que su anatema fuera una reacción airada a las críticas severas de muchos científicos de la época (en particular, cartesianos) a su teoría de la gravitación. Para muchos de los primeros lectores de los *Principia* (científicos de primera fila), la idea misma de la gravitación universal aparecía como una ficción inaceptable.

Así, por ejemplo, Cohen ha constatado[9] que en una reseña francesa de los *Principia* publicada en el *Journal des Sçavans* de 1688 (la revista científica más importante de la época) se califica el supuesto de la gravitación universal de «hipótesis arbitraria». Es seguro que Newton leyó esa reseña, pues tenía el *Journal des Sçavans* en su biblioteca, y es probable que se enojara bastante al leer los términos en que los franceses liquidaban su teoría de la gravitación, que tanto esfuerzo le había costado.

Tampoco podía alegrarse mucho por las reacciones críticas de dos de los máximos científicos contemporáneos, Leibniz y Huygens. Ambos consideraban la gravitación universal como una entidad imaginaria. En 1690 le escribía Huygens a Leibniz en una carta sobre temas de física: «Por lo que respecta a la Causa de la Marea que da Mr. Newton, no me satisface en absoluto, ni tampoco todas sus demás Teorías que él construye sobre su principio de atracción, el cual me parece absurdo... Y frecuentemente me ha asombrado que él se haya podido dar el trabajo de tantas investigaciones de cálculos difíciles, que no tienen más fundamento que este mismo principio» [cf. Gerhard, *Briefwechsel...*, p. 156].

Y Leibniz escribía en 1715 en una carta al abad Conti, refiriéndose a los fundamentos de la teoría newtoniana:

«Su filosofía me parece un poco extraña, y no creo que pueda establecerse. Si todo cuerpo es grave, es necesario que la gravedad sea una cualidad oculta escolástica o el efecto de un milagro» [cf. Newton, *Comercium epistolicum*, *Opera...*, t. IV, p. 596].

Ni Leibniz ni Huygens lograron nunca quedar convencidos de la validez de la concepción newtoniana. Más aún, les repugnaba la idea misma de la atracción a distancia (sin intermediario físico), porque sospechaban tras ella un retorno a las «fuerzas ocultas» o «espirituales» y, por tanto, a una visión animista del Universo, tan característica

---

[8] La primera versión del *Scholium Generale* es de 1713 –veintiséis años posterior a la primera publicación de los *Principia*.

[9] Cf. Cohen, *Introduction to Isaac Newton's Principia*, p. 156.

de la época renacentista y que tanto esfuerzo les había costado vencer a los «mecánicos» del Barroco. Es históricamente comprensible que a muchos de ellos les pareciera la «filosofía natural» de Newton como la de un reaccionario oscurantista.

Ahora bien, es importante notar que los adversarios más eminentes de Newton no rechazaban la *ley de la gravitación* como expresión matemática de la relación entre los fenómenos. Lo que les molestaba era la introducción del *concepto de gravitación* como principio de explicación, pues para ellos era este concepto mismo lo que precisaba una explicación.

La réplica de Newton a esas críticas podría quizá resumirse así: «No podéis negar la ley de la gravitación, puesto que se apoya en observaciones empíricas y cálculos exactos. Pero si lo que pedís es una *explicación* de la gravitación, entonces estáis exigiendo demasiado de mí. Pues desconozco la causa de la gravitación, *y a este respecto* no pienso inventarme hipótesis.»

Sobre la causa de la gravitación, Newton no quería pronunciarse —por lo menos en los *Principia*. Su posición a este respecto queda aún más clara en su tercera carta a Bentley, quien le había preguntado cómo concebía la gravitación: «La gravedad debe ser causada por un agente que actúa constantemente de acuerdo con ciertas leyes; pero que este agente sea material o inmaterial lo he dejado a la consideración de mis lectores» [*Newton, Comercium epistolicum, Opera...*, t. IV, p. 596][10].

Resumiendo, el rechazo que hace Newton enfáticamente en el *Scholium Generale* está indisolublemente referido a la discusión sobre el concepto de gravitación y a la exigencia de sus críticos de que debía explicar la gravitación. En este contexto, y sólo en éste, debe interpretarse su famoso *dictum* contra las hipótesis.

Sin embargo, tampoco hay que caer en el extremo contrario de interpretación y suponer que la enfática declaración «Hypotheses non fingo» no es más que un «lapsus linguae» de un Newton agriado por la discusión. Es cierto que la frase en cuestión y el párrafo que la acompaña surgieron primordialmente con ocasión de la fuerte polémica que Newton sostenía en torno a la gravitación; pero no es menos cierto que Newton, en su época madura, fue cada vez más parco en el uso del término «hipótesis», y no sólo por razones polémicas. Si bien aparecen formuladas algunas «hypotheses» en los *Principia*, téngase en cuenta que no son más que *tres* en una voluminosa obra de carácter innovador. Y también debería observarse que en su obra publicada

[10] Nótese aquí que Newton deja abierta la puerta a una interpretación espiritualista u «ocultista» de la gravitación. Si fuera sólo por esta carta, podríamos pensar que se trata de una cortesía al teólogo Bentley. Pero más adelante expondremos más datos que nos inducen a creer que Newton había pensado seriamente en una explicación espiritualista de la gravitación y de otras leyes fundamentales. Por eso, la sospecha de Leibniz, Huygens y otros «ilustrados» no era enteramente infundada.

Newton nunca calificó las leyes fundamentales de la dinámica (los «tres principios de Newton») de «hipótesis». En su época madura, Newton siempre distinguió netamente las leyes generales que consideraba establecidas de lo que *él llamaba* «hipótesis». Y si bien es cierto que postuló algunas «hypotheses» (en *su* sentido), las manejó también con suma precaución.

3. Hipótesis como preguntas

Buena parte de la discusión reciente sobre «Hypotheses non fingo» viene viciada de origen por la falacia implícita de considerar que Newton usaba un concepto de «hipótesis» idéntico o muy semejante al que se usa hoy día, por ejemplo cuando se habla del «modelo hipotético-deductivo» de las teorías científicas. Si se aplica el concepto actual de «hipótesis» es casi trivial concluir que Newton postuló un gran número de hipótesis, tanto en los *Principia* como en otros escritos.

Sin embargo, una lectura atenta de Newton revela que, por lo menos en su época madura, él hacía una neta distinción, que los metodólogos actuales no suelen hacer, entre dos tipos de enunciados generales sobre la Naturaleza: «leges» e «hypotheses». Las primeras son proposiciones confirmadas empíricamente («inferidas por inducción general a partir de los fenómenos»), prácticamente indubitables, como las «leges motus» o la ley de gravitación universal; las segundas son mucho más problemáticas, pues carecen de confirmación, y si acaso se formulan alguna vez deben manejarse con suma precaución. Las «hypotheses» deben tomarse a lo sumo como estímulo inicial para el investigador, no como base ni como resultado inatacable de la investigación sistemática.

En su obra tardía Newton incluso cambió de término para las «hypotheses» (seguramente para evitar conflictos) y las denominó *queries:* la única justificación metodológica de las «hypotheses» era para Newton el ser *preguntas* hechas por el investigador a la Naturaleza; pueden figurar en el estadio inicial de la investigación, no entre los resultados finales. Muestra clara de esta concepción es la *Opticks*, redactada bastante más tarde que los *Principia.* La *Opticks* está dividida claramente en dos partes: la primera, sistemática, expone los resultados obtenidos por el propio Newton y por sus antecesores (las leyes de reflexión y de refracción, por ejemplo), que Newton consideraba seguros; en esta parte no aparece ninguna «hypothese». La segunda parte, en cambio, consta de más de treinta «queries», algunas muy especulativas, que expresan las preguntas que se hacía Newton acerca de una serie de fenómenos y experimentos poco o mal explicados. Las «queries» son sólo otro nombre para las hipótesis.

Como instrumento heurístico, Newton admitía las hipótesis,

siempre y cuando se concibieran como *preguntas* –no como principios ni como resultados. La expresión más clara de esta idea no se debe al propio Newton, sino a su discípulo y amigo Cotes, quien por supuesto sabía muy bien lo que opinaba su maestro, generalmente parco en explicaciones. En el *Prólogo* a la segunda edición de los *Principia*, Cotes expone claramente el punto de vista newtoniano acerca de las hipótesis. Después de haber atacado a aristotélicos y cartesianos por sus explicaciones meramente verbales de los fenómenos, expone el «buen método», el de los «filósofos experimentales», que es naturalmente el método de Newton: «Estos [los newtonianos] ciertamente derivan las causas de todas las cosas a partir de los principios más simples posibles; pero entonces no asumen nada como principio que no esté probado por los fenómenos. No inventan hipótesis ni las admiten dentro de la filosofía a no ser como preguntas cuya verdad puede discutirse [*Newton's Principia*, p. XX].

Nótese que el equilibrado Cotes es aquí mucho menos radical en su expresión que el Newton del *Scholium Generale:* no hay que postular hipótesis, *a menos que se las tome como preguntas sobre cuya validez puede discutirse.* Esta concepción iba a ser aplicada en la práctica en la *Opticks*, donde las hipótesis, para evitar confusiones, tomarían incluso la forma gramatical de preguntas. La concepción madura de Newton sobre las hipótesis es clara y coherente; sólo podemos asombrarnos de que muchos de los comentaristas actuales hayan pasado por alto esa conexión tan directa y reveladora entre el Prólogo de Cotes y las «Queries» de la *Optica.*

4. La insuficiencia de las explicaciones mecánicas

La imagen popular de un Newton contrario a las hipótesis suele ir acompañada de la imagen, aparentemente más plausible, de un Newton mecanicista, que concebía el Universo como un enorme mecanismo de relojería y que buscaba en todos los dominios la aplicación exclusiva de las leyes de la mecánica. Esta imagen puede que corresponda a muchos de los predecesores, contemporáneos y sucesores de Newton, pero decididamente no al propio Newton. Si bien en sus *Principia* había proporcionado la mejor teoría empírica conocida hasta la fecha partiendo de principios puramente mecánicos, Newton nunca creyó –al reves de muchos de sus contemporáneos, como Huygens, Leibniz y los cartesianos– que una explicación completa y general de la Naturaleza pudiera conseguirse a partir de un fundamento puramente mecánico. Lo que luego se convino en llamar «la concepción newtoniana del Universo», que tanto entusiasmó a los intelectuales de los siglos XVIII y XIX (la fe en que «todo puede reducirse a la mecánica»), no habría sido suscrita por su supuesto fundador. Newton no sólo no fue partidario del mecanicismo, sino que fue su decidido adversario.

Para comprender mejor la posición de Newton a este respecto, el intérprete moderno debe tener en cuenta que, en la época de Newton, «explicación mecánica» prácticamente equivalía a «explicación natural»; o, dicho de otro modo, dado que la única ciencia natural realmente constituida como disciplina explicativa era la mecánica[11], resulta históricamente comprensible que no se concibiera otro modo «natural» de explicar los fenómenos que el «modo mecánico». Por ello, al rechazar Newton el mecanicismo, tenía que rechazar también, como veremos, el «naturalismo», entendiendo por tal la creencia, usual hoy día entre los científicos, de que todo fenómeno o proceso observable en la Naturaleza es («en principio», como se dice) explicable a partir de alguna teoría perteneciente a las ciencias naturales reconocidas.

La actitud de Newton frente al mecanicismo estuvo, sin duda, influida por ciertas corrientes ideológicas de la época, en particular por la *escuela neoplatónica* de Cambridge, que tuvo su apogeo durante la segunda mitad del siglo XVII. La Escuela de Cambridge reaccionó muy polémicamente a la difusión del mecanicismo, sobre todo en la forma propagada por Hobbes, y se propuso como objetivo fundamental hallar argumentos en favor de una concepción espiritualista del Universo. P. M. Rattansi ha constatado en su trabajo sobre las corrientes neoplatónicas de los siglos XVI y XVII[12] que Newton, durante la redacción de su primer trabajo sobre la gravitación, leía simultánea y ávidamente los escritos de Henry More (uno de los principales representantes de la Escuela de Cambridge). Rattansi cree probable que esas ideas neoplatónicas obraran en Newton como antídoto al mecanicismo de Hobbes y Descartes.

Para mostrar la insostenibilidad del naturalismo mecanicista, Newton, siguiendo a los neoplatónicos de Cambridge, trató de ofrecer argumentos de carácter más bien empírico que metafísico o ideológico. Entre sus argumentos puramente empíricos estaba, por ejemplo, el de que la estructura y ordenación especiales del sistema planetario no se podían explicar a partir de supuestos exclusivamente mecánicos. Newton había comprobado que todos los planetas conocidos en su tiempo se movían aproximadamente sobre el mismo plano, todos en el mismo sentido, y que esto también era el caso para los diez satélites por él conocidos. No le parecía posible imaginar una causa mecánica para una regularidad tan asombrosa. Así lo expresa en el *Scholium*

[11] No puede considerarse que la astronomía pre-newtoniana (es decir, la astronomía antes de ser absorbida por la mecánica) ofreciera algún *corpus* teórico de *explicación;* se trataba «meramente» de descripciones. Ni siquiera en el estadio más avanzado de la astronomía preclásica, las leyes de Kepler, los científicos de la época sintieron que se había llegado a una «explicación del Universo». Por eso Newton llamó a las leyes de Kepler «*Phaenomena*», no «*Principia*», ni «*Leges*».

[12] Cf. Rattansi, *Some Evaluations of Reason in Sixteenth and Seventeenth-Century Natural Philosophy*, pp. 154 y ss.

Generale: «No se puede concebir que causas meramente mecánicas pudieran dar lugar a tantos movimientos regulares, dado que los cometas recorren todas las partes de los cielos en órbitas muy excéntricas. [...] Este sistema tan hermoso del sol, los planetas y los cometas sólo puede provenir del consejo y el dominio de un Ser inteligente y poderoso» [*Newton's Principia*, p. 554].

Otro argumento empírico en favor de una explicación espiritualista del Universo que da Newton a renglón seguido sólo tiene sentido partiendo de su teoría de la gravitación: si todos los cuerpos se atraen mutuamente, ¿cómo es posible que todos los astros no choquen reduciéndose a una única masa cósmica? «Y para que los sistemas de las estrellas fijas no caigan, por su gravedad, unos sobre otros, él [a saber, el Agente Inteligente] los ha colocado a distancias inmensas unos de otros» [*Newton's Principia*, p. 544].

Pero tampoco el problema del origen del Universo podía ser resuelto por una filosofía mecanicista. En la *Query* 31 de su *Opticks* sostiene Newton que es «antifilosófico», debido a una mala metafísica, «pretender que él [es decir, el Universo] puede surgir del Caos por las meras leyes de la naturaleza... Una uniformidad tan maravillosa en el sistema planetario debe considerarse como el efecto de una elección. Y lo mismo debe valer para la uniformidad en los cuerpos de los animales» [*Opera...*, t. IV, pp. 261-262].

Y en su primera carta a Bentley, quien le había pedido aclaraciones filosóficas sobre su sistema, confirma: «Los movimientos que ahora tienen los planetas no pudieron surgir de ninguna causa natural solamente, sino que se los imprimió un Agente Inteligente» [*Opera...*, t. IV, p. 431].

¿Significan estas declaraciones antinaturalistas que, para Newton, las causas últimas de los fenómenos naturales ya no podían ser objeto de la investigación científica, que su indagación debía dejarse a la teología metafísica? En absoluto; para Newton, la «ciencia de la Naturaleza» *no* era identificable con la Mecánica ni con la «ciencia natural» en nuestro sentido. Aunque algo escapara a la Mecánica, no por eso dejaba de ser asequible al método científico, al método «inductivo», como lo llama Newton. Está bien aplicar la Mecánica hasta donde sea posible; pero, superado el límite de la Mecánica, el científico no debe cerrar los ojos y abandonar el resto a una metafísica especulativa. Por el contrario, debe buscar, con la misma metodología de siempre, nuevos principios explicativos. La explicación causal-mecánica de la Naturaleza puede y debe ser completada por principios no-mecánicos, basados en la hipótesis de una realidad espiritual. De hecho, Newton no admitía la separación tajante, y ya corriente en su época, entre ciencia y teología. Es posible que incluso concibiese su propia obra como el primer paso hacia una «teología empírica», basada en experimentos y matemáticas, y que, procediendo por «inducción», permitiera un conocimiento científico y paulatino de

Dios. Esto explicaría también el hecho (algo desconcertante para los estudiosos modernos de Newton) de que el autor de los *Principia* y la *Opticks* estuviera, a la vez, profunda y genuinamente interesado en cuestiones teológicas.

5. La teleología newtoniana

Pero veamos más detenidamente cómo concebía Newton una metodología científica que superara el marco del mecanicismo y permitiera abordar las cuestiones teológicas.

Ante todo, distingue dos grandes tipos de principios explicativos o «leyes generales»: *principios activos y principios pasivos.* Los principios pasivos son de carácter puramente mecánico, es decir, constituyen el fundamento de la Mecánica. (En la terminología de Newton, podría definirse la Mecánica precisamente como la «ciencia de los principios pasivos».) El ejemplo típico de principio pasivo es el *principio de inercia*, su primera *lex motus.* Los principios pasivos son el máximo tipo de generalizaciones empíricas que se pueden justificar mecánicamente. El descubrimiento de principios pasivos es importante; pero ellos solos no explican todo. «La fuerza de inercia es un principio pasivo... Por medio de este principio solo nunca hubiera podido haber ningún movimiento en el mundo. Fue necesario algún otro principio para poner los cuerpos en movimiento, es necesario algún otro principio para conservar el movimiento» [*Query* 31, *Opticks*, *Opera...*, t. IV, p. 258].

A esos principios adicionales, que permiten una explicación más completa de la Naturaleza y que no pueden concebirse mecánicamente, los denomina «principios activos». El ejemplo más famoso es nuevamente el principio de gravitación, es decir, el principio según el cual todos los cuerpos del Universo se ejercen una acción a distancia cuya magnitud viene fijada por la ley de gravitación. Sin la admisión de este principio como algo externo a la Mecánica (pues en verdad no es deducible de las *leges motus*), Newton no habría concebido la teoría matemática de la gravitación.

También en otros campos, que en la época de Newton estaban poco estudiados, como el dominio de los efectos eléctricos y magnéticos o el de las reacciones químicas, deberían buscarse principios activos de explicación. La búsqueda de tales principios en los campos menos conocidos de la ciencia es un *leitmotiv* de las *Queries* de la *Optica:* «Estas Partículas no sólo tienen una *Fuerza de Inercia*, acompañada de las leyes pasivas del movimiento que resultan naturalmente de esa fuerza, sino que también... las mueven ciertos principios Activos, tales como el de la gravedad, y el que causa la fermentación y la cohesión de los cuerpos. A estos principios no los considero propiedades Ocultas..., sino leyes generales de la Naturaleza

a través de las cuales se forman las cosas en sí mismas: su verdad aparece ante nosotros a través de los fenómenos, aunque sus causas aún no se hayan descubierto» [*Opera...*, t. IV, pp. 260-261].

Por lo que conozco, Newton no especificó nunca las características concretas de los principios activos; probablemente, él mismo no tenía las ideas claras al respecto, y en tal situación, «hypotheses non finxit». De hecho, Newton sólo dio ejemplos concretos de principios activos, no una caracterización general. Sus ejemplos son, ante todo, la gravitación, pero también las fuerzas que intervienen en las reacciones químicas y las fuerzas de cohesión dentro de los sólidos; del contexto de la *Query* 31 se desprende que también contaba entre los principios activos a las fuerzas eléctricas y magnéticas y a las «fuerzas vitales».

Aunque Newton no proporcionase un criterio general y operativo para distinguir los principios activos de los pasivos, lo que queda claro de su texto es que veía el origen o fundamento de los primeros en un principio superior y universal proveniente, en último término, de la voluntad de Dios. Ya hemos visto cómo en su primera carta a Bentley supone que los movimientos armónicos de los planetas sólo pueden explicarse por la voluntad de un agente inteligente. Y en la *Query* 31 concluye su discusión de los principios activos con las siguientes palabras: «Ahora bien, con la ayuda de estos principios, parece que se compusieron todas las cosas materiales a partir de las partículas duras y sólidas antes mencionadas; asociadas de diversas maneras, en la primera creación, por el consejo de un Agente inteligente. Pues fue El quien las creó, para ponerlas en orden [*Opera...*, t. IV, p. 261].

Se trata aquí de una explicación teleológica de la función de los principios activos dentro del plan divino. Los principios pasivos de la Mecánica deben subordinarse a los principios activos, subordinados a su vez al Plan General de Dios. Sólo bajo esa perspectiva jerárquica puede explicarse el Universo.

Ahora bien, la justificación teleológica de los principios activos no debe engañarnos acerca de cómo los consideraba Newton dentro de su obra científica. Para él, los principios activos eran tan *empíricos* como los pasivos. No se trataba de principios metafísicos. Los principios activos eran asequibles a la investigación experimental en el mismo sentido que los pasivos, sólo que un poco más difíciles de estudiar; y la pregunta por la finalidad de los principios activos, una cuestión tan empírica como la de averiguar las leyes del movimiento. En la primera versión del *Scholium Generale* es muy explícito: «Ciertamente las causas finales tienen su lugar en la filosofía natural; y ciertamente es legítimo preguntarse con qué fin se creó este mundo» [Hall y Hall, *Unpublished Scientific Papers...*, p. 363].

A modo de conclusión, parece plausible esbozar el siguiente cuadro general de la metodología newtoniana. A partir de las observaciones y experimentos realizados cuidadosamente en determi-

nados dominios de la Naturaleza se obtienen, en un primer paso inductivo, las regularidades de «bajo nivel» que Newton, en los *Principia*, llama «fenómenos», tales como las leyes de Kepler en astronomía o la ley de refracción de Snell en óptica. Mediante un segundo paso inductivo esas regularidades empíricas se incluyen en leyes mucho más generales, que pueden aplicarse a otros dominios distintos de los estudiados originariamente. Esto es posible gracias a las consideraciones de analogía: analogía y leyes generales van íntimamente asociadas. Un instrumento útil para la obtención de leyes generales es la formulación de «hipótesis» sin comprobación empírica. Las hipótesis deben tomarse sólo como *preguntas* («queries») que se hacen a la Naturaleza a fin de obtener una respuesta en determinado sentido. Pero las hipótesis, a ser posible, no deben figurar en un sistema científico concluso; sólo sirven para la preparación del sistema. Todas las leyes generales que aparecen en un sistema científico acabado deben haberse inferido inductivamente a partir de los «fenómenos», y por tanto, en definitiva, de observaciones y experimentos. Algunas de las leyes generales de un sistema científico (como las leyes del movimiento) son de carácter puramente *mecánico;* pero otras (como la ley de la gravitación o las leyes de las reacciones químicas) contienen nociones (gravitación, fermentación) que no pueden interpretarse mecánicamente. Esas nociones sólo resultan comprensibles si se admiten principios *teleológicos* en la investigación científica. Estos, a su vez, se justifican con conceptos *teológicos.*

La metodología newtoniana pretende ir, pues, «de abajo a arriba», y no al revés, como en Descartes o Leibniz. Y no se limita a principios causal-mecánicos («naturales») de explicación, sino que acude, cuando lo considera necesario, a principios teleológicos. No puede trazarse una línea tajante de separación entre ciencia y teología. En principio, la investigación de las propiedades y los propósitos de Dios también ha de ser posible mediante el método «inductivo»: «Y asimismo vale para Dios, discurrir sobre quien a partir de los Fenómenos, pertenece a la filosofía experimental. De los fenómenos se infieren las causas próxima de las cosas: de ellas, las causas superiores, hasta que llegamos a la causa suprema» [Hall y Hall, *Unpublished Scientific Papers...*, página 348].

Si quisiera resumirse la metodología newtoniana en un término-cliché, aunque esto es siempre peligroso, podría quizá decirse que se trata de un «inductivismo teleológico». Esta concepción metodológica de uno de los más grandes científicos de la historia probablemente disguste al paladar metodológico actual. Pero no deja de ser un dato interesante para el filósofo de la ciencia el de que alguien pueda llegar a resultados científicos universalmente aceptados propugnando al mismo tiempo una metodología y una filosofía que los cánones modernos consideran inaceptables.

3.3. LA EVOLUCION DE LA MECANICA NEWTONIANA

0. Introducción

En el capítulo 2.4 se introdujo el concepto de red teórica como estructura compuesta de estructuras más simples, los elementos teóricos, cada uno de los cuales, a su vez, está constituido por cuatro entidades: un núcleo, un dominio de aplicaciones, una comunidad científica y un intervalo histórico. En aquel capítulo y en los subsiguientes hicimos uso de la noción de red teórica y de sus componentes, especialmente el núcleo, para reconstruir aspectos sincrónicos de la ciencia, tanto en general como en ejemplos concretos. Pero ya apuntamos allí que el concepto de red teórica introducido podía servir también de base para reconstrucciones diacrónicas de la ciencia. En el presente capítulo tratamos de sustentar aquella afirmación nuevamente enfrentándonos a un ejemplo concreto. Para ello haremos uso sustancial no sólo del componente más formal de los elementos teóricos, el núcleo, sino también de los tres restantes componentes. Ellos nos permitirán definir una estructura aún más general que la de red teórica, a saber, la de *evolución teórica*, destinada a dar cuenta de los aspectos diacrónicos globales de lo que Kuhn llamó «ciencia normal». La idea básica es la siguiente. Consideramos que la evolución histórica de una teoría (mientras no rompe el marco de un «paradigma» en el sentido kuhniano) puede representarse como una sucesión de redes teóricas interconectadas de cierta manera –o, lo que es equivalente, pero quizá más gráfico, como una red teórica *cambiante*. En términos más metafóricos, visualizamos la evolución «normal» de una teoría como una red «viviente», que crece, cambia y se ramifica en diversas direcciones a través de la historia.

Como de costumbre, contrastaremos esta representación formal general ante un caso concreto. El ejemplo escogido tiene la virtud de que, además de ser uno de los mejor conocidos históricamente, puede servir como paradigmático para otros casos, al menos dentro de la física. Se trata de la evolución de (parte de) la mecánica clásica durante sus primeros cien años de existencia, es decir, desde Newton hasta el joven Laplace.

Aparte del objetivo concreto de reconstruir la evolución de una teoría física determinada, un objetivo más general que se quiere alcanzar con esta aplicación es el de mostrar que es plausible y fructífero usar nociones metateóricas generales para aclarar los aspectos más históricos y pragmáticos de la ciencia –una posibilidad que ha sido negada *a priori* por numerosos filósofos e historiadores contemporáneos.

El presente capítulo consta de dos partes. En la primera, se introducen las nociones generales requeridas; en la segunda, se reconstruye con ellas la evolución de la mecánica clásica.

1. Conceptos metateóricos diacrónicos

En un primer paso conviene introducir sin definición formal, como conceptos primitivos, dos nociones pragmáticas muy generales, para las que sólo damos unas breves aclaraciones intuitivas: a) la noción de *conjunto de ejemplos paradigmáticos reconocidos*, y b) las relaciones epistémicas de *admitir* y *creer* una proposición dada.

a) La noción de *ejemplo paradigmático* es esencial para nuestra metateoría diacrónica. Se retrotrae a Wittgenstein, aunque este autor no la usó en un contexto metacientífico. El primero en hacerlo así fue Kuhn. De él, la noción fue recogida y reelaborada formalmente por Sneed y Stegmüller. No repetiré aquí todo lo que estos autores han dicho sobre ejemplos paradigmáticos; sólo añadiré las siguientes precisiones. Se considera como ejemplo paradigmático para una teoría a veces un solo modelo parcial de esa teoría, a veces una *clase* de modelos parciales «homogéneos» (o sea, a veces un elemento de una «aplicación propuesta» en nuestro sentido, a veces la aplicación propuesta entera). Pero, en cualquier caso, la relación «...es un ejemplo paradigmático de...» debe elucidarse como una relación entre un individuo y una clase que contiene ese individuo. Además, debe haber cierta relación de «homogeneidad» o «semejanza», que podemos simbolizar por «$\approx$» entre el ejemplo paradigmático y los demás elementos de la clase que lo contiene. Por tanto, si tenemos un conjunto P de ejemplos paradigmáticos de una clase determinada C, P deberá ser un subconjunto de C y todos los elementos de C deberán estar entre sí en la relación $\approx$. La calificación adicional de «reconocido» se refiere implícitamente a un grupo de personas que reconocen los elementos de P como ejemplos paradigmáticos.

b) Dada cierta comunidad científica CC, y dada una proposición cualquiera p, el enunciado «CC admite p» deberá entenderse como una abreviación de: «la mayoría de miembros de CC considera que p es una proposición bien confirmada o corroborada por medio de procedimientos de contrastación típicamente usados por CC». Por otro lado, el enunciado «la persona H cree que p» significa meramente lo que significa en el lenguaje corriente; no hay referencia implícita ni a una comunidad científica ni a procedimientos específicos de contrastación.

En lo que sigue vamos a presuponer las nociones generales introducidas en 2.4. Recordamos que K es una variable para «núcleo»; A, una variable para «dominio de aplicaciones propuestas»; CC, una variable para «comunidad científica»; h una variable para «intervalo histórico»; T, una variable para «elemento teórico»; R, una variable para «red teórica», e $I(R)$ indica la totalidad de aplicaciones de la red. Pasamos a definir las nociones diacrónicas.

*D*1: R *precede inmediatamente* a R' si y sólo si:

(1) $R \neq R'$.
(2) $CC[R] = CC[R']$.
(3) $h[R] \lesssim \cdot h[R']$.
(4) No existe ningún R_i con $R_i \neq R$ y $R_i \neq R'$ tal que R_i satisfaga la condición (2) con respecto a R y R', y $h[R] \lesssim \cdot h[R_i] \lesssim \cdot h[R']$.

*D*2: E es una *evolución teórica* si y sólo si: E es una sucesión finita de redes teóricas tal que, para dos redes cualesquiera R_i, R_{i+1} pertenecientes a E se cumple:

(1) R_i precede inmediatamente a R_{i+1}.
(2) Para cualquier T^j_{i+1} perteneciente a R_{i+1} existe un T^j_i perteneciente a R_i tal que T^j_{i+1} es una especialización de T^j_i.

La condición (2) es esencial: para tener una verdadera evolución teórica es preciso que cada elemento teórico de una red construida en un momento dado esté conectado a través de la relación de especialización con algún elemento teórico (que puede ser idéntico) de la red teórica precedente. Nótese, sin embargo, que este concepto de evolución teórica no excluye la posibilidad de que una nueva red teórica contenga *menos* elementos teóricos que la red que la precede, en vez de tener más, como supondríamos en un caso normal. Incluso si esta posibilidad no parece muy verosímil, no hay nada en la historia real de la ciencia que nos haga excluir *a priori* esta posibilidad de una «regresión teórica». Por tanto, nuestro esquema formal debe admitirla.

Podemos hacer un análisis más detallado de las evoluciones teóricas si introducimos ciertas distinciones más finas. En particular, podemos considerar las relaciones epistémicas de *CC* con respecto a las aplicaciones propuestas de sus redes teóricas. Dada cierta red R durante el intervalo h, algunas de las aplicaciones contenidas en $I(R)$ serán admitidas por *CC* durante h como aplicaciones bien confirmadas. A este subconjunto de $I(R)$ lo llamaremos «el dominio *firme* o confirmado de aplicaciones», y lo simbolizaremos por $F(R)$. Ahora bien, *puede* que haya otros miembros de $I(R)$ que no estén bien confirmados como auténticas aplicaciones durante h, pero de los que un subgrupo dentro de *CC*, llamémosle CC_o[13], *supone* que resultarán ser verdaderas aplicaciones de la teoría. A este subconjunto complementario de $I(R)$ lo llamaremos «dominio *supuesto* de aplicaciones», y lo simbolizaremos como $S(R)$. Por el concepto mismo de $I(R)$ es una verdad analítica que: $F(R) \cup S(R) = I(R)$. E, idealmente, debería ser el caso que: $F(R) \cap S(R) = \varnothing$. Podría considerarse que el límite exacto entre $F(R)$ y $S(R)$ es lo que determina, como criterio de identidad, un intervalo histórico particular h: el paso de algunos elementos de $S(R)$ a $F(R)$ significaría un cambio de período histórico. (En este contexto no parece tener sentido hablar de una «historia universal» para todas las teorías en todas las disciplinas.)

Naturalmente, la noción de dominio confirmado de aplicaciones, $F(R)$ no debe ser tomada en un sentido absoluto. Es relativa a una evolución teórica dada e incluso a un período particular dentro de esa evolución teórica. Aplicaciones que se consideran firmes en un período dado, posteriormente pueden ser cuestionadas o incluso eliminadas del dominio de aplicaciones. Hay por lo menos dos ejemplos de esta situación en el caso de evolución teórica que formalizaremos después. Primero, las mareas fueron consideradas por Newton y sus colaboradores como aplicaciones firmes de la mecánica de partículas; el elemento teórico que desarrollaron para dar cuenta de las mareas fue la llamada «teoría del equilibrio para las mareas». Pero antes de que pasara un siglo, Laplace expulsó las mareas del dominio de la mecánica de partículas y las trató como una aplicación de la hidrodinámica con su «teoría ondulatoria». Este proceso tuvo lugar debido a que sólo una porción de los fenómenos ligados a las mareas pudo tomarse como una aplicación firme de la teoría del equilibrio de partículas; el resto mostró ser recalcitrante a un tratamiento con ese elemento teórico. Otro ejemplo es la teoría lunar (como elemento teórico de la mecánica). En efecto, aunque el sistema Tierra-Luna-Sol se consideró como aplicación firme de la mecánica newtoniana de partículas durante el primer período de la evolución teórica, posteriormente Euler llegó a ponerlo en duda. Con estas observaciones sobre la relatividad de la noción de aplicación firme

[13] CC_o puede ser una sola persona.

sólo se quiere recordar que el propósito del marco conceptual aquí introducido no es atacar los problemas irresueltos de la teoría de la confirmación.

*D*3: *E* es una *evolución teórica progresiva* si y sólo si:

(1) *E* es una evolución teórica.
(2) $\forall R_i, R_j \in E: i < j \rightarrow F(R_i) \subseteq F(R_j)$.

*D*4: *E* es una *evolución teórica perfecta* si y sólo si:

(1) *E* es una evolución teórica progresiva.
(2) $\forall R_i \in E \exists R_j \in E: i \leqslant j \wedge S(R_i) \subseteq F(R_j)$.

En *D*3 se capta la idea de que una evolución teórica es progresiva siempre que su dominio de aplicaciones firme aumente, o al menos no disminuya, con cada transición de una red teórica a la siguiente. Este concepto puede compararse a la noción de «programa de investigación progresivo» que introdujo Lakatos en su monografía «La falsación y la metodología de los programas de investigación científica».

En *D*4 se formaliza una especie de evolución teórica «utópica»: una evolución teórica que no sólo es progresiva, sino que procede de tal modo que cualquier aplicación propuesta *supuesta* en el momento presente resulte aplicación *confirmada* en algún momento del futuro. Probablemente no hay ningún ejemplo histórico de evolución teórica perfecta, y nunca habrá tal cosa; no obstante, algunas evoluciones teóricas puede que se aproximen más a este ideal que otras y, en este sentido, puede que este concepto sea históricamente relevante.

Aunque el trasfondo de las ideas aquí expuestas sobre el modo como formalizar la evolución «normal» de las teorías está inspirado en Kuhn, hasta este punto no se ha introducido ninguna noción específicamente kuhniana. Podemos acercarnos más a la metateoría de Kuhn sobre las evoluciones teóricas si introducimos el concepto de *paradigma* (para una evolución teórica).

Como ha mostrado Stegmüller en detalle en su *Theorienstrukturen und Theoriendynamik*, hay que distinguir dos componentes diversos en el concepto kuhniano de paradigma: uno de carácter formal-teórico (que correspondería a nuestra noción de núcleo) y otro de carácter pragmático-empírico (correspondiente a un dominio particular de aplicaciones propuestas). La noción primitiva de «ejemplo paradigmático reconocido» se necesita para reconstruir el segundo componente. Este deberá ser un subconjunto de $I(R)$, o sea, por la definición de $I(R)$ como reunión de todos los dominios A_i de los elementos teóricos de R, deberá ser un conjunto de clases de equivalencia «homogéneas» de modelos parciales. Dado un A_i, k_i será una variable para las clases

de aplicaciones homogéneas que constituyen A_i. Naturalmente, el mismo $I(R)$ en su totalidad estará constituido por tales clases de aplicaciones homogéneas.

D5: $\langle K_o; A_o \rangle$ es un *paradigma* para la evolución teórica E si y sólo si:

(1) K_o es un núcleo.
(2) $A_o \subseteq \mathscr{P}(M^o_{pp})$[14].
(3) Para cualquier red R perteneciente a E y cualquier $T_i = \langle K_i, A_i, CC_i, h_i \rangle$ perteneciente a R, se cumple que:
  (a) K_i es una especialización nuclear de K_o.
  (b) Para cada $k^j_i \in A_i$ existe un $P^o_{ij} \in A_o$ tal que CC_i reconoce P^o_{ij} como subconjunto paradigmático de k^j_i.

La condición *D5*-(3)(*b*) significa que cada dominio de aplicaciones propuestas en una evolución teórica puede dividirse en clases homogéneas de tal modo que cada una de estas clases tiene su propio subconjunto paradigmático de aplicaciones, el cual forma parte del paradigma «total» A_o.

D6: E es una *evolución teórica kuhniana* si y sólo si:

(1) E es una evolución teórica.
(2) Existen K_o, A_o, tales que $\langle K_o; A_o \rangle$ es un paradigma para E.

Para el lector que conozca el concepto de «Kuhn-theory» introducido por Sneed en su artículo «Philosophical Problems in the Empirical Science of Science» y retomado por Balzer y Sneed en su artículo conjunto de 1978, es pertinente hacer notar que la presente definición de «evolución teórica kuhniana» está inspirada en aquel concepto. No obstante, hay algunas diferencias notables entre ambas nociones. En primer lugar, al contrario de la noción sneediana, aquí no se exige que $\langle K_o, A_o \rangle$ sea él mismo un elemento teórico de ninguna red de E; en particular, no se presupone que $\langle K_o, A_o \rangle$ es históricamente el *primer* elemento teórico de la evolución. En el presente marco, esta cuestión se deja abierta, pues creo que se trata de un asunto históricamente contingente, que hay que decidir en cada caso concreto. En segundo lugar, no se postula (al contrario de Sneed) que *todo* A_o sea un paradigma para *cada* dominio de aplicaciones propuestas de E; sólo se asume que hay diversas clases «pequeñas» en A_o que son paradigmas para las correspondientes clases «pequeñas» en cada A_i; esto también parece concordar mejor con la noción intuitiva de

[14] M^o_{pp} es el conjunto de modelos parciales correspondientes a K_o.

ejemplo paradigmático (es decir, no es paradigmático para *cualquier cosa*).

Con esto hemos llegado al punto en que podemos contrastar el aparato formal introducido ante un ejemplo. Vamos a tratar de reconstruir la evolución de la mecánica newtoniana de partículas como una sucesión de redes teóricas.

2. La evolución de la mecánica newtoniana de partículas durante sus primeros cien años (de fines del siglo XVII a fines del siglo XVIII)

En lo que sigue presentamos un esbozo de las características principales de la historia de la mecánica desde sus comienzos hasta la época de Laplace. Naturalmente, sólo se trata de un esbozo; pero se pretende que concuerde con lo esencial de dicha evolución.

La mecánica *newtoniana* de partículas no debe ser identificada con la mecánica *clásica* de partículas. El primer concepto es más restrictivo que el segundo. Resulta de añadir al Segundo Principio de Newton, que es la ley fundamental de la mecánica clásica de partículas, el principio de acción y reacción para fuerzas centrales. Es decir, la mecánica newtoniana de partículas es la red teórica cuyo elemento básico viene dado por el predicado que en el capítulo 2.5 hemos llamado «dinámica de acción y reacción estricta» (junto con sus especializaciones y ligaduras correspondientes). La razón por la que nos restringimos en nuestro análisis a esta parte de la red de la mecánica clásica de partículas es que no parece que durante el siglo estudiado se considerara la posibilidad de sistemas más generales que los subsumibles bajo el predicado «*DARE*».

La evolución de la mecánica newtoniana *de partículas* estuvo sin duda conectada con la evolución de otras teorías mecánicas (hidrodinámica, acústica, mecánica del sólido rígido). Por lo que respecta al dominio de aplicaciones, en algunos casos se da un desplazamiento gradual de la mecánica de partículas a otras teorías mecánicas; por ejemplo, los movimientos de partículas en medios resistentes gradualmente pasaron del dominio de la mecánica de partículas al de la hidrodinámica. Sin embargo, aquí no vamos a considerar estas interconexiones entre las diversas evoluciones teóricas. Ello formaría parte de un estudio mucho más amplio, el de la evolución del marco teórico de la mecánica entera. De momento, el procedimiento metodológicamente recomendable es tratar de aislar las diversas evoluciones teóricas (es decir, las evoluciones de cada una de las teorías que forman el marco) y reconstruir la estructura de cada una de ellas –lo cual, en sí mismo no es una tarea trivial– *antes* de empezar a investigar las relaciones mutuas. En consecuencia, en el presente caso, se reconstruirá la evolución de la mecánica newtoniana de partículas como una única estructura en evolución. Naturalmente, esto es una

idealización; pero se espera que la idealización no sea demasiado simplista.

Antes de pasar a la representación formal de la red newtoniana en evolución vamos a dar cuenta informalmente de su contenido. Esta exposición se basa en la literatura historiográfica relevante, que se menciona al final del capítulo bajo el rubro de «Fuentes históricas». Así, pues, las aserciones históricas concretas que se hacen en la presente reconstrucción se basan en el supuesto de que las descripciones e interpretaciones de los historiadores mencionados son, por lo general, correctas. En algunos casos especialmente significativos se mencionará la obra historiográfica particular de la que se han extraído los datos.

Se ha dividido la evolución de la mecánica newtoniana de partículas desde fines del XVII hasta fines del XVIII en cuatro períodos (=intervalos históricos) principales; por supuesto, esta división es de carácter esquemático, pero creemos que tiene buen apoyo histórico. El primer período es el *revolucionario*. Estuvo limitado a Newton y a sus colaboradores más cercanos en Inglaterra (Halley, Flamsteed, Clarke, Taylor). Durante el segundo período, el formalismo de Newton fue adoptado, refinado y ulteriormente aplicado sobre todo por el *grupo de Basilea* (Herman, Euler, los Bernoullis). El tercer período es el de los «*geomètres*» *franceses* (Clairaut, D'Alembert, Lagrange y también Euler); ellos hicieron grandes esfuerzos por resolver difíciles problemas de varios cuerpos en la mecánica celeste –no siempre con éxito. El cuarto y último período considerado se caracteriza principalmente por dos aspectos completamente diversos: las impresionantes soluciones que obtuvo *Laplace* de algunos importantes y difíciles problemas de la mecánica celeste y los éxitos de *Coulomb* al tratar las fuerzas electrostáticas y magnetostáticas por analogía con la fuerza de gravitación.

Usando el aparato conceptual introducido anteriormente vamos a describir las características básicas de cada período con un poco más de detalle.

En primer lugar, nótese que, desde un punto de vista puramente cronológico, los cuatro períodos distinguidos se traslapan parcialmente. Esto apoya la concepción general de que, para el tratamiento de cuestiones diacrónicas en la ciencia, la relación de orden puramente topológica $\precsim$ es más adecuada que la relación numérica $<$ entre años.

Durante el *primer período* se construye en el Primer Libro de los *Principia* de Newton el núcleo básico de la red (los conceptos y las leyes fundamentales de Newton). Este núcleo se especializa inmediatamente a un elemento teórico que sólo considera fuerzas dependientes de la distancia (nuestro predicado *DDD* del cap. 2.5), el cual, a su vez, se especializa a fuerzas gravitacionales (nuestro *DGP*) en el Tercer Libro de los *Principia*. La construcción y aplicación exitosa del

elemento gravitacional fue el principal resultado de este período revolucionario. Casi todas las aplicaciones firmes de este período son aplicaciones de este elemento teórico[15]. Esta es la razón por la que en este período la teoría newtoniana casi se identificó con la teoría de la gravitación (la cual es, desde nuestro punto de vista actual, solamente *un* elemento de la red teórica entera). Ciertamente, Newton pensó que su red teórica original podía expandirse insertándole elementos teóricos no-gravitacionales con el fin de dar cuenta de los fenómenos ópticos, químicos, eléctricos y magnéticos; pero en el momento eso sólo era una vaga promesa sin base firme.

Las aplicaciones más sorprendentemente confirmadas que se obtuvieron durante el primer período pertenecen a la mecánica celeste: todos los planetas conocidos hasta el momento; algunos cometas, el sistema Tierra-Luna; Júpiter con sus satélites; Saturno con los suyos. Pero no debemos olvidar las aplicaciones terrestres del elemento gravitacional, que también le dieron buen apoyo a la teoría: principalmente, la caída libre cerca de la superficie terrestre, los péndulos simples y algunos fenómenos de mareas (aunque no todos ellos –véase Defant, «Flutwellen und Gezeiten des Wassers»).

Además de estas aplicaciones firmes del elemento gravitacional, son dignos de notarse dos casos interesantes que fueron aplicaciones empíricas directas del núcleo básico (de las leyes fundamentales): primero, la aplicación paradigmática de la teoría rival de Newton, la mecánica cartesiana, a saber, el choque de partículas, fue derivada directa y naturalmente del núcleo básico newtoniano sin necesidad de ulteriores especializaciones (cf. Newton, *Principia*, Corolarios III, IV, VI, y también Home, *op. cit.*); en segundo lugar, Taylor usó el núcleo newtoniano para atacar el problema de la cuerda vibrante, obteniendo un éxito parcial. «Exito parcial» significa aquí que Taylor sólo pudo explicar los casos más simples de la cuerda vibrante (cf. Szabó, «Die Geschichte der Theorie der schwingenden Saite»; Truesdell, *Essays in the History of Mechanics*). Newton también atacó con cierto éxito el problema del movimiento de una partícula en un medio resistente como el aire, empleando para ello fuerzas dependientes de la velocidad.

Por lo que respecta a las aplicaciones supuestas en este estadio de la teoría, Newton y sus colaboradores supusieron explícitamente que investigaciones futuras lograrían mostrar que todos los fenómenos celestes, incluyendo los sistemas de varios cuerpos y todos los ejemplos de cometas, encajan en el elemento gravitacional (cf. Truesdell, *op. cit.*). Además, Newton creía que en el futuro se podrían desarrollar nuevas especializaciones que explicarían una gran varie-

[15] Al menos hay una excepción a esto: el choque elástico (cf. Home, «The Third Law in Newton's Mechanics»).

dad de fenómenos: el comportamiento de los rayos de luz, la cohesión química, la electricidad y el magnetismo. El mismo hizo serios esfuerzos en este sentido para la primera clase de aplicaciones (cf. Newton, *Opticks*; Truesdell, *op. cit.*). Las otras tres clases de aplicaciones las presentó simplemente como «queries» (=interrogantes) a la Naturaleza en su *Opticks*.

El *segundo período* se caracteriza esencialmente por la elaboración detallada de la red newtoniana por parte de la Escuela de Basilea. Se desarrolló con gran éxito una nueva línea de investigación: fuerzas dependientes de la distancia, pero no-gravitacionales, destinadas a dar cuenta de sistemas oscilatorios. Desde nuestro punto de vista, esto significa la adición exitosa de una nueva especialización a la red anterior. La culminación de esta línea de trabajo teórico puede verse en la elaborada teoría de Euler sobre las diversas clases de osciladores armónicos (cf. Truesdell, *op. cit.*).

Una línea distinta dentro de este período es la representada por los sucesivos esfuerzos realizados por D'Alembert, Euler y Daniel Bernoulli para llegar a un tratamiento completamente satisfactorio de la cuerda vibrante (discreta). Lagrange también supuso durante este período que el caso más general de la cuerda continua podía tratarse dentro del marco proporcionado por la mecánica de partículas. Pero más tarde las críticas de D'Alembert persuadieron a los newtonianos de que la cuerda continua no pertenece propiamente al dominio de aplicaciones de la mecánica de *partículas*. Así fue como esta aplicación particular acabó por ser expulsada de esta red teórica (cf. Truesdell, *op. cit.*).

Durante el *tercer período*, Clairaut, Euler, D'Alembert y Lagrange dedicaron sus mayores esfuerzos a la eliminación de todas las divergencias detectadas entre núcleo y aplicación en la mecánica celeste. Sin embargo, no siempre tuvieron éxito en su empresa, dado que los problemas con los que se enfrentaron eran matemáticamente muy difíciles. Ciertamente, una espléndida confirmación del elemento gravitacional para los fenómenos celestes fue el tratamiento dado por Clairaut al complicado sistema compuesto por el cometa Halley, Júpiter, Saturno y el Sol, lo cual le permitió hacer una predicción muy exacta de la órbita del cometa Halley (cf. Pannekoek, *A History of Astronomy*). No obstante, otros problemas de la mecánica celeste resultaron ser más recalcitrantes. En particular, el sistema Tierra-Luna-Sol llevó a algunos investigadores de la comunidad a poner en duda la adecuación del elemento gravitacional y buscar algo nuevo: por un tiempo, Clairaut supuso que había que añadir un factor $\frac{1}{r^4}$ a la ley de gravitación para dar cuenta de los datos recalcitrantes que se habían obtenido al medir las órbitas respectivas; y Euler llegó incluso al punto de cuestionar que el sistema Tierra-Luna-Sol pudiera hacerse

encajar en nada parecido a la ley de gravitación (cf. Pannekoek, *op. cit.*). A fin de simplificar la representación formal de la evolución teórica que se ofrece más abajo, en ella no tendremos en cuenta estos dos interesantes desarrollos. Sin embargo, no habría ninguna dificultad especial en formalizarlos con el presente aparato conceptual. La propuesta de Clairaut puede interpretarse en nuestros términos como un cambio sustancial en la red teórica existente, consistente en reemplazar el elemento gravitacional aceptado por otro nuevo; es interesante notar que la propuesta de Clairaut no fue considerada por la comunidad newtoniana como una desviación revolucionaria de la mecánica newtoniana, lo cual, por cierto, muestra que la ley de gravitación no debe considerarse como formando parte del núcleo básico de la teoría. En cuanto a Euler, su actitud negativa hacia el sistema Tierra-Luna-Sol significa que estaba dispuesto a sacar ese sistema del dominio de aplicaciones del elemento gravitacional de la red; tampoco esto se sintió como una propuesta revolucionaria en contra de la mecánica newtoniana de partículas, sino más bien como la esperanza de que se lograría inventar una especialización distinta apta para resolver el problema.

Un resultado completamente distinto es el que corresponde a la «otra línea» de la evolución teórica: el éxito de Euler al poder añadir ciertas clases de sistemas de varios cuerpos al dominio de aplicaciones de las fuerzas dependientes de la distancia introducidas para los sistemas oscilatorios (cf. Truesdell, *op. cit*).

El *cuarto período* es la fase culminante de la evolución teórica considerada. Ella se debe sobre todo a la obra de Laplace, pero también, en menor medida y en un sentido diferente, a los hallazgos de Coulomb. La impresionante habilidad matemática de Laplace le permitió eliminar prácticamente todas las dificultades de la mecánica clásica que habían preocupado gravemente a la generación anterior. Uno de los resultados más sobresalientes de Laplace fue su confirmación de que, a pesar de todo, el sistema Júpiter-Saturno-Sol y el sistema Tierra-Luna-Sol efectivamente encajan en el elemento gravitacional con la exactitud requerida. Un aspecto negativo de la obra de Laplace fue su expulsión de las mareas del dominio de aplicación de la mecánica de partículas.

Otros ejemplos de adiciones exitosas al dominio del elemento gravitacional fueron los cálculos de Olbers de las órbitas de todos los cometas observados (cf. Pannekoek, *op. cit.*) y el descubrimiento de Urano por Herschel por la misma época.

En cuanto a las investigaciones de Coulomb sobre los fenómenos electrostáticos y magnetostáticos tienen el siguiente significado para nuestra metateoría. Lo que Coulomb hizo fue introducir dos nuevas especializaciones de la red teórica básica: especializó el elemento teórico de fuerzas dependientes de la distancia en las fuerzas elestrostáticas y magnetostáticas que obedecen a la llamada ley de Coulomb,

y además confirmó la aplicabilidad de estos nuevos elementos teóricos al caso de sistemas de dos partículas.

Después de este breve e informal resumen de la evolución teórica, esperamos que su reconstrucción formal aparezca más clara. Primero hay que explicar el simbolismo especial utilizado. Los pares ordenados representarán elementos teóricos de cada una de las redes sucesivas; para simplificar la exposición, se omitirán referencias explícitas al período histórico y a la comunidad científica correspondiente a cada elemento teórico. En la representación gráfica las líneas que conectan elementos teóricos diferentes expresan especializaciones (siendo los elementos teóricos inferiores especializaciones de los superiores).

Denotaremos los núcleos de los elementos teóricos de cada red con los predicados conjuntistas que determinan los conjuntos respectivos de modelos actuales. Es decir, usaremos para los núcleos K_i la misma notación que para sus conjuntos M_i correspondientes. Esto significa, en particular, que no consideraremos las condiciones de ligadura por separado. Así, por ejemplo, «*DGP*» («dinámica gravitacional de partículas») se referirá al núcleo gravitacional entero, o sea, la ley de gravitación junto con las ligaduras correspondientes. En esta reconstrucción de la evolución de la mecánica no viene al caso distinguir las leyes propiamente dichas de las ligaduras, por lo que podemos simplificar la notación.

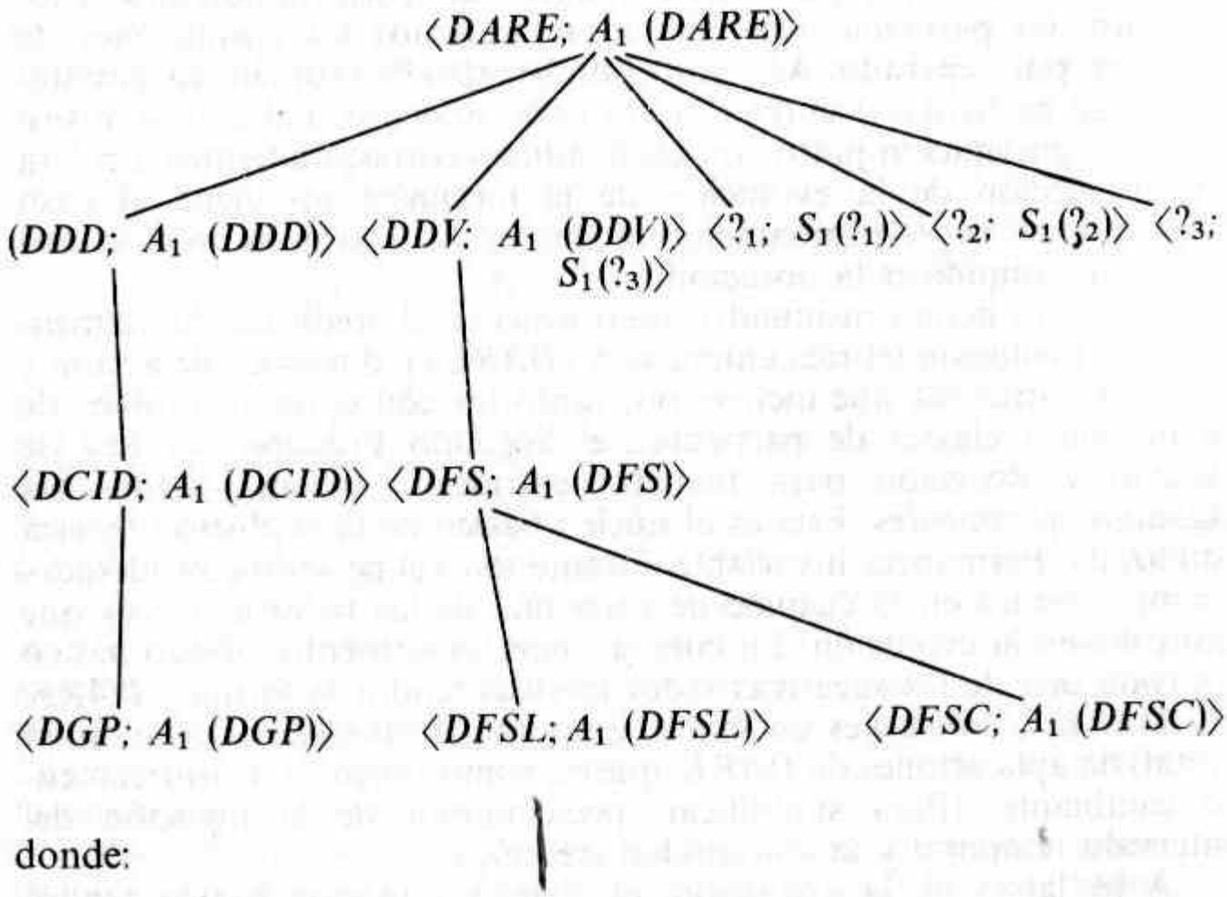
Como ya hemos insinuado anteriormente, el predicado fundamental de la evolución teórica entera será «*DARE*» («dinámica de acción y reacción estricta»), que incluye por tanto los conceptos primitivos de la mecánica clásica de partículas, el Segundo Principio, la Ley de Acción y Reacción para fuerzas centrales, y además todas las ligaduras pertinentes. Este es el núcleo básico de la evolución teórica analizada. Permanece invariable durante todo el período considerado. Siempre estará en la cúspide de cada una de las redes arbóreas que constituyen la evolución. En consecuencia, el elemento teórico básico en cada una de las sucesivas redes teóricas tendrá la forma: $\langle DARE; A_i\ (DARE)\rangle$, donde i es un índice temporal: $A_i\ (DARE)$ es el dominio (total) de aplicaciones de *DARE*, que es, como sabemos, históricamente cambiante. (Para simplificar, prescindimos de la mención del intervalo temporal y la comunidad científica.)

A lo largo de la evolución, el elemento teórico básico tendrá diversas aplicaciones, y éstas a su vez tendrán sus especializaciones, y así sucesivamente. Cada uno de los elementos especializados tendrá la forma $\langle X; A_i(X)\rangle$, donde X es siempre el predicado conjuntista que define el conjunto de modelos correspondiente. F_i y S_i son dos operadores de cada elemento teórico que hacen corresponder a cada núcleo (a sus leyes) las aplicaciones firmes y las supuestas, respectivamente. Así, por ejemplo, «Tierra-Luna$\in F_1\ (DGP)$» significa que el sistema Tierra-Luna pertenece al dominio de aplicaciones firmes de la

dinámica gravitacional de partículas durante el primer período de la evolución histórica. Por otra parte, «sistemas planetarios de varios cuerpos$\in S_1$ (*DGP*)» significa que durante el primer período sólo se suponía, sin confirmación, que los sistemas planetarios de varios cuerpos pertenecían al dominio de aplicaciones de la dinámica gravitacional de partículas. Recuérdese, además, que los rótulos «Tierra-Luna» o «mareas» o «sistemas de varios cuerpos» se refieren formalmente no a aplicaciones singulares, sino a conjuntos de aplicaciones («clases homogéneas»); en algunos casos, como en el de «Tierra-Luna», estos conjuntos pueden ser unitarios, pero esto no es el caso normal.

Pasamos finalmente a la representación formal de la evolución.

Primer período: $h_1 = [168...\text{-}172...]$



donde:

«*DARE*» = «dinámica de acción y reacción estricta» (leyes fundamentales de Newton».

«*DDD*» = «dinámica dependiente de la distancia» (en esta evolución este predicado se tomó como especialización de *DARE*).

«*DDV*» = «dinámica dependiente de la velocidad».

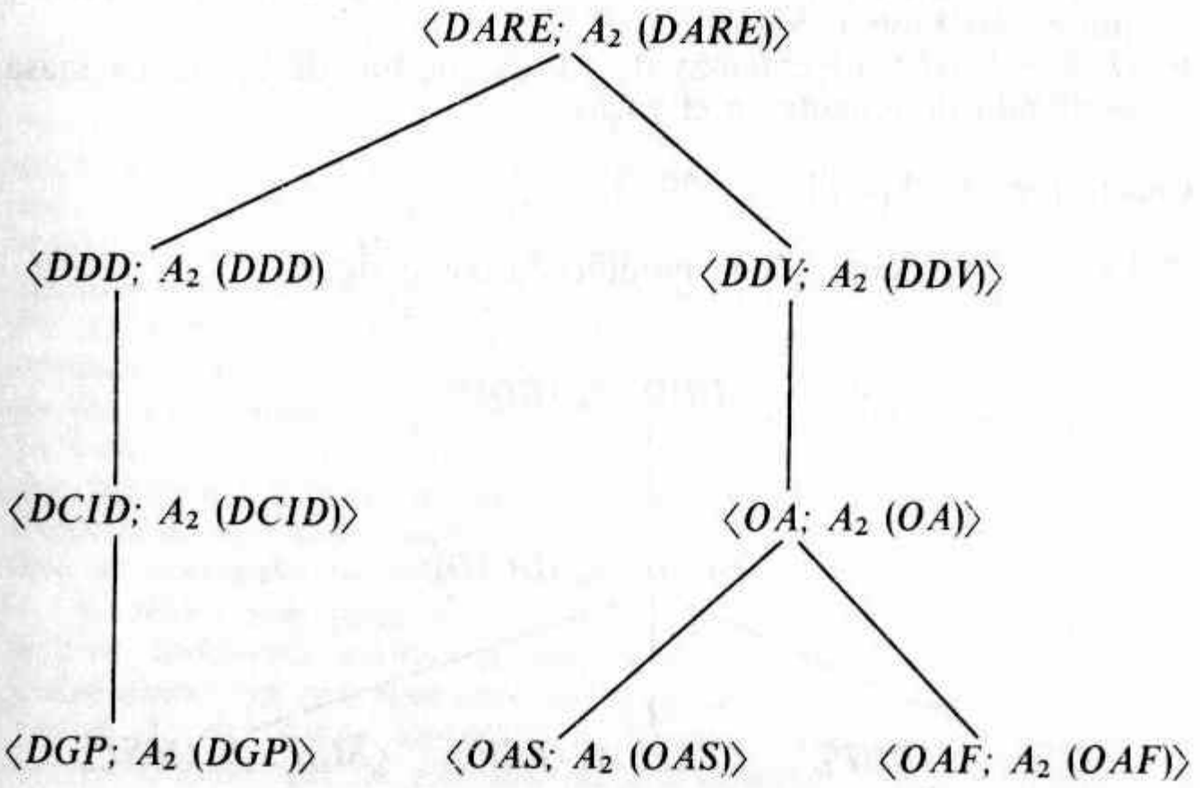
«*DFS*» = «dinámica de fricción simple».

«*DFSL*» = «dinámica de fricción simple lineal» (el exponente de la velocidad de la que depende la fuerza de fricción se toma $n=1$).

«*DFSC*» = «dinámica de fricción simple cuadrática» (el exponente de la velocidad se toma $n=2$).
«*DCID*» = «dinámica cuadrático-inversa de la distancia».
«*DGP*» = «dinámica gravitacional de partículas».
Y $A_1\ (DARE) = A_1(DDD) \cup A_1(DDV) \cup S_1(?_1) \cup S_1(?_2) \cup S_1(?_3)$, con
$F_1(DARE) = \{$choques de partículas; un subconjunto de casos de cuerdas vibrantes$\} \cup F_1(DGP) \cup F_1(DFS)$, siendo
$F_1(DDD) = F_1(DCID) = F_1(DGP) = \{$todos los sistemas planeta-Sol conocidos en el momento; Júpiter-satélites; Saturno-satélites; Tierra-Luna; algún cometa-Sol; caída libre cerca de la superficie terrestre; proyectiles sin resistencia del aire; péndulos; un subconjunto de los fenómenos de mareas$\}$.
$S_1(DDD) = S_1(DCID) = S_1(DGP) = \{$fenómenos de mareas no explicados; todos los sistemas planetarios de varios cuerpos$\}$.
$F_1(DDV) = F_1(DFS) = F_1(DFSL) \cup F_1(DFSC) = \{$algunos proyectiles a través del aire$\}$.
$S_1(DDV) = \{$todos los proyectiles a través de un fluido$\}$.
$S_1(?_1) = \{$fenómenos ópticos$\}$.
$S_1(?_2) = \{$cohesión química$\}$.
$S_1(?_3) = \{$fenómenos eléctricos y magnéticos$\}$.

Segundo período: $h_2 = [172...\text{-}174...]$.

La red anterior queda modificada de la siguiente manera:



donde

«OA» = «oscilador armónico»,
«OAS» = «oscilador armónico simple» (sistemas oscilando libremente),
«OAF» = «oscilador armónico forzado» o «frenado» (sistemas oscilatorios, en los que las fuerzas no sólo dependen de la posición, sino también de la velocidad impresa).

Los principales cambios que se dan durante este período en el dominio general de aplicaciones, $A_2(DARE)$, son como siguen:

$F_2(DARE) = F_1(DARE) \cup \{$todos los casos de cuerdas vibrantes discretas$\} \cup A_2^*(DGP) \cup F_2(OAS) \cup F_2(OAF)$, donde

$A_2^*(DGP) \subset S_1(DGP)$,

$F_2(OAS) = \{$partículas aisladas oscilando libremente en el vacío$\}$,

$F_2(OAF) = \{$partículas aisladas oscilando forzadamente en el vacío$\}$.

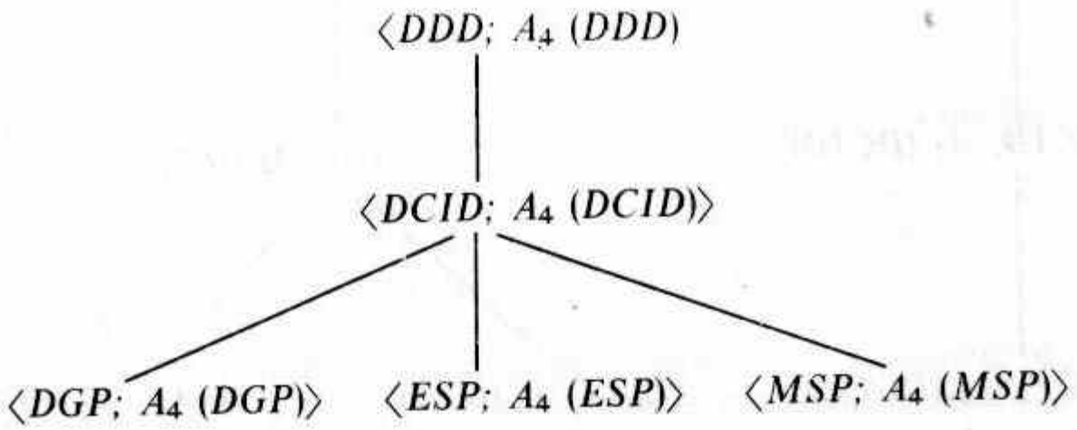
Tercer período: $h_3 = [174...\text{-}177...]$.

Los cambios producidos en la red precedente sólo afectan el dominio de aplicaciones, no la estructura nuclear (teórica).

$F_3(DGP) = F_2(DGP) \cup \{$el sistema formado por el cometa de Halley junto con Júpiter, Saturno y el Sol$\}$.
$F_3(OAS) = F_2(OAS) \cup \{$sistemas de varios cuerpos de la misma masa oscilando libremente en el vacío$\}$.

Cuarto período: $h_4 = [177...\text{-}179...]$.

La red precedente queda modificada como sigue:



donde

«*ESP*» = «electrostática de partículas» (ley de Coulomb para cargas electrostáticas),
«*MSP*» = «magnetostática de partículas» (ley de Coulomb para polos magnéticos).

En cuanto al dominio de aplicaciones, tenemos los siguientes cambios:

$F_4(DGP) = F_3(DGP) \cup \{$Tierra-Luna-Sol; Júpiter-Saturno-Sol; Urano-Sol; cada uno de los cometas-Sol$\} - \{$mareas$\}$,
$F_4(ESP) = \{$sistemas de dos partículas cargadas eléctricamente$\}$,
$F_4(MSP) = \{$sistemas de dos partículas como polos magnéticos$\}$, y
$F_4(ESP) \cup F_4(MSP) \subset S_1(?_3)$.

Acerca del contenido de esta última red hay que observar que la expulsión de las mareas del dominio de aplicaciones de la mecánica newtoniana de partículas fue debida a Laplace, quien propuso a cambio que estos fenómenos fueran tratados por una teoría hidrodinámica. Sin embargo, es de notar que no todos los miembros de la comunidad newtoniana aceptaron esta revisión, y durante muchos años algunos newtonianos siguieron tratando de subsumir las mareas bajo el esquema teórico de la mecánica de partículas (cf. Defant, *op. cit.*). Así, pues, el tratamiento de las mareas continuó siendo una cuestión disputada.

Es fácil comprobar que cada una de las redes consideradas en esta evolución es una red arbórea en el sentido preciso en que esta noción fue introducida en el capítulo 2.4, ya que, entre otras razones, puede mostrarse que cada uno de los núcleos que aparece en cada una de ellas es, por lo menos, una especialización nuclear de *DARE*. Asimismo, la sucesión de las cuatro redes consideradas constituye una evolución teórica en el sentido preciso de la definición *D*2 de este capítulo; además, si prescindimos de momento de la cuestión disputada de las mareas y la decisión de Laplace, podemos considerar esta evolución teórica como (aproximadamente) progresiva en el sentido de *D*3. En cambio, lo que ciertamente no es es una evolución teórica perfecta, dado que muchas de las aplicaciones propuestas y supuestas por Newton a la larga demostraron no ser aplicaciones firmes (esto es especialmente válido para los fenómenos ópticos y químicos por los que se interesaba Newton).

Si ahora nos preguntamos por el paradigma que pueda haber guiado toda esta evolución, resulta más difícil hacer una elección indisputable, ya que esta cuestión depende de una interpretación no-trivial de los datos históricos a nuestra disposición. Parece casi incuestionable que la porción teórica (nuclear) del paradigma que

guió la evolución estuvo constituida por las leyes fundamentales de Newton (*DARE*) y sólo por éstas. La cuestión es más difícil de decidir para el caso de las aplicaciones paradigmáticas. Existen algunas aplicaciones que fueron indudablemente paradigmáticas para los desarrollos subsiguientes: los casos planetarios que Newton mismo logró subsumir con éxito bajo las leyes de la teoría. También se podrían añadir los choques y los cuerpos en caída libre. No obstante, un punto más interesante es que, si no estoy interpretando mal los datos, parece que *no sólo* fueron las aplicaciones explícita y exitosamente tratadas por Newton las que jugaron un papel paradigmático para el trabajo subsiguiente en la evolución teórica: probablemente también tuvieron un carácter paradigmático algunas aplicaciones que Newton se propuso tratar, pero que no pudo resolver con éxito (como sistemas de tres cuerpos en el sistema planetario) e incluso aplicaciones en las que Newton no parecía estar particularmente interesado (partículas oscilando libremente). Por otra parte, está claro que aplicaciones más complicadas que aparecieron a lo largo de la evolución teórica no deberían ser consideradas como paradigmáticas (para *esta* evolución teórica); ejemplos: las mareas, sistemas complicados de varios cuerpos, oscilaciones forzadas, fenómenos eléctricos y magnéticos.

Las afirmaciones que se acaban de hacer son, naturalmente, aserciones de contenido sustancialmente histórico; podrían ser refutadas por investigaciones historiográficas futuras. Ahora bien, la cuestión parece ser de tal naturaleza que es difícil de resolver definitivamente, porque no sólo depende de estudios históricos detallados, sino también de la elucidación de los criterios propuestos para que algo deba considerarse como ejemplo paradigmático, problema que aún no ha sido atacado sistemáticamente. En cualquier caso, y suponiendo que la interpretación precedente de los datos no es errónea, el punto interesante para la *filosofía* de la ciencia que emerge de esta discusión es la hipótesis de que algunas aplicaciones pudieron jugar el papel de ejemplos paradigmáticos para la evolución de la mecánica newtoniana sin haber sido tratadas con éxito (o ni siquiera propuestas) por el propio Newton. Si esta interpretación resultara ser correcta, entonces habría que adoptar un concepto más flexible y «liberal» de aplicación paradigmática que el ofrecido por Kuhn y precisado por Stegmüller en sus escritos respectivos sobre el tema. Las aplicaciones paradigmáticas no serían necesariamente sólo aquellas que fueron explicadas con éxito por el fundador de la teoría en el estadio inicial de la evolución. Asimismo el primer elemento de las redes teóricas de una evolución no tendría por qué coincidir con el paradigma de la evolución (aunque podría hacerlo). Esto último contradiría lo supuesto por Balzer y Sneed al introducir su concepto de «Kuhn-theory» antes mencionado. El primer elemento teórico invariable de cada red de la evolución por un lado y el paradigma de la misma por el otro serían

entonces dos nociones conceptualmente distintas e independientes, y la cuestión de si están o no estrechamente conectados entre sí sería una cuestión históricamente contingente. El análisis de la evolución teórica que hemos presentado en este capítulo parece apoyar estas hipótesis.

El paradigma de la evolución de la mecánica newtoniana de partículas parece ser formalmente el siguiente:

$$P[MNP]=\langle DARE;\ \{\text{choques}\}\cup F_1(DGP)\cup S_1^*(DGP)\cup F_2(OAS)\rangle,$$

donde $S_1^*(DGP)=\{$ciertos sistemas de varios cuerpos en el sistema solar$\}$.

Vemos que este paradigma *no* coincide con el elemento teórico básico común a todas las redes de la evolución considerada. Por lo tanto, la tarea de determinar el paradigma debería considerarse como conceptualmente independiente de la tarea de determinar el elemento teórico básico.

Por otro lado, parece fundado admitir que P realmente satisface las condiciones de $D5$ para poder ser un paradigma para la evolución de la mecánica newtoniana de partículas. La única condición que podría presentar problemas es $D5$-(3) (*b*). Pero ante los datos me parece admisible la afirmación de que a cada aplicación hallada en esta evolución teórica le corresponde algún subconjunto del dominio de aplicaciones de $P[MNP]$ que efectivamente jugó para ella el papel de ejemplo paradigmático. Claro que ésta es una cuestión que depende fuertemente de los hechos históricos y para la que de momento sólo se pueden ofrecer criterios intuitivos. Para concluir, digamos sólo que parece muy plausible describir la evolución de la mecánica newtoniana de partículas a través de los cuatro períodos considerados como una evolución teórica kuhniana en el sentido de $D6$.

Fuentes históricas

Defant, A.: «Flutwellen und Gezeiten des Wassers», en *Handbuch der Physik*, t. 78, *Geophysik* II, 1957.

Home, R. W.: «The Third Law in Newton's Mechanics», en *The British Journal for the History of Science*, 7 (1968-69).

Hund, F.: *Geschichte der physikalischen Begriffe*, Mannheim, 1972.

Newton, I.: *Philosophiae Naturalis Principia Mathematica* (comp. por A. Koyré e I. B. Cohen), Cambridge, Mass., 1972.

—: *Opticks. Opera quae extant omnia*, t. IV (comp. por S. Horsley), Stuttgart, 1964.

Pannekoek, A.: *A History of Astronomy*, Nueva York, 1961.

Reichen, Ch.-A.: *A History of Astronomy*, Londres, 1968.

Shapley, H., y H. Howarth: *A Source Book in Astronomy*, Nueva York, 1929.
Szabó, I.: «Die Geschichte der Theorie der schwingenden Saite», *Humanismus und Technik*, 15, 1971.
Truesdell, C. A.: *Essays in the History of Mechanics*, Berlín-Heidelberg-Nueva York, 1968.
Truesdell, C. A.: «History of Classical Mechanics. Part I, to 1800», *Die Naturwissenschaften*, 1976.

3.4. LA EPISTEMOLOGIA FISIOLOGICA DE HERMANN VON HELMHOLTZ

0. Introducción

La filosofía del positivismo lógico, que tanto ha influido en la epistemología y la filosofía de la ciencia de las últimas décadas, no surgió en un vacío histórico: todos sabemos que tiene sus ancestros. Entre sus antecedentes históricos, los que más se suelen citar son Hume, Mach, el nacimiento de la lógica moderna, el primer Wittgenstein y las grandes revoluciones en física. No pretendo aquí negar la importancia de estas influencias; sin embargo, me parece que no cubren toda la prehistoria del positivismo lógico. Con harta frecuencia se ignora otro elemento significativo: la influencia de los grandes desarrollos en la fisiología de los sentidos del siglo XIX sobre los modos en que el Círculo de Viena y grupos análogos formularon sus problemas y concepciones. Esta influencia puede trazarse claramente en el fenomenalismo de los años veinte, cuya expresión más conspicua es la primera obra mayor de Carnap, *Der logische Aufbau der Welt*. Es probable que un lector actual que lea esta obra se sienta sorprendido ante la gran cantidad de esfuerzo y trabajo de detalle que Carnap dedica a problemas, conceptos y construcciones hipotéticas de la fisiología y psicología de la percepción sensorial –especialmente si se recuerda la evolución ulterior de Carnap. Considero que esta clase de fenomenalismo no es sino una teoría del conocimiento elaborada en términos psicofisiológicos. Además, se trata de una teoría del conocimiento cuya metodología demuestra contener elementos kantianos. El *Aufbau* de Carnap le debe a la tradición kantiana más de lo que probablemente le hubiera gustado reconocer a su propio autor.

Naturalmente, los intentos de Carnap y otros parecidos característicos de la primera fase del positivismo lógico pueden retrotraerse a Mach. El *Análisis de las sensaciones* de Mach es un curioso cóctel de planteamientos fisiológicos y epistemológicos. Ahora bien, la teoría machiana del conocimiento, a su vez, proviene de una tradición que subrayaba las consecuencias epistemológicas de la moderna fisiología sensorial y sus conexiones con Kant. Esta tradición fue particularmente fuerte en los países de lengua alemana y reveló dos características que no se hallaban presentes en la teoría del conocimiento del empirismo en sentido estricto: primera, una conexión estrecha con investigaciones especiales, experimentales, dentro del campo de la percepción sensorial; segunda, un análisis detallado y crítico de la epistemología kantiana. El «empiriocriticismo» de Mach no puede comprenderse bien en su sentido si no se tienen en cuenta sus propias investigaciones psicofisiológicas, así como las de científicos tales como Hering, Fechner y, sobre todo, Helmholtz. En este sentido, Mach pertenece a un desarrollo bien definido en la historia de las ideas: a un programa epistemológico para fundamentar todos los tipos de conocimiento humano genuino en un análisis filosófico cuidadoso de resultados empíricos concretos obtenidos en el estudio de la fisiología de los sentidos. Este desarrollo histórico tiene su precursor en Kant, sus pioneros en Johannes Müller y sus colaboradores, su prosecución radical en el empiriocriticismo de Mach y Avenarius, y su culminación en el fenomenalismo de Carnap; el pivote central de esta evolución es Hermann von Helmholtz. Tanto desde un punto de vista histórico como metodológico, a Helmholtz se le puede considerar como la figura crucial en la evolución de este programa encaminado a basar la teoría del conocimiento en la fisiología de los sentidos. En lo que sigue vamos a discutir los aspectos esenciales de la contribución de Helmholtz a esta tradición.

1. La transformación de un científico natural en epistemólogo

En una carta a Fick, Helmholtz definió el objetivo de la filosofía, tal como él la entendía, con las siguientes palabras:

«Creo que la única manera de ayudar a la filosofía a que se ponga de nuevo en marcha es haciendo que se dedique seria y esforzadamente a la investigación de los procesos del conocimiento y los métodos científicos. [...] Sin embargo, para tal investigación crítica lo más importante es un conocimiento exacto de los procesos de la percepción sensorial»[16].

[16] Cf. H. Hörz y S. Wollgast, *Einleitung zu Hermann von Helmholtz' Philosophische Vorträge und Aufsätze*, p. XXVI.

Este fuerte impulso a usar resultados y conceptos científicos (en este caso fisiológicos) con el fin de aclarar cuestiones epistemológicas generales caracteriza toda la obra de Helmholtz. Construir una teoría del conocimiento sobre estas bases no era para Helmholtz el «hobby» de un científico en sus ratos de ocio o en sus deseos de popularización científica, sino una cuestión muy seria, estrechamente conectada con su trabajo científico cotidiano. Es esta actitud la que hace de él no sólo un gran físico, matemático y fisiólogo, sino también un auténtico filósofo.

Ahora bien, podemos preguntarnos ¿qué fue lo que hizo que un físico y fisiólogo como Helmholtz, a pesar de no tener ningún entrenamiento especial en la filosofía académica, se dedicara toda su vida con tanta energía a la elaboración de una epistemología general?

Aparte de las virtudes personales de Helmholtz (como su capacidad casi increíble de trabajo y su impresionante energía intelectual), hay dos clases de factores históricos que pueden explicar su fuerte impulso para saltar de la investigación científica altamente especializada a las cuestiones epistemológicas más profundas. Unos factores son de carácter general, los otros son más específicos. Vamos a describirlos brevemente.

En su historia, siempre estimulante, de las ideas europeas en el siglo XIX, Theodore Merz describe muy claramente la medida en que todas las actividades científicas en Alemania (especialmente durante la primera mitad del siglo XIX), al revés de lo que ocurría en los demás países europeos, estaban influidas por reflexiones y discusiones genuinamente filosóficas, incluso decididamente metafísicas. La totalidad de la vida cultural estaba permeada de discusiones filosóficas. Durante el segundo cuarto del siglo las controversias sobre las cuestiones del vitalismo y materialismo fueron particularmente candentes. El desarrollo intelectual de Helmholtz tuvo lugar en esta atmósfera. Por supuesto que este factor solo no es suficiente para explicar su carrera filosófica. No obstante, determinó el marco en el que se dio la misma.

Dentro de este marco general de influencias, la relación concreta entre la filosofía y la fisiología de los sentidos jugó un papel especialmente importante. El segundo cuarto del siglo XIX fue la gran época de Johannes Müller y su escuela. Helmholtz pertenecía a esta escuela. Los intereses de Müller y sus colaboradores no se limitaban a cuestiones específicamente fisiológicas. Müller era un kantiano convencido y pensaba que su «ley de las energías específicas de los nervios sensoriales» había proporcionado un fundamento experimental a la epistemología kantiana. Sus contemporáneos no tomaron esta posición como una evaluación esotérica de las consecuencias filosóficas de su obra. Un kantiano tan genuino como Friedrich Lange también lo sentía así:

«La fisiología de los órganos de los sentidos es la forma desarrollada o corregida del kantismo, y el sistema de Kant puede considerarse como un programa que condujo a los recientes descubrimientos en este campo»[17].

Esta identificación también era evidente para Helmholtz. Al principio, estando aún bajo la fuerte influencia de Müller, se consideró a sí mismo como un kantiano experimental, como un científico natural que trataba de mostrar, por medio de investigaciones fisiológicas específicas, que el esquema de Kant en principio era correcto. En su obra posterior, Helmholtz se apartó deliberadamente de la filosofía de Kant. Sin embargo, esto fue en su mayor parte debido a sus propios análisis de los fundamentos de la geometría, más que al desarrollo ulterior de sus estudios fisiológicos. Puede decirse que si bien en su epistemología geométrica abandonó el kantismo, en su epistemología fisiológica siguió básicamente fiel a dicha corriente. La «forma revisada de kantismo» (para adoptar la frase de Lange) no era para él otra cosa más que la interpretación fisiológica de la epistemología de la escuela crítica, y como tal seguía siendo válida. La «forma revisada de kantismo» era la concepción que emergió al combinar la filosofía con la fisiología. En una conferencia que Helmholtz dio a mediados de siglo en memoria de Kant dijo:

«El punto de contacto más estrecho entre la filosofía y las ciencias naturales es la disciplina de las percepciones sensoriales de los seres humanos»[18].

A lo que podemos añadir:

«La fisiología de los sentidos es el campo limítrofe en el que se entremezclan los dos grandes departamentos del conocimiento humano, que se suelen distinguir bajo los nombres de ciencias naturales y ciencias del espíritu, en el que se nos imponen problemas... que sólo se pueden resolver a través de un esfuerzo común»[19].

Esta filosofía interpretada fisiológicamente es el campo que Kant había intuido, pero que no había sido capaz de establecer realmente por carecer de los hallazgos empíricos necesarios. Kant sentó las bases del programa y éste fue llevado a cabo por la fisiología del siglo XIX. Así es como Müller y Helmholtz veían su relación intelectual con Kant; y en el presente contexto es completamente irrelevante preguntar si su interpretación de Kant era correcta o no. La cuestión interesante es más bien ésta: ¿En qué medida determinó la reconstrucción fisiológica del programa de Kant toda la obra de Helmholtz como fisiólogo filosófico?

---

[17] Cf. F. Lange, *Geschichte des Materialismus*, t. II, p. 409.

[18] Cf. H. v. Helmholtz, «Über des Sehen des Menschen», en *Philosophische Vorträge und Aufsätze*, p. 48.

[19] Cf. H. v. Helmholtz, *Vorträge und Reden*, t. 1, p. 16.

Helmholtz consideraba que la ley de las energías específicas que Müller había descubierto y que él mismo había desarrollado ulteriormente era tan significativa para una concepción filosófica general como la ley de la gravitación de Newton (cf. R. Kahl, *Introduction to Selected Writings of Hermann von Helmholtz*, *passim*). Ahora bien, en su opinión, la ley de las energías específicas no es sino una aplicación sistemática del principio kantiano de las formas *a priori* de la intuición al caso de las formas de percepción determinadas por cada uno de los órganos sensoriales, o incluso (en la versión generalizada del propio Helmholtz) por cada nervio sensorial (cf. Helmholtz, «Goethes Vorahnungen kommender naturwissenschaftlicher Ideen», en sus *Philosophische Vorträge und Aufsätze*, pp. 358 ss.). El principio que Kant había adelantado en términos generales y sin fundamentación experimental iba a ser cuidadosamente justificado e interpretado a través de la fisiología de los sentidos. Esta idea se le ocurrió a Müller sólo como proyecto vago; para Helmholtz se convirtió en todo un sistema filosófico y en un proyecto vital.

2. Los elementos básicos de la epistemología de Helmholtz

Los intereses filosóficos de los científicos naturales de la época, el entrenamiento fisiológico a través de su maestro Müller y la consideración de la teoría kantiana de las formas *a priori* de la intuición, todo esto constituyó el trasfondo de los elementos a partir de los cuales Helmholtz iba a construir su propia teoría del conocimiento. Veamos cómo se organizaron estos elementos en una estructura conceptual coherente.

Un análisis cuidadoso del enfoque epistemológico de Helmholtz revela los tres siguientes componentes como los más significativos: un simbolismo, un realismo estructural y un activismo epistemológico. Indudablemente, estos componentes, en particular el primero y el tercero, pueden encontrarse ya en Kant, aunque en una versión más abstracta. Por otro lado, en la versión fisiológica de Helmholtz, ellos constituyen el primer paso hacia el empiriocriticismo y hacia el fenomenalismo de Carnap. Estas teorías no son sino formas más precisas y desarrolladas de la teoría del conocimiento inaugurada por Helmholtz, quitándole a esta última la presuposición de una forma cualquiera de realismo sustantivo.

De los tres componentes mencionados, el simbolismo es el más importante, dado que es la precondición conceptual de los otros dos. Helmholtz da por supuesto que las investigaciones detalladas llevadas a cabo por Müller, por sí mismo y por otros en el campo de la fisiología de los sentidos ha mostrado que la naturaleza de las diferencias de modalidad dentro de la percepción sensorial pueden reducirse directamente a diferencias cualitativas (fisiológicas) entre los

diversos órganos sensoriales y los diversos nervios individualmente considerados –y no a las diferencias reales entre los estímulos externos. Müller había adelantado esta tesis ante todo para las «grandes» diferencias de percepción que nos permiten distinguir sentidos diferentes (por ejemplo, la diferencia entre la visión y la audición). Helmholtz extendió el principio de Müller a cada una de las diferencias cualitativas dentro de un mismo sentido (pongamos por caso las diferencias que hay entre dos distintas sensaciones de color).

El peligro filosófico que Helmholtz deseaba evitar y que parecía estar implícito en esta concepción de la percepción era el de un subjetivismo total en la teoría del conocimiento. El primer paso en este intento fue su simbolismo. Por oposición al realismo ingenuo, el simbolismo nos libera de la idea de que las sensaciones son reproducciones de los objetos externos. Las sensaciones en sí mismas consideradas no tienen por qué tener nada en común con sus fuentes externas. De esto, sin embargo, no es lícito inferir, como hace el idealismo subjetivista, que no existen fuentes de las sensaciones que sean independientes del sujeto. Las sensaciones ciertamente no son reproducciones, pero sí *símbolos* de la realidad externa: son signos para un «algo» independiente. O bien, en las palabras que el propio Helmholtz usó en un ejemplo concreto:

> «Las sensaciones de luz y color sólo son símbolos para ciertas relaciones en la realidad..., nosotros no sabemos nada de la naturaleza real de las relaciones externas designadas por ellas»[20].

Esta doctrina ciertamente es familiar en la historia de la filosofía. No obstante, el punto que se trata de subrayar aquí es el hecho de que el fisiólogo Helmholtz consideraba que sus propios resultados empíricos (y los de sus colaboradores) habían demostrado, más allá de cualquier duda razonable, que esta doctrina filosófica es simplemente verdadera.

Ahora bien, aunque cada sensación tomada individualmente no es más que un signo para un objeto externo que no posee ninguna semejanza esencial con ella, una colección de cierto número de sensaciones interconectadas puede constituir una estructura que ya no es meramente un signo, sino una verdadera reproducción de algo externo, objetivo. Lo representado es entonces no un solo estado de cosas externo como fuente de sensaciones singulares, sino más bien la *estructura nómica*, el marco de relaciones objetivas en el que se inserta cada una de las fuentes desconocidas. Si bien no podemos admitir la doctrina realista con respecto a cada uno de los pares *sensación-objeto*, estamos justificados en admitir otra clase de realismo: un realismo a nivel superior, si se quiere, un realismo con respecto

[20] Cf. Helmholtz, *Vorträge und Reden*, vol. 1, p. 30.

a las estructuras representadas (cf. Helmholtz, «Die Tatsachen in der Wahrnemung», en *Philosophische Vorträge und Aufsätze*, p. 256 y *passim*). Esta idea lleva directamente a una consecuencia metodológica que Carnap habría de explotar sistemáticamente muchos años después: la idea de que los objetos propios de conocimiento científico son estructuras sensoriales y no objetos singulares. No hace falta advertir que por esa época Carnap ya no hablaba más de realismo.

Para Helmholtz, en cambio, el realismo era un componente esencial de una teoría científica del conocimiento. Además, consideraba que el realismo epistemológico podía y debía estar fundado no en argumentos puramente filosóficos, sino en resultados fisiológicos que mostraban el carácter activo de los procesos epistémicos. El hecho demostrado fisiológicamente de que los procesos humanos de adquisición de conocimiento no se dan de una manera puramente pasiva, sino que contienen aspectos contructivos activos, era para Helmholtz la premisa científica del realismo filosófico:

> «Los esfuerzos de las escuelas filosóficas por fundamentar la creencia en la existencia de la realidad, [tenían que] fracasar mientras partían sólo del observador pasivo del mundo exterior; [...] no se daban cuenta de que las acciones humanas producidas por la voluntad constituyen una parte imprescindible de nuestras fuentes de conocimiento»[21].

El aspecto activo del proceso epistémico consiste, de acuerdo con Helmholtz, en la habilidad humana para modificar o desplazar agregados sensoriales de su experiencia por medio de la llamada *innervación*, es decir, mediante la estimulación arbitraria (voluntaria) de los nervios motóricos (cf. Helmholtz, «Die Tatsachen...», *op. cit.*, p. 257). A la clase de todos los agregados de sensaciones que pueden darse a través de la innervación en un momento dado, Helmholtz la llama la clase de los *presentables* («Präsentabilien») en ese momento (cf. *ibid.*, p. 259). Esta clase es primordialmente algo construido por el sujeto. Sin embargo, existe algo que el sujeto no puede determinar a voluntad por medio de su actividad motórica: el hecho empírico de que en un momento *distinto* y con respecto a las mismas innervaciones frecuentemente se dará una clase *distinta* de presentables. Por ejemplo, una serie de innervaciones puede llevarme en un momento dado a abrir el grifo del lavabo y tener un cierto conjunto de sensaciones visuales, tactiles y auditivas; este conjunto de sensaciones constituyen los presentables en ese momento. Pero en un momento distinto, y a pesar de actuar las mismas innervaciones, los presentables pueden ser distintos, porque no hay agua, porque se trata de un falso grifo o porque estoy soñando. En general, y tomado estrictamente, será casi siempre imposible para el sujeto reproducir exactamente la

[21] Cf. Helmholtz, «Goethes Vorahnungen...», *op. cit.*, p. 362.

misma clase de presentables que había percibido previamente, aun cuando las innervaciones sean exactamente las mismas. Esta divergencia entre la identidad de las innervaciones y la no-identidad de las clases de presentables correspondientes a momentos distintos es, según Helmholtz, el significado empírico concreto del realismo: el hecho de que, en momentos distintos, se den clases distintas de presentables y que la innervación activa del sujeto cognoscente puede modificar los elementos de cada clase a voluntad (puedo abrir o no el grifo), pero no las clases mismas. Una vez producida por la innervación voluntaria una cierta sensación, ésta irá asociada a otras sensaciones concomitantes, siendo el conjunto global algo que escapa a mi voluntad y que puede presentarse de manera distinta en otro momento. La existencia de las diversas clases de presentables es independiente de la innervación. (Nótese, sin embargo, que para este argumento en pro del realismo son esenciales las nociones de tiempo subjetivo y de memoria, por las cuales podemos darnos cuenta de que hemos tenido diferentes clases de presentables en diferentes momentos ante las mismas innervaciones.)

Simbolismo, estructuralismo, activismo: éstos son los elementos que Helmholtz combinó para obtener un tipo particular de realismo en epistemología –un realismo que tuvo sus orígenes en los resultados fisiológicos del segundo cuarto del siglo XIX. El estadio siguiente en esta evolución histórica está representado por ciertos filósofos con intereses en fisiología, como Avenarius, y ciertos fisiólogos con intereses filosóficos, como Mach. Esta evolución condujo del realismo estructural de Helmholtz al empirismo puro. Pero éste es el capítulo siguiente de la historia.

3.5. LA GENESIS DEL POSITIVISMO EN SU CONTEXTO CIENTIFICO

0. Introducción

En la literatura filosófica actual de todos los matices y en todas las latitudes es muy frecuente el uso del epíteto «positivista» para referirse, generalmente en tono peyorativo, a veces incluso muy apasionado, a filósofos o corrientes que, por alguna razón u otra, no le son simpáticos al autor que usa el epíteto en cuestión. Se da por supuesto que todo el mundo sabe cuáles son las características del positivismo. Sin embargo, es difícil admitir la validez de este supuesto, dada la variedad de usos contradictorios del término «positivismo» en la literatura. En realidad, nadie ha emprendido una verdadera clarificación de la naturaleza del positivismo, a pesar de su (supuesta) difusión en el mundo filosófico actual.

Cierto que existen algunas «definiciones» del positivismo por parte de algunos de sus detractores. Pero, prescindiendo de la dudosa objetividad en tales determinaciones, el problema es que ellas no contribuyen en gran cosa a una comprensión de la naturaleza de la concepción positivista; ello se debe a que el positivismo no consiste en un conjunto de tesis establecidas por escrito en algún sitio, sino más bien en una determinada actitud que ha evolucionado mucho al través del tiempo.

Las corrientes realmente significativas en la historia del pensamiento no pueden «definirse» asignándoles un par de rasgos generales. Con ello lo único que se consigue es un cliché, apto a lo sumo para manuales de divulgación. Lo que debe intentarse es determinar la peculiar evolución histórica de la corriente («corriente» entendida aquí en un sentido cuasi-literal), analizando todas las fases por las que

atraviesa y las modificaciones que sufre. Sólo así puede comprenderse algo de sus características peculiares. Esto es válido en general, pero en especial lo es para el positivismo, pues éste consiste más en una actitud que en un sistema.

Dar un primer paso en semejante intento de determinación «hidrográfica», podríamos decir, de la corriente positivista de los siglos XVIII y XIX es lo que se pretende en este artículo.

Por otro lado, aquí nos limitaremos a un aspecto del positivismo que, si bien esencial, ciertamente no es el único componente del mismo: sus relaciones con la investigación de fundamentos en las ciencias empíricas. Una adecuada comprensión de este componente es sin duda decisiva con respecto a las formas más modernas de positivismo, por ejemplo, la del Círculo de Viena, que se han autoconsiderado fundamentalmente como filosofías de la ciencia. Intentaremos aquí trazar la historia de esta componente y mostrar, entre otras cosas, que los orígenes históricos de la filosofía positivista de la ciencia deben buscarse no en el supuesto fundador del positivismo como sistema filosófico, Augusto Comte, sino en los trabajos de investigación de los fundamentos de las ciencias empíricas (especialmente de la mecánica) emprendidos *antes* y sobre todo *después* de Comte.

Pueden distinguirse por lo menos tres grandes fases en la evolución histórica del positivismo anterior al Círculo de Viena: un «proto-positivismo» o positivismo germinal anterior a Comte, ubicado en Francia desde mediados del siglo XVIII hasta la era napoleónica; el positivismo clásico de Comte y sus discípulos, con el que está estrechamente conectado el inductivismo de John Stuart Mill y de la mayoría de los metodólogos británicos de la era victoriana; y finalmente el positivismo crítico alemán del último tercio del siglo XIX, predecesor del positivismo lógico del Círculo de Viena.

A cada una de estas tres grandes fases del positivismo van asociadas manifestaciones secundarias o corrientes «laterales», por ejemplo la de los «ideólogos» de la Francia revolucionaria, el evolucionismo positivista de Spencer en Inglaterra y Haeckel en Alemania a mediados del XIX y la escuela «energética» alemana de Helm y Ostwald en la transición del XIX al XX.

No queremos entrar aquí en los detalles de las diversas manifestaciones históricas del positivismo, sino sólo trazar a grandes rasgos la historia de la germinación y el desarrollo de la temática que sería central en el positivismo lógico del siglo XX: el tema del carácter y los fundamentos del conocimiento científico.

1. El «proto-positivismo» francés anterior a Comte

Los inicios de un modo de pensar positivista se encuentran ya, sin lugar a dudas, en los «géomètres» franceses, es decir, en los físicos

matemáticos del siglo XVIII, formados en el fermento científico y filosófico producido en Francia por la confluencia (en parte violenta) de tres corrientes encontradas: la física newtoniana, el mecanicismo geométrico cartesiano y el empirismo británico.

Pueden distinguirse dos generaciones sucesivas de «géomètres»: la formada alrededor de D'Alembert a mediados del XVIII y la de Lagrange y Laplace poco antes de la Revolución. En el pensamiento y en la actividad de ambas generaciones científicas pueden detectarse ya con bastante claridad los rasgos fundamentales del positivismo posterior, por lo que no sería del todo desencaminado considerar a D'Alembert, Turgot y Condillac como los verdaderos fundadores del positivismo –suponiendo que tenga sentido en este caso buscar «fundadores»[22]. De todos modos, para evitar objeciones terminológicas es preferible reservar para estos autores precomtianos la designación de «proto-positivistas».

En su historia general del positivismo, Leszek Kolakowski rastrea los orígenes del positivismo mucho más atrás, hasta el nominalismo de la Baja Edad Media[23], y denomina «positivistas» a autores como los mecanicistas del XVII (Mersenne, Gassendi) y Hume, a quien considera, en particular, el padre del positivismo moderno.

Esta es, en mi opinión, una caracterización demasiado amplia del positivismo. Está claro que la actitud positivista tiene importantes puntos de contacto con otras corrientes «anti-metafísicas» como el occamismo, el mecanicismo y, sobre todo, el empirismo; pero no es históricamente adecuado identificarla con ellas. El positivismo posee rasgos específicos, debidos a su contexto histórico particular, que permiten separarlo de las corrientes mencionadas. Es una característica esencial del positivismo el presentarse como una filosofía de las *ciencias* empíricas, y no meramente como una teoría del *conocimiento ordinario* al estilo de los empiristas clásicos. Es una constante en el positivismo la preocupación por la metodología científica y por los análisis detallados de la estructura de las teorías científicas. Ahora bien, esta preocupación por la metodología científica sólo pudo surgir a partir del momento en que se hubo constituido una ciencia natural exacta «consciente de sí misma», es decir, consciente de su triple autonomía con respecto a la matemática pura, a la filosofía y al

[22] Esta opinión, que ya aparece insinuada en dos contemporáneos del positivismo crítico alemán de fines de siglo, Dilthey y Georg Misch, ha sido recientemente reafirmada por los historiadores de la ciencia Armin Hermann y Walter Kaiser en su artículo «Der Positivismus in der Physik des 18. und 19. Jahrhunderts», en *Rete*, 1972, t. I, cuaderno 2.

[23] L. Kolakowski: *Filosofia pozytywistyczna*, Varsovia, 1966. Trad. alemana de P. Lachmann, *Die Philosophie des Positivismus*, Munich, 1971. Esta es la exposición más sistemática y general que conozco de la historia del positivismo; a pesar de ello, una laguna importante en esta obra es la escasa consideración que se da en ella a las conexiones (a mi entender, esenciales) entre el pensamiento positivista y el desarrollo de las ciencias naturales.

conocimiento empírico común. Esto no ocurrió hasta mediados del siglo XVIII[24]. En conexión con este fenómeno histórico, y no antes, deben buscarse los orígenes del positivismo como forma especial de las corrientes antimetafísicas.

Dicho de otro modo, sólo a partir del momento en que el intelectual occidental fue plenamente consciente de la existencia de la explicación científica del mundo como algo radicalmente distinto de los tipos anteriores de explicación (sentido común, metafísica –incluida la metafísica materialista o mecanicista–, teología) pudo surgir la actitud «filocientífica» propia del positivismo, y su deseo de exponer ese tipo de explicación «en su estado puro», libre de las impurezas acientíficas que todavía contenía. En este sentido se distingue claramente el positivismo del empirismo clásico: este último está basado más en un análisis del conocimiento común que en una preocupación por las ciencias exactas.

Dentro del poderoso movimiento escéptico, antirreligioso, antimetafísico, cuasi-materialista, surgido con la *Encyclopédie* fue configurándose gradualmente como subforma especial la actitud proto-positivista. Desde un punto de vista ideológico puede considerarse la *Encyclopédie* como el efecto del influjo del empirismo británico sobre el cartesianismo francés, con la salvedad de que los enciclopedistas mismos no querían saber nada de los momentos espiritualistas contenidos tanto en el empirismo como en el cartesianismo. Muchos de los enciplopedistas se inclinaban declaradamente hacia el materialismo, que suponían «demostrado» por la física. No obstante, los primeros positivistas se distanciaron ya conscientemente de un materialismo estricto al estilo de Hélvétius: el materialismo era para ellos una hipótesis casi tan especulativa como el supuesto de un mundo sobrenatural y, en cualquier caso, igual de inútil para la ciencia.

Estas primeras formas de positivismo no se basaban en ningún sistema filosófico: eran decididamente «anti-sistemáticas». El sistema positivista surgirá tan sólo con Comte. De todos modos, de los escritos de autores como D'Alembert, Turgot, Condorcet, pueden entresacarse ya algunos de los aspectos básicos del positivismo posterior:

i) rechazo de cualquier pregunta por la esencia de las causas físicas;

[24] Baste recordar que Newton todavía consideraba que su obra era un primer paso para resolver problemas ontológicos, metafísicos e incluso teológicos generales, como atestigua el *Schclium Generale* de sus *Principia* y aún más concretamente su correspondencia con Bentley. La generación de Newton y Leibniz no habría comprendido el sentido de la distinción neta que se hace actualmente entre «filósofos» y «científicos». D'Alembert, en cambio, seguramente la habría aceptado.

ii) limitación de la tarea propia de la ciencia al establecimiento de relaciones lógico-matemáticas entre fenómenos;
iii) rechazo de toda explicación teológica, metafísica o teleológica de los fenómenos;
iv) fe en el progreso continuado de la comprensión científica del mundo –la única forma válida de conocimiento.

Lo más característico de los proto-positivistas del siglo XVIII es su estrecha conexión con la investigación matemática de la naturaleza. No es una mera casualidad que el más eminente de estos proto-positivistas, D'Alembert, fuera a la vez uno de los físicos matemáticos más importantes de su época; y más sintomático todavía es que se dedicara particularmente a la clarificación lógica de los fundamentos de la mecánica y criticase enérgicamente el concepto contemporáneo de *fuerza* (lo que sería también el *leitmotiv* de los trabajos positivistas de fundamentación, más de cien años más tarde).

En el Prólogo a su *Traité de Dynamique* (1743) subraya D'Alembert que su objetivo primordial es liberar la mecánica de todas las oscuridades metafísicas que la acosan en su estado presente.

> He proscrito totalmente las fuerzas inherentes a los cuerpos en movimiento, seres oscuros y metafísicos, que no son aptos más que para difundir las tinieblas en una ciencia que en sí misma debería ser clara (*Traité*, p. XXVIII).

Un aspecto positivista posterior que, en cambio, *no* se halla expresado claramente en los proto-positivistas es la necesidad de configurar unitariamente todas las ciencias según un aparato conceptual único. Los primeros en postular explícitamente esta necesidad serían los «ideólogos», una corriente filosófico-psicológica surgida con la Revolución Francesa como retoño de los proto-positivistas anteriores. Cabanis, uno de los representantes más notables de los «ideólogos», escribía alrededor del cambio de siglo:

> Es sin duda una idea bella y grandiosa la de considerar todas las ciencias y artes en mutua conexión, como todo indivisible, o como las ramas de un mismo tronco, unidas por su origen común, pero más unidas aún por el fruto que han de aportar todas por igual, la plenitud y la felicidad del hombre (*Oeuvres philosophiques*, p. 124).

La idea de la unificación conceptual de las ciencias (con su connotación más de exigencia ético-social que de realidad fáctica) resurgirá con el positivismo crítico alemán y llegará a su culminación con el formidable intento de la «Enciclopedia de la Ciencia Unificada» de los años 1930 y 1940.

Los «ideólogos» y otros filósofos y científicos emparentados con ellos de fines del XVIII, intentaron elaborar clasificaciones unificadoras de las ciencias entonces existentes; el punto de partida metodológico

era la ordenación de las ciencias particulares según el grado de abstracción y generalidad alcanzado. Mientras que los proto-positivistas habían extraído sus ideas metodológicas esclusivamente de la física matemática, con el cambio de siglo van entrando en la perspectiva positivista otras ciencias: primero la química (después de los trabajos de Lavoisier y Laplace), luego la medicina y la fisiología (que, en opinión de los «ideólogos», ya habían alcanzado el *status* de ciencia maduras), finalmente también las ciencias sociales.

En realidad, se da aquí una especie de paradoja histórica: si en el proto-positivismo de los «ideólogos» aparece explícitamente la idea de la unificación de las ciencias, es justamente porque en su época las ciencias ya habían empezado a disgregarse de hecho; la generación de D'Alembert no sentía la necesidad de tematizar el problema de la unificación de las ciencias, ya que el supuesto más o menos consciente de esa generación era que, en definitiva, sólo podía haber una ciencia: la mecánica. Con el cambio de siglo, sobre todo después de Lavoisier, este supuesto empezó a parecer menos obvio, y las corrientes positivistas existentes por esas fechas sintieron la necesidad de afirmarlo explícitamente.

A fines del siglo XVIII, y por iniciativa de la Convención Nacional revolucionaria, la antigua «Escuela Central de Obras Públicas» fue rebautizada y transformada en la famosa «École Polytechnique», que había de dedicarse mucho más intensamente a las investigaciones de ciencia pura. El alma de esta institución fue Lagrange, y en ella se formaron la mayoría de los científicos franceses posteriores.

La creación de la «École Polytechnique» significa un acontecimiento decisivo para la historia del positivismo decimonónico. Augusto Comte, cuya formación científica provenía de esa Escuela, manifestaría repetidas veces cuán fuerte fue el influjo ejercido sobre su pensamiento por el espíritu general que imperaba en la «École Polytechnique».

También el conde de Saint-Simon, que en las concepciones sociales de Comte iba a jugar análogo papel al que jugó la «École Polytechnique» en su metodología científica, mantuvo estrechos contactos con profesores y alumnos de la «École» por los años 1820, es decir, por la época en que Comte estudiaba allí.

2. El positivismo clásico

Durante años fue Comte el secretario de Saint-Simon. No cabe ninguna duda de que Saint-Simon ejerció una gran influencia en las ideas político-sociales de su joven amigo. No está del todo claro cuáles son las doctrinas de Comte que provienen directamente de Saint-Simon; pero la famosa «ley de los tres estadios» de la historia humana (teológico, metafísico, positivo), que ya había sido insinuada

por Turgot, fue postulada explícitamente por Saint-Simon; Comte se limitó a desarrollarla y a tratar de apoyarla en material histórico. También el término «positivo», como sinónimo de «científico», aparece ya en Turgot y en Saint-Simon. Y la idea de la fundación de una ciencia de la sociedad tan exacta como la física tiene su claro origen en Saint-Simon.

En cualquier caso, es indiscutible que es a Comte a quien corresponde el mérito (o demérito) de haber fundado el positivismo como sistema filosófico y como metodología de supuesta validez universal.

Iring Fetscher, en su introducción a la traducción alemana del *Discours sur l'esprit positif*, interpreta la empresa comtiana en su totalidad como el intento de forjar un sistema definitivo de filosofía de la historia. La aspiración básica de Comte era lograr una aplicación convincente del método de las ciencias naturales, que ya había hecho tan grandes progresos en otros campos, al dominio de la historia y de los fenómenos sociales. Concuerdo plenamente con la interpretación de Fetscher. El interés más genuino de Comte no estaba centrado en los fundamentos de las ciencias naturales, sino en una ciencia de la sociedad aún por construir. El propósito de Comte era llegar a ser para la sociología lo que Newton había sido para la mecánica y Lavoisier para la química.

A diferencia de los proto-positivistas anteriores por un lado y de los positivistas críticos posteriores por otro, Comte no se dedicó a la investigación de fundamentos en las ciencias naturales, a pesar de que su formación científica quizá se lo hubiera permitido.

Comte consideraba el estado de las ciencias naturales de su época, sobre todo la física y la química, como definitivamente maduro, y no esperaba ninguna sorpresa por ese lado. De ahí el tono dogmático, acrítico, casi sacerdotal y, en definitiva, aburrido con que Comte y sus discípulos exponen las bases de las ciencias naturales, en total contraposición con las fases anteriores y posteriores del positivismo. Esto explica también por qué el positivismo de Comte tuvo mucha mayor significación para el desarrollo de las ciencias sociales e incluso de la literatura, que para las ciencias naturales.

Para ser justos con Comte, no obstante, debe tenerse en cuenta que su presentación de la metodología científica no es sólo producto de su idiosincrasia personal, sino también reflejo de la situación general de las ciencias físicas y hasta cierto punto también de las biológicas durante la primera mitad del siglo XIX. La mayor parte de las ciencias naturales, sobre todo en Francia, se hallaban inmersas por la época de Comte en un estadio que, siguiendo la terminología de Thomas Kuhn, podríamos caracterizar de «ciencia normal», es decir, no se ponían en cuestión los fundamentos de las teorías científicas establecidas, se elaboraban primordialmente los detalles técnicos de las mismas y la imagen general de la empresa científica era la de un

progreso lineal «paso a paso». Esta imagen de la ciencia era probablemente la que imperaba en la «École Polytechnique» y la que recibió Comte en sus años de estudiante.

Sea como sea, la filosofía de Comte nos interesa en el contexto presente de la relación entre positivismo y desarrollo científico sólo en la medida en que representa una forma bien delineada del cientificismo moderno. Según Comte, todo desarrollo en la sociedad humana depende en última instancia del desarrollo científico. La historia de la ciencia es el núcleo de la historia general de la especie humana. No puede comprenderse bien el sentido de la historia universal si antes no se ha clarificado la evolución de las formas del conocimiento empírico.

Esta evolución sigue tres estadios: el teológico, el metafísico y el positivo. Toda ciencia, y por tanto también toda sociedad, debe atravesar estos tres estadios. Las diferencias entre los estadios vienen determinadas por el modo diverso como el hombre concibe al mundo. En el estadio *teológico*, el hombre intenta explicar los fenómenos naturales suponiéndolos efecto de la voluntad de espíritus o fuerzas sobrenaturales. En el estadio *metafísico* se interpretan los fenómenos como efectos de fuerzas o entidades abstractas, ya no más personificadas. En el estadio *positivo*, que es el de una ciencia o de una sociedad maduras, se describen y predicen con toda exactitud los fenómenos mediante leyes naturales –sin buscar explicaciones causales «tras» los fenómenos–; las leyes naturales son el producto exclusivamente de la observación y de la reflexión racional.

El verdadero objetivo de las ciencias *no* es buscar las causas ocultas de los fenómenos, sino sólo describirlos sistemáticamente, para poder hacer buenas predicciones. Las predicciones nos permiten actuar sobre la naturaleza; con ello se promueve el progreso tecnológico, la base de todo progreso humano.

La ley de los tres estadios la completó Comte con otra ley general acerca de la ordenación dinámica de las ciencias: la no menos famosa «ley enciclopédica», la cual fija un orden temporal en las ciencias según la complejidad de su objeto. La primera ciencia empírica que ha llegado al estadio positivo es aquella cuyo objeto muestra una estructura máximamente simple y regular: la astronomía. Después de ella vienen, por orden, la física, la química, la fisiología o biología y la sociología.

La posición que ocupa la matemática en este esquema no es del todo clara. En su *Discours* pone Comte la matemática en la cúspide, antes de la astronomía. En el *Cours de philosophie positive* afirma, en cambio, que la matemática debe ocupar un lugar especial en el esquema, puesto que no es una ciencia entre otras, sino el lenguaje conceptual de todas ellas. La matemática *no* es una *parte* de la ciencia, sino su fundamento conceptual. Esta segunda concepción de la matemática está más de acuerdo con nuestro paladar actual; curiosamente, no obstante –y esto obedece a una confusión conceptual de la

época–, entiende Comte por matemática no sólo la aritmética, el cálculo y la geometría, sino también la llamada «mecánica racional», es decir, los *Principia* de Newton y su secuela, a pesar de que él era consciente del carácter empírico, no-apriorístico, de la mecánica.

La ley enciclopédica consiste resumidamente en la afirmación de que el desarrollo de cada una de las ciencias depende del estado en que se halle coetáneamente la ciencia que la precede en la lista. No puede haber, por ejemplo, buena física sin una astronomía madura, ni buena biología sin una química constituida. A pesar de ello, cada una de las ciencias tiene su propia metodología autónoma, puesto que su objeto es también distinto del de las demás. Aquí vemos hasta qué punto las concepciones de Comte reflejan la situación de las ciencias en su época; poco antes de Comte se había inaugurado definitivamente la época de la especialización científica: los científicos ya no eran «sabios universales» que aspirasen a una comprensión global del universo, sino profesionales satisfechos con el estudio detallado de una parcela de la realidad, cada vez más impermeables a lo que pudieran hacer o decir los profesionales de otras parcelas, en la medida en que no tuviera un interés inmediato para su propio campo.

Con su doctrina de que cada ciencia tiene su propio método y objeto, Comte simplemente describió en un marco conceptual general lo que los científicos de su época ya *sentían*. Y esta doctrina, lo mismo que todas sus sutiles y rígidas clasificaciones en estadios y subestadios, ciencias y subciencias, es uno de los puntos que le separa más abiertamente de los proto-positivistas anteriores y, sobre todo, de los positivistas críticos posteriores.

Otro punto de divergencia consiste en la interpretación comtiana de las leyes naturales. Comte, y aún más decididamente su discípulo Pierre Laffitte, sostenían que las leyes básicas de una ciencia ya madura no pueden ponerse en cuestión. Deben ser consideradas inmodificables, de lo contrario resulta imposible el progreso científico. Nuevamente tenemos aquí un reflejo de la autocomprensión de la «ciencia normal» de la época. En el lenguaje de Kuhn podríamos decir: para Comte era inconcebible un progreso científico revolucionario; el progreso para él sólo podía consistir en un desarrollo lineal dentro de los cauces prefijados por los paradigmas de la ciencia normal. El lema comtiano «orden y progreso» no sólo debía aplicarse a la organización social, sino también a la ética científica. Esto era sólo la expresión filosófica de la actitud más o menos subconsciente de la mayoría de los científicos, particularmente de los físicos y químicos, durante la primera mitad del XIX.

Nada de todo esto, ni el sistematismo filosófico, ni la clasificación enciclopédica de las ciencias, ni el supuesto de la validez incontrovertible de las leyes generales, ni el ingenuo progresismo científico, se encontrará en la corriente de fines de siglo que hemos denominado «positivismo crítico».

La aversión a todo sistema filosófico, incluido un sistema de corte positivista, y la sospecha de que los fundamentos de la ciencia son mucho más dudosos de lo que se quiere admitir, características que ya se insinúan en algunos proto-positivistas del XVIII, resurgirán con renovada fuerza en los positivistas críticos alemanes del último tercio del XIX, sobre todo en su portavoz más eminente, Ernst Mach. En este sentido tiene los positivistas críticos más rasgos en común con la generación de D'Alembert que con la de Comte. Es notable la poca consideración que sentían Mach y Avenarius, por ejemplo, hacia la escuela de Comte o la de Mill y Spencer. En toda la obra de Mach he encontrado un solo lugar donde cita a Comte, y lo hace con un comentario negativo.

Los positivistas críticos no creían que los fundamentos de la ciencia fueran intocables ni les interesaban las clasificaciones escolásticas de las ramas científicas existentes. Creían, por el contrario, que ni los fundamentos ni las divisiones académicas existentes eran adecuados. En la admisión incondicional de la ciencia en su estado presente veían una nueva forma de dogmatismo y una nueva metafísica. Su programa era el de una reconstrucción crítica y unificada del conocimiento empírico, crítica en el sentido de eliminar toda oscuridad metafísica de la ciencia, unificada en el sentido de considerar que la base del conocimiento empírico debe ser común a todas las ramas científicas. Combatir la metafísica seguía siendo su objetivo, al igual que en Comte, sólo que ellos veían la metafísica allí donde Comte no la suponía: en los fundamentos de la ciencia misma, y en particular en la ciencia aparentemente más «madura», la mecánica.

3. Investigación sobre los fundamentos de la física

Como todo fenómeno histórico importante, esta nueva actitud hacia la ciencia no surgió de un modo casual. Tiene sus premisas históricas en la propia evolución de la ciencia decimonónica, particularmente de la física y de la fisiología de los sentidos. Creo que el punto de inflexión hacia la nueva actitud científico-filosófica puede fecharse alrededor de 1860. Determinar exactamente las causas del cambio es una tarea difícil, que requeriría probablemente una gruesa monografía. Aquí sólo podemos intentar un bosquejo a grandes rasgos de la situación histórica.

El cambio de actitud científica queda caracterizado esencialmente por una nueva consideración de la mecánica newtoniana. Esta ya no se ve como algo absolutamente seguro y firme; sus conceptos básicos, espacio y tiempo, masa y fuerza, dejan de ser considerados como nociones evidentes.

Puede que este escepticismo creciente con respecto a la supuesta evidencia y validez universal de la mecánica tuviera su origen en la

constitución entre 1850 y 1865 de una nueva rama de la física, la termodinámica fenomenológica, cuyos principios y conceptos fundamentales aparecían como totalmente ajenos al aparato conceptual newtoniano espacio-tiempo-masa-fuerza. También el creciente malestar ante la incapacidad de dar una formulación estrictamente mecánica a los resultados obtenidos en el electromagnetismo puede haber contribuido a hacer tambalear la fe en los fundamentos de la mecánica.

Pero no se trata aquí de lanzar hipótesis históricas, sino de constatar simplemente el hecho de que a principios del último tercio del XIX una serie de grandes físicos, en su mayoría alemanes, empezaron a ocuparse sistemáticamente de los fundamentos conceptuales y epistemológicos de la mecánica. Creo que los trabajos más significativos en este sentido son los de Helmholtz, Kirchhoff, Mach y Hertz, por orden cronológico. A excepción de Mach, no puede calificarse a estos investigadores de «pósitivistas» sin más; no obstante, sus trabajos de investigación acerca de los fundamentos fueron decisivos para el surgimiento del positivismo crítico alemán, y su ideario metodológico (sobre todo en el caso de Kirchhoff y Hertz) se hallaba muy próximo al del positivismo o estaba condicionado por él.

De Helmholtz a Hertz puede observarse una inclinación creciente a tratar los fundamentos de la mecánica con una actitud crítico-positivista. En todos estos autores hallamos puntos de vista de tono claramente positivista: una actitud negativa hacia los conceptos de fuerza y de causa física; la admisión de la necesidad de lograr una unificación lógica de las diversas ciencias; la tesis de que sólo los hechos directamente observables son de fiar; el rechazo decidido de cualquier forma de metafísica; la acentuación de valor práctico-biológico de la ciencia: la tarea primordial de la ciencia no es *explicar* los fenómenos, sino *describirlos* con la máxima precisión posible para hacer predicciones que nos permitan actuar en consecuencia.

Ya en Helmholtz, el menos positivista de los autores mencionados, encontramos una preocupación constante por el problema de lograr una unificación real de las ciencias. Helmholtz fundaba su esperanza de realizar esta empresa en las nuevas investigaciones que se estaban llevando a cabo a mediados de siglo justamente en la fisiología de los sentidos –esperanza que resurgirá en Mach y mucho más tarde en el *Aufbau* de Carnap–, y a las que el propio Helmholtz contribuyó en gran medida.

En una conferencia con el característico título «Sobre la relación de las ciencias naturales con la totalidad de la ciencia», en 1862, lamentaba Helmholtz apasionadamente la disgregación de las ciencias particulares, que ya en su época era cada vez más acusada. Consideraba de necesidad imperiosa el establecimiento de conexiones lógicas entre las diversas ciencias, tanto por razones internas («para promover mejores resultados en la labor científica») como por razones externas

(«para mantener un equilibrio sano entre las fuerzas espirituales»). Dado que las ciencias habían crecido tanto en extensión, el establecimiento de la conexión requerida no podía lograrse de una manera directa, sino buscando una base profunda y común que sistematizase todo el conocimiento científico.

Helmholtz veía claramente la dificultad de semejante empresa y, a diferencia de muchos de sus contemporáneos, no creía que esa base pudiera hallarse en la mecánica en su forma presente; más bien la buscaba en una especie de «protofísica fisiológica». Ya en Helmholtz aparece claramente el escepticismo con respecto a la mecánica, el cual luego será un componente esencial del positivismo crítico. Le parecía muy dudoso el valor de las explicaciones mecánicas de fenómenos no mecánicos (electromagnetismo, fenómenos térmicos), que era el objetivo básico de la mayoría de los científicos de la época. En su propia Introducción a la *Mecánica* de Hertz escribiría:

> Físicos ingleses, como Lord Kelvin con su teoría de los átomos en remolino y Maxwell con su hipótesis de un sistema de celdas cuyo contenido se halla en rotación, en la que se basa su intento de explicar mecánicamente los fenómenos electromagnéticos, se sienten manifiestamente más satisfechos con este tipo de explicaciones [mecánicas] que con la *simple descripción más general posible de los hechos*... Debo confesar que yo mismo me he mantenido hasta ahora en este tipo de descripción, con la que me siento más seguro (H. Hertz, *Prinzipien der Mechanik*, pp. XXI-XXII).

Tampoco Gustav Kirchhoff se sentía satisfecho alrededor de 1875 con las pretensiones de la mecánica de explicar todos los fenómenos naturales. Todavía en 1865 se adhería a la concepción de Laplace, según la cual el objetivo de la mecánica es buscar en todas partes las fuerzas que expliquen causalmente los movimientos y cambios de los sistemas físicos. Once años más tarde había cambiado totalmente de opinión, sobre todo porque el concepto de *fuerza* le parecía sospechoso. En el Prólogo a sus *Lecciones sobre mecánica* escribe:

> Suele definirse la Mecánica como la ciencia de las *fuerzas*, y las fuerzas como las *causas* que producen o *tienden* a producir movimientos. Cierto que esta definición ha sido de gran utilidad para el desarrollo de la Mecánica... Pero está preñada de la oscuridad inherente a los conceptos de causa y de tendencia. Por esta razón propongo como tarea propia de la Mecánica la de *describir* los movimientos que ocurren en la naturaleza y describirlos del modo más completo y más simple posible. Quiero decir con ello que lo único que nos debe interesar es averiguar *cuáles* son los fenómenos que ocurren y no determinar sus *causas*.

Según Kirchhoff, pues, el objetivo propio de la mecánica no es la *explicación* de los movimientos, sino sólo su *descripción* exacta. Dicho de otro modo, el acento no se pone en la explicación *dinámica* del

universo (determinación de fuerzas), sino en su descripción *cinemática* (determinación de relaciones espacio-temporales).

Casi simultáneamente a Mach y Avenarius propone Kirchhoff también un *principio de economía conceptual* destinado a justificar el aparato conceptual de la ciencia. Según este principio, cuanto más simple sea la forma de las leyes físicas, más aptas serán ellas para lograr descripciones y predicciones controlables... Las consideraciones de simplicidad justifican la introducción de conceptos «abstractos» como el de fuerza. Tales conceptos pueden introducirse en la mecánica en calidad de cómodas abreviaciones, pero sólo en la medida en que se esté seguro de que pueden reducirse en principio a los conceptos empíricos básicos de espacio, tiempo y materia. En consecuencia, Kirchhoff da un paso audaz para su época: *define* la fuerza simplemente como el producto de la masa por la aceleración; convierte por tanto el segundo principio de Newton en una mera tautología.

Por esta época sostenía ya Ernst Mach las mismas opiniones con respecto a los fundamentos de la mecánica. Sólo que Mach era más radical que Kirchhoff (y también más ignorado por el mundo académico). Ya en 1868, ocho años antes de la definición de fuerza propuesta por Kirchhoff, había dado un paso más en el «reduccionismo cinemático»: en su artículo «Sobre la definición de la masa» (aparecido en el *Carls Repertorium der Experimentalphysik*, una especie de revista general de física) rechaza Mach el concepto de masa como concepto básico (primitivo) de una formulación adecuada de la mecánica: la masa es reducible definicionalmente a magnitudes directamente observables, a saber, las aceleraciones relativas de dos cuerpos en proximidad espacial.

Creo que vale la pena detenerse un poco en esta cuestión algo «técnica» de la definición machiana de masa, puesto que en ella están contenidos en germen los principios metodológicos que Mach desarrollaría posteriormente en sus obras más filosóficas[25].

Para empezar, rechaza por oscura la identificación usual de la masa con la «quantitas materiae»:

> La idea de la «cantidad de materia» no es adecuada para determinar y aclarar el concepto de masa, puesto que ella misma carece de suficiente claridad. Esto sigue siendo el caso aun cuando se recurra, como han hecho algunos autores, al número de los hipotéticos átomos [contenidos en un cuerpo] (E. Mach, *Die Mechanik in ihrer Entwicklung*, 9.ª ed., p. 210).

[25] Este artículo de juventud de Mach (donde también define ya la fuerza como producto de la masa por la aceleración) apareció en forma muy comprimida en la revista mencionada y su sentido no es fácil de comprender para quien no conozca la epistemología general de Mach. Por ello no es de extrañar que fuera prácticamente ignorado. En su *Mecánica*, 15 años más tarde, volvió Mach a exponer en forma más clara su definición de masa; a esta segunda versión nos atenemos aquí, aunque no existen diferencias esenciales entre ambas.

En vez de la confusa noción de cantidad de materia, parte Mach del tercer principio de Newton (acción = reacción) para determinar primeramente la igualdad de masas; es decir, primero define el concepto de igualdad de masas y luego el concepto general de masa como magnitud: «Decimos que dos cuerpos tienen igual masa si cuando interactúan se inducen mutuamente aceleraciones iguales y de sentido opuesto», y añade una observación típicamente positivista: «Con ello no hemos hecho más que *denominar* una relación *fáctica*» (*op. cit.*, p. 211).

Es decir, la experiencia nos enseña directamente que los cuerpos del universo pueden agruparse en clases de equivalencias determinadas por la igualdad en valor absoluto de las aceleraciones que se provocan mutuamente los elementos de cada clase. Para *denominar* con una abreviación cómoda dichas clases de equivalencia utilizamos el término «igualdad de masas»: dos cuerpos poseen *igual masa* si y sólo si pertenecen a la misma clase de equivalencia de aceleraciones mutuas. Esto es una mera definición, y una definición tal, que en el *definiens* sólo aparecen conceptos cinemáticos (conceptos «observacionales»).

El siguiente paso es la definición del concepto de masa en general o, lo que viene a ser lo mismo, la introducción de una escala (un orden) de masas. Por convención, se escoge una de las clases de igualdad de masas y se asigna a sus elementos el valor unidad. Para determinar la masa de cualquier otro cuerpo C no perteneciente a dicha clase se toma un representante A de la misma y se le coloca en proximidad espacial de C. Observamos sus aceleraciones mutuas y definimos: «C tiene la masa m si y sólo si la aceleración inducida en A por C es m veces mayor que la aceleración inducida en C por A.» Nuevamente es un hecho empírico que las escalas de masas que se obtienen tomando otras clases de equivalencia como unidades son todas proporcionales entre sí. Y como justificación general de su procedimiento concluye Mach: «Mi definición resulta del intento de determinar *las relaciones de dependencia entre los fenómenos* y de eliminar toda oscuridad metafísica» (*op. cit.*, p. 212).

La definición machiana de masa puede leerse todavía hoy en manuales de mecánica clásica[26]. Investigaciones relativamente recientes en filosofía de la ciencia han puesto en claro que la supuesta definición propuesta por Mach no es una definición en sentido lógico estricto, sino algo mucho más débil. No parece posible una definición adecuada de masa a partir de conceptos puramente cinemáticos[27]

---

[26] Cf., por ejemplo, K. R. Symon, *Mechanics*, Reading, Mass., 1960, pp. 5-7.

[27] Cf., a este respecto, por ejemplo, P. Suppes, *Introduction to Logic*, Nueva York, 1957, p. 298. En mi opinión, una clarificación definitiva del *status* lógico de la «definición» de Mach no se logró hasta la obra de J. D. Sneed, *The Logical Structure of Mathematical Physics*, Dordrecht, Holanda, 1971, pp. 59-60, 135.

Pero no se trata aquí de adentrarse en las discusiones de la actual filosofía de la ciencia, sino de determinar el sentido histórico de la propuesta de Mach. Se trata de un intento de reconstrucción de la mecánica de signo claramente positivista en el siguiente sentido: para la determinación conceptual de los fenómenos mecánicos sólo se admiten en principio nociones espacio-temporales, pues sólo éstas son asequibles a la observación directa; todos los demás conceptos de la mecánica, incluidos los de masa y fuerza, deben ser reducibles a nociones cinemáticas, de lo contrario son «oscuridades metafísicas».

Las propuestas de Mach fueron ignoradas durante largos años. Sólo después de que una autoridad como Kirchhoff hubo allanado el camino empezaron a ser tomadas en consideración, particularmente después de la publicación de la *Mecánica* del propio Mach, en 1883. En esta obra afirma Mach una vez más la definibilidad cinemática de masa y fuerza y critica a fondo los conceptos newtonianos de espacio y tiempo absolutos, que para él eran puras fantasías metafísicas.

La crítica machiana de los conceptos fundamentales de la mecánica habría de causar un profundo impacto en Einstein y preparar su camino hacia la teoría de la relatividad restringida. El papel que jugó Mach en la génesis de este gran cambio de la historia de la física ha sido reconocido claramente por el propio Einstein en su *Autobiografía:*

> No deberíamos asombrarnos de que prácticamente todos los físicos del último siglo viesen en la mecánica clásica una base firme y definitiva de la física entera, incluso de la ciencia en su totalidad, y de que intentasen una y otra vez fundamentar en la mecánica también la teoría maxwelliana del electromagnetismo, que se iba afianzando paulatinamente; [...] en su pensamiento consciente se aferraban a la mecánica como base segura. Fue Ernst Mach quien en su *Historia de la mecánica* hizo tambalear esa fe dogmática; justamente en este sentido ejerció este libro sobre mí una profunda influencia en mi época de estudiante. Veo la verdadera grandeza de Mach en su escepticismo insobornable y en su independencia de criterio; en mis años jóvenes me impresionó mucho la posición epistemológica de Mach. (En P. A. Schilpp (ed.), *Albert Einstein. Philosopher-Scientist*, La Salle, Illinois, 1949, p. 20.)

La expulsión sistemática llevada hasta sus últimas consecuencias del concepto de fuerza de la mecánica fue la tarea emprendida por Heinrich Hertz en sus *Principios de la mecánica* en 1894. Hertz, que había sido discípulo de Helmholtz y de Kirchhoff, reconoce en la Introducción a su libro que debe su concepción general de los fundamentos de la mecánica a los trabajos de Mach.

El objetivo central de Hertz en esta obra es construir de modo riguroso y sistemático los fundamentos de la mecánica partiendo de una base puramente empírica, «libre de metafísica». Espacio, tiempo y partícula o punto-masa (¡no masa!) son en Hertz los conceptos empíricos fundamentales. Con ellos, y sólo con ellos, formula Hertz

los axiomas de su mecánica. Fuerza y energía no figuran en la base conceptual propuesta por Hertz.

A diferencia de los investigadores anteriores, Hertz distingue estrictamente entre las cuestiones apriorísticas (puramente matemáticas) y las empíricas (propiamente físicas) en un tratamiento adecuado de la mecánica. Esta división estricta se refleja incluso en el modo concreto de exposición escogido por Hertz: en la primera parte del libro se tratan las cuestiones apriorísticas y en la segunda parte las empíricas, netamente separadas. La necesidad de distinguir claramente entre ambos tipos de cuestiones en toda obra científica sería más tarde uno de los requisitos básicos del positivismo lógico. Es sabido que las concepciones metodológicas de Hertz contribuyeron en medida importante a la filosofía de la ciencia del *Tractatus Logico-Philosophicus* de Wittgenstein (Hertz es, junto con Russell y Frege, uno de los pocos autores a quienes Wittgenstein hizo el honor de citar en su *Tractatus*), y a través de él influyó en la filosofía de la ciencia del Círculo de Viena. Pero éste es otro capítulo de la historia.

4. El positivismo crítico

Todos estos trabajos en los fundamentos de las ciencias, especialmente de la mecánica, se hallan en estrecha conexión con el surgimiento de la corriente filosófica que hemos denominado «positivismo crítico» y que, centrada en Alemania, tendría prontas repercusiones en otros países, tales como Gran Bretaña y Norteamérica. Los trabajos de fundamentos mencionados fueron en parte la premisa histórica, pero en parte también la consecuencia de la nueva actitud filosófica. No es fácil determinar en cada caso el orden exacto de las influencias.

También es difícil determinar exactamente cuándo surgió el positivismo crítico. Puestos a buscar una fecha de nacimiento, podría darse como tal la «definición de masa» de Mach en 1868 o quizá la formulación dada por Avenarius del principio de economía epistémica de 1876. En cualquier caso, antes de la publicación del *Análisis de las sensaciones* de Mach en 1885 eran poco conocidas del público general las nuevas ideas positivistas.

El *Análisis de las sensaciones* fue por su peculiar carácter (estimulante mezcla de investigaciones científicas concretas y tesis epistémico-metodológicas generales) una obra muy leída en la época, tanto en medios estrictamente científicos como en sectores más amplios del público. A la primera edición de 1885 siguieron seis ediciones más, aún en vida del autor (Mach murió en 1916). De entre los filósofos coetáneos fueron sobre todo los empirocriticistas Avenarius y Petzoldt los más influidos por la epistemología machiana; entre los científicos de la época que se sintieron fuertemente atraídos por Mach pueden mencionarse: William K. Clifford, Karl Pearson (íntimo amigo

de Mach, a quien está dedicado el *Análisis de las sensaciones*), William James (como psicólogo, no como filósofo), Max Planck y el joven Einstein.

Las ideas de Mach acerca de los fundamentos del conocimiento humano no eran, sin duda, completamente originales. Físicos como Helmholtz y Ewald Hering, que también se interesaban por la psicofisiología, habían anticipado ya algunas de ellas. En el campo estrictamente filosófico, Avenarius ya había inaugurado en 1876 muchas de las tesis «económico-biológicas» de la nueva epistemología en su librito acerca del *Principio del mínimo esfuerzo como principio-guía de todo conocimiento positivo*. Estos y otros autores indagaban las conexiones existentes entre física y psicología e intentaban una solución al llamado problema psicofísico a través de una concepción unitaria, según la cual la base del conocimiento no debe buscarse ni en los conceptos físicos ni en los psíquicos; antes bien, el aparente abismo entre física y psicología debe salvarse al nivel de un substrato común primario.

De todos modos, el *Análisis de las sensaciones* fue la primera obra sistemática en este campo, la primera que planteó claramente el programa. El problema que intentaba solucionarse, tal como aparece planteado en el *Análisis*, puede describirse concisamente así: cómo hallar conceptos y leyes que sean de carácter más básico que los de la física y la psicología, y de los cuales puedan deducirse (o «construirse») estos últimos. El programa de Mach para solucionar este problema era el siguiente. Las nociones y las leyes requeridas deben buscarse en el material proporcionado por nuestra experiencia fisiológica; esto es lo único que nos viene dado de forma inmediata, y por sus propiedades no pertenece ni al dominio físico ni al psíquico. Los conceptos de la física y de la psicología deben construirse matemáticamente a partir de la experiencia sensible. Ya la primera frase en el Prólogo a la primera edición del *Análisis* caracteriza bien el programa de Mach:

> Profundamente convencido de que la ciencia en general y la física en particular esperan las más importantes aclaraciones sobre sus fundamentos de la biología y, concretamente, del análisis de las sensaciones orgánicas, me he visto conducido repetidas veces a este campo.

Pero el objetivo último de este programa no era sólo la clarificación de los fundamentos de las diversas ciencias, sino su unificación. Para ello era necesario partir de una idea clara y firme del proceso cognoscitivo en general: la idea de que el acto cognoscitivo y su objeto son idénticos por principio, por mucho que al nivel más complejo de las diversas ciencias parezcan divergir.

Todo conocimiento auténtico descansa sobre la experiencia sensorial, y ésta consiste en un gran número de elementos discretos, a los que Mach llama a veces «sensaciones». Hay que manejar el término

«sensación» con cuidado, sobre todo no entenderlo en modo pasivo, como si hubiera un «sujeto» previo a las sensaciones que «las sintiera»; lo único que *hay* es ese conjunto de sensaciones.

Las sensaciones pueden agruparse según ciertas relaciones de similitud en diversas categorías: sensaciones cromáticas, tactiles, térmicas, etc. En principio, estas categorías no tienen nada que ver entre sí. Afortunadamente, la experiencia sensorial está constituida de tal manera que pueden establecerse ciertas correlaciones más o menos regulares de dependencia entre los diversos grupos de sensaciones. Esto permite la construcción de complejos más o menos estables, los «objetos sensibles».

Según como estén estructuradas las diversas relaciones que se establezcan entre las sensaciones, se obtendrán los diversos objetos de las ciencias particulares. En una reconstrucción ideal de la base de la ciencia habría que representar dichas relaciones en forma de funciones (en el sentido matemático: relaciones unívocas en uno de sus argumentos) y expresar las dependencias regulares entre las sensaciones mediante tales funciones como ecuaciones en el tiempo (el tiempo fenoménico también es para Mach no más que un conjunto de sensaciones específicas). Se obtendrían entonces dos grandes grupos de funciones de sensaciones, distinguibles entre sí por su forma analítica general: la clase de las funciones de las que diríamos que constituyen objetos o procesos físicos y la clase de las funciones de las que diríamos que constituyen procesos psíquicos. Los objetos físicos vendrían determinados por conjuntos (de conjuntos) de funciones «físicas» que satisfacen ciertas propiedades formales; análogamente se definirían los objetos psíquicos. Entre las funciones «físicas» y las «psíquicas» han de establecerse también ciertas correlaciones; ésta es la tarea propia de la ciencia llamada *psicofisiología*[28].

El punto decisivo aquí es que los argumentos de todas las funciones o relaciones consideradas son siempre sensaciones tomadas como elementos neutrales. A partir de ellas se constituye el mundo físico o el psíquico, según el tipo de funciones de sensaciones que se establezcan. El desarrollo exacto de esta empresa es, según Mach, la tarea de una epistemología científica, definida como ciencia de las funciones o relaciones de sensaciones. Una vez constituida esta ciencia (para lo cual Mach pretende dar los primeros pasos), se vería claramente que física, fisiología y psicología no se distinguen entre sí por su objeto, que es siempre el mismo (conjuntos de sensaciones), sino por el modo de ordenar y correlacionar los elementos básicos. Esta es, en definitiva, la misma empresa, a cuya realización formal y

[28] Esta es naturalmente una exposición muy simplificada del sistema conceptual de Mach. Sin embargo, el propio Mach no es en su libro mucho más explícito y riguroso por lo que respecta a este punto. Una reconstrucción semiformal más detallada y exacta de la base del sistema machiano se hallará en mi libro *La estructura del mundo sensible*, Barcelona, 1973, Cap. I, especialmente pp. 46-49.

efectiva se dedicaría Carnap cuarenta años más tarde en su *Logische Aufbau der Welt*, con la única diferencia (aunque esencial) de que Carnap dispondría ya de un instrumento del que Mach carecía por completo: la lógica de los *Principia Mathematica*.

Este programa de reconstrucción de la totalidad del conocimiento científico causó una gran impresión en los intelectuales de la época. Después de la publicación del *Análisis*, un número creciente de filósofos, psicólogos, fisiólogos y físicos se dedicaron a investigaciones análogas a las de Mach, con un espíritu de confluencia entre las diversas disciplinas que sería inconcebible en nuestra época. Así podía escribir Mach con justificada satisfacción en 1902 en el prólogo a la cuarta edición de su libro:

> Hoy veo en fin que un gran número de filósofos... y también algunos científicos, sin conocerse unos a otros, han caminado en direcciones convergentes. Si por esto el valor de mi trabajo particular queda reducido a poca cosa, en cambio puedo asegurar que no persigo una quimera, sino que he colaborado en una obra general.

En nuestro siglo las formas de pensamiento positivista, fundamentalmente centradas alrededor del Círculo de Viena en los años 1920 y 1930, entrarían en una nueva fase –quizá la última fase histórica del positivismo. Es indudable que el positivismo lógico del periodo de entreguerras ofrece aspectos totalmente novedosos que poco o nada tienen que ver con las corrientes anteriores y que no consisten en un mero desarrollo de algo preexistente. Pero es asimismo indudable que muchas de las tesis y de los puntos de vista centrales del nuevo positivismo, sobre todo por lo que respecta a la filosofía de las ciencias empíricas, están contenidos o prefigurados ya en las formas anteriores, particularmente en los trabajos críticos de los físicos alemanes del último tercio del XIX. Sin un estudio detallado de esas formas anteriores no es posible una comprensión plena del positivismo del siglo XX.

4. HACIA UNA ONTOLOGIA DE LA CIENCIA

La ontología se ocupa de dar respuesta a la pregunta general: «¿Qué tipos de cosas hay?», y a preguntas más particulares, acompañantes o derivadas de esta pregunta general. Esta equivale a preguntarse por las maneras en que podemos agrupar las cosas que suponemos que existen. Dicho de un modo más vago y metafórico, se trata de preguntarse por los «cajones» en los que podemos encasillar «la realidad». O sea, la ontología trata de sostener o refutar que ciertas cosas particulares caen bajo ciertas categorías generales previamente determinadas. Esta es la forma implícita de las afirmaciones o negaciones de existencia que hace la ontología. En este sentido hacemos ontología a cada paso; tanto en la vida cotidiana, cuando negamos que existan centauros, por ejemplo, como en las ciencias empíricas, cuando afirmamos que existen neutrinos. También hacemos ontología en filosofía, a un nivel conceptual más general, como cuando afirmamos o negamos la existencia de universales, pongamos por caso. Algunos filósofos suelen reservar el término «ontología» para el tratamiento de este último tipo de cuestiones; pero esta limitación no tiene por qué ser más que el producto de una deformación profesional. No sólo en filosofía, sino también en el sentido común y en las ciencias hacemos continuamente «compromisos ontológicos».

Para que tales compromisos sean auténticos y firmes es necesario que las categorías ontológicas asumidas posean un grado razonable de claridad, precisión y aplicabilidad. La mejor garantía de ello es un análisis sistemático del significado de los términos que usamos para expresar dichas categorías. Por ello, la ontología está íntimamente ligada a la semántica.

El filósofo de la ciencia puede enfrentarse a cuestiones ontológicas en un doble sentido. En primer lugar, puede encontrarse con compromisos ontológicos hechos de manera generalmente implícita y frecuentemente confusa por la propia ciencia. La tarea de analizar y aclarar esos compromisos podría (y quizá debería) ser hecha por los propios científicos, quienes en principio deberían estar interesados en tener una idea clara acerca de cuáles son las categorías ontológicas que usan, o que les conviene usar, en su trabajo cotidiano. Sin embargo, el científico «normal» suele ser muy descuidado en sus compromisos ontológicos –posiblemente por buenos motivos prácticos: la consideración de temas ontológicos le llevaría una parte demasiado considerable del tiempo y de la energía que necesita para otras cosas. Por ello es legítimo y fructífero que el filósofo de la ciencia se ocupe de analizar los compromisos ontológicos que, en principio, el científico hace o debería hacer. A este nivel de análisis lo podemos llamar «ontología de primer orden».

Ahora bien, el propio filósofo de la ciencia tiene que hacer *sus* compromisos ontológicos específicos –como todo el mundo. Y también conviene que obtenga la máxima claridad al respecto, explicitándolos y analizándolos. Así como, idealmente, el científico debería tener claras las categorías ontológicas que usa para investigar los fenómenos físicos, biológicos o humanos, así también el filósofo debería explicitar las categorías ontológicas que él mismo usa para estudiar las teorías científicas, u otros productos culturales análogos. A esta segunda clase de empresa la podemos llamar «ontología de segundo orden». Como suele ser el caso con este tipo de distinciones de niveles, la ontología de segundo orden puede y suele tener consecuencias para la ontología de primer orden.

Tanto la ontología de primer orden como la de segundo (pero especialmente esta última) son todavía campos muy poco explorados. Y, además, repletos de confusiones y vaguedades. Hay que hacer aún un gran esfuerzo de clarificación y crítica conceptual en ambos campos. En el primer capítulo de esta última parte se pretende hacer una contribución a la ontología de segundo orden (con algunas implicaciones para la de primer orden), y los dos capítulos posteriores están dedicados a ontología de primer orden.

4.1. ONTOSEMANTICA DE LAS TEORIAS CIENTIFICAS

0. El objetivo de un análisis ontosemántico de las teorías científicas

Un producto cultural característico de la especie humana son las teorías científicas. Es un producto, además, al que, por su desarrollo tanto cuantitativo como cualitativo, los seres humanos atribuyen cada vez mayor importancia. Es natural, por tanto, que nos veamos llevados a plantear ciertas preguntas acerca de las teorías científicas. Algunas de estas preguntas son: «¿Qué son las teorías científicas?», «¿Cómo son?», «¿Cómo devienen?» Por su forma, éstas son típicas cuestiones metacientíficas, es decir, filosóficas. La tercera pregunta, en el sentido de «cómo nacen, evolucionan y mueren las teorías científicas», ha sido ampliamente discutida en las últimas décadas, y pertenece al estudio diacrónico de la ciencia (historia de la ciencia y filosofía de la historia de la ciencia). La segunda pregunta, en el sentido de «cuál es la forma o estructura de las teorías científicas», ha sido tratada con rigor y detalle crecientes, sobre todo en los últimos años, dentro de la filosofía formal de la ciencia. Respuestas más o menos precisas a esta pregunta las han dado diversos enfoques dentro de lo que podríamos denominar «análisis estructural de la ciencia». La primera pregunta, en el sentido de «qué clase de entidad es una teoría científica», es una pregunta ontológica. La filosofía de la ciencia no le ha dado, en general, más que respuestas someras y vagas en el mejor de los casos; en el peor, la ha ignorado totalmente. El presente capítulo propone preparar el terreno para una discusión generalizada del tipo de respuestas posibles a esa cuestión ontológica; y para ello

intenta hacer uso de algunos elementos del esquema conceptual de Frege.

El hecho de que el grueso de la filosofía de la ciencia haya encarado con desgana o negligencia la cuestión ontológica acerca de las teorías científicas es seguramente debido a una confusión metodológica: muchos autores, entre ellos el presente, han creído que una respuesta razonablemente clara a la primera pregunta, «¿Qué son las teorías científicas?», sólo puede venir dada en términos de la segunda, «¿Cuál es la estructura de las teorías científicas?», o hasta de la tercera, «¿Cómo evolucionan las teorías científicas?» No es casualidad que tantas obras importantes de la filosofía reciente que pretenden dar una caracterización general de la ciencia aparezcan rotuladas con títulos que empiezan por el término «La estructura...». Existe la creencia ampliamente difundida de que una respuesta sensata a la pregunta «¿Qué es una teoría?» o incluso, en general, «¿Qué es *x*?», para cualquier *x*, vendrá necesaria y suficientemente dada en términos de una caracterización de las estructuras sincrónicas y/o diacrónicas inherentes a *x*.

Esta actitud ha tenido manifiestamente efectos positivos en el sentido de evitar la persecución de quiméricas esencias de los objetos, especialmente en el caso de objetos abstractos, como son las entidades teóricas. Sin embargo, también ha tenido el efecto negativo de restringir indebidamente el alcance posible de los cuestionamientos filosóficos. Simplemente, no es cierto que todo lo que podamos preguntar con sentido acerca de las cosas sean preguntas acerca de la estructura de esas cosas –a menos que el concepto de estructura se estire tanto que resulte trivial. Además de las preguntas cuya respuesta requiere un análisis puramente estructural, existen preguntas cuya respuesta exige un análisis ontológico o, más exactamente, ontológico y semántico, o, para emplear una abreviación pedante, *ontosemántico*. Esto es particularmente válido para las preguntas filosóficas.

Tomemos un ejemplo cualquiera para hacer plausible esta idea. A la pregunta «¿Qué es una molécula de hidrógeno?» podemos responder de varias maneras. Es cierto que, para el físico, lo más interesante es una respuesta en términos de estructuras que supuestamente describen la constitución atómica y subatómica de una molécula de hidrógeno. No obstante, esto no es todo lo que uno puede preguntarse respecto a las moléculas de hidrógeno. Hay cuestiones que, aunque suelen interesar más al filósofo que al físico y suelen ser más oscuras, no por ello dejan de tener sentido. Por ejemplo, podemos preguntarnos: ¿A qué tipo de objetos se refiere el término «molécula de hidrógeno»? ¿Debe interpretarse el término «una molécula de hidrógeno» como un objeto físico localizado en el espacio y el tiempo, como una ficción heurísticamente útil, como un conjunto de sensaciones interconectadas del físico perceptor, o, en fin, como qué otra cosa?

¿Son las moléculas de hidrógeno cosas en el mismo sentido en que son cosas las mesas, las sillas y los árboles? ¿Son objetos abstractos o concretos? ¿Depende el significado de «molécula de hidrógeno» de la teoría particular en que se usa o bien es al menos parcialmente independiente de ella?... Estos son ejemplos, a mi entender, paradigmáticos de cuestiones ontosemánticas. Creo que el planteamiento de este tipo de preguntas tiene pleno sentido, pero no dispongo de ningún argumento para justificar esta creencia y debo reconocer que esto tampoco me preocupa mucho. Aquí simplemente voy a presuponer que es válida.

Pues bien, entiendo la pregunta «¿Qué son las teorías científicas?» como cuestión ontosemántica en el sentido ilustrado. Es decir, se trata de subsumir las cosas que llamamos teorías científicas y sus posibles partes componentes bajo alguna o algunas categorías ontológicas previamente asumidas, y ello es concomitante con la tarea de proporcionar un análisis del significado de las expresiones lingüísticas de las teorías y sus componentes. En el presente contexto, el análisis ontológico y el semántico van indisolublemente ligados; de hecho, son las dos caras de la misma moneda. La dilucidación del significado de las expresiones de teorías es la base necesaria para su subsunción bajo cualesquiera categorías ontológicas; y, viceversa, para esclarecer los posibles significados debemos asumir ciertas categorías ontológicas. Es por eso que hablo de un análisis ontosemántico.

Frege se ocupó en toda su obra filosófica de cuestiones ontosemánticas. Se ocupó de ellas a un muy alto nivel de generalidad y a través de un pasmoso esfuerzo por alcanzar la máxima claridad posible. Su objetivo primordial era la fundamentación de las matemáticas; pero pretendió deliberadamente que los elementos básicos de su análisis fueran universalmente aplicables. Y creo que se puede caracterizar su esquema general de análisis como ontosemántico, porque las categorías ontológicas y semánticas propuestas no aparecen por separado, sino formando un todo coherente –o, al menos, ésa fue su intención (cf. R. S. Wells, «Frege's Ontology»).

El alto nivel de generalidad y abstracción en el que se mueve la ontosemántica de Frege quizá no es muy adecuado para responder a todas las cuestiones involucradas en la pregunta «¿Qué son las moléculas de hidrógeno?», pero en cambio creo que sí es apropiado para tratar de la naturaleza de cosas tan generales y abstractas como son las teorías científicas. Es cierto que el esquema ontosemántico general de Frege, en su versión original, contiene algunas oscuridades e imprecisiones. Sin embargo, éstas pueden subsanarse, y en cualquier caso, precisamente para ese nivel de abstracción, el esquema de Frege es todavía hoy por hoy el mejor instrumento de análisis del que disponemos. En este capítulo me dispongo a mostrar algunas de sus potencialidades y dificultades para el caso de las teorías científicas (entiendo por «teorías científicas» las teorías de las disciplinas

llamadas «empíricas» o «factuales»). Pero antes es conveniente resumir los elementos básicos de la ontosemántica fregeana que van a ser utilizados con este fin.

1. El aparato ontosemántico básico de Frege

Las cuatro categorías ontosemánticas básicas de Frege que vamos a usar son las de objeto, función, referencia y sentido. Son categorías irreducibles en el sistema fregeano; sólo se pueden ejemplificar y elucidar, pero no definir. El par ontológico objeto-función es exhaustivo y mutuamente excluyente: todo lo que «hay» es o bien objeto o bien función, pero nada puede ser las dos cosas a la vez. En cambio, el par semántico referencia-sentido no es exhaustivo ni excluyente: además de referencias y sentidos hay otras relaciones entre lo denotante y lo denotado; por ejemplo, lo que Frege llama el «colorido» («*Färbung*») y Dummett el «tono» («*tone*»). Y puede haber sentidos que sean referencias; por ejemplo, en el uso indirecto de las oraciones.

Los objetos son cosas completas, autónomas, existentes por sí mismas; las funciones son cosas incompletas, que para completarse necesitan de objetos y, una vez completadas, dan lugar a nuevos objetos más complejos. Esta división ontológica está inspirada, por supuesto, en la división lógica entre functores y argumentos. Pero esto no es importante ahora.

Hay diversos tipos de objetos. La clasificación más general es entre objetos concretos y objetos abstractos, implícita en Frege, como ha mostrado Dummett. La distinción no es todo lo clara que sería de desear, y parece que hay casos limítrofes entre ambos tipos de objetos (cf. Dummett, *Frege-Philosophy of Language*, p. 480 *et passim*). El criterio que explicita Dummett para distinguir los objetos abstractos de los concretos es el siguiente: un objeto es abstracto cuando su nombre es inteligible *sólo si* existe una expresión funcional que se aplica apropiadamente a ese nombre y a través de la cual lo entendemos (cf. Dummett, *op. cit.*, p. 486 *et passim*). Ejemplos de objetos abstractos importantes para Frege son los números y los valores veritativos. Más importantes en el presente contexto son otros tipos de objetos abstractos: los sentidos, los correlatos objetuales de las funciones (para los que Frege no introduce ningún término especial) y lo que Frege llama los «recorridos» («*Wertverläufe*») de las funciones. Estos últimos se aproximan, aunque no son exactamente iguales, a nuestra noción actual de conjunto. Para nuestros propósitos presentes basta suponer que se trata efectivamente de conjuntos en su sentido actual –ya sea de conjuntos de objetos o bien de conjuntos de tuplos de objetos. A la lista de objetos abstractos tratados por Frege se puede añadir un tipo de objetos que él no considera, pero que son

esenciales aquí: estructuras. Las estructuras son objetos, puesto que son cosas completas, autónomas; y son abstractas, dado que, como veremos ulteriormente, sólo son inteligibles a través de las expresiones funcionales que las «definen» (expresiones de la forma: «... es un espacio vectorial», «... es una mecánica de partículas»). Las estructuras son objetos complejos heterogéneos, en el sentido de que constan de objetos de niveles inferiores: conjuntos y correlatos objetuales de funciones.

Una clase de objetos concretos que nos interesa especialmente son las expresiones lingüísticas. Como ha puesto de manifiesto Shwayder (cf. D. S. Shwayder, «Uses of Language and Uses of Words», p. 137), hay que distinguir en Frege tajantemente entre tres tipos de expresiones lingüísticas: nombres propios (en el sentido amplio de Frege: «*Eigennamen*»); expresiones funcionales («*Funktionennamen*»), y el signo asertórico de la *Begriffschrift*. Este último no nos interesa aquí. Los nombres propios se refieren a objetos, ya sean concretos o abstractos; las expresiones funcionales se refieren a funciones. Es característico de los nombres propios en sentido amplio que empiecen por un artículo determinado o por un demostrativo. En cambio, las expresiones funcionales suelen comenzar con un artículo indeterminado. Por eso dice Frege que, si bien «un triángulo» se refiere a una función (en este caso, un concepto), en cambio, «el concepto de triángulo» no se refiere a un concepto ni a una función, sino a un objeto: al correlato objetual de la función «un triángulo». Se trata, por supuesto, de un objeto sumamente abstracto, para el que no disponemos de nomenclatura particular.

Una clase de expresiones lingüísticas que aquí nos interesarán especialmente son las que se utilizan para formular teorías científicas. Para abreviar, a esta clase de expresiones las llamaremos «formulaciones de teoría». Las formulaciones de teoría pueden darse en un lenguaje formal o bien, lo más frecuente, en un lenguaje natural tecnificado. En cualquier caso, admitiremos que debe haber una forma canónica de formulación de teorías, más allá de las vicisitudes y accidentes que sufren las formulaciones en los libros de texto corrientes. En principio, podemos admitir que esta formulación canónica es la formulación axiomática, aunque más adelante consideraremos la posibilidad de tener que matizar esta asunción.

Hay dos clases de funciones en el sentido de Frege: conceptos y relaciones. Los primeros son funciones monádicas cuyos valores son valores veritativos; las segundas pueden ser de naturaleza muy variada y pueden tener o no valores veritativos entre sus términos.

Entre las relaciones merecen especial mención aquellas que constituyen la referencia y el sentido de las expresiones lingüísticas. Hay una clase de relaciones que relacionan las expresiones con sus sentidos; otras, que relacionan los sentidos con los referentes de las expresiones; otras, en fin, que relacionan las expresiones con sus

referentes (para una reconstrucción formal de estas tres clases de relaciones, cf. J. L. González, *Una reconstrucción conjuntista de la semántica de Frege*). Es una tesis básica de Frege la de que este último tipo de relación resulta de la composición lógica de los dos primeros: es a través del sentido que una expresión logra referirse a su referente; o también: el sentido determina el referente de las expresiones. Al último tipo de relación lo podemos denominar «referencia»; los dos primeros carecen de nombres específicos. Idealmente, cada una de estas clases de relaciones debería ser una clase unitaria: es decir, la relación entre cada expresión y cada sentido, y la relación entre cada sentido y cada referente deberían estar unívocamente determinadas. Sin embargo, en una visión realista de la semántica de los lenguajes naturales no las podemos considerar así, debido a las frecuentes variaciones subjetivas e históricas en el uso de los lenguajes.

Los referentes de las expresiones lingüísticas nominales son objetos; pueden ser objetos de cualquier tipo, concretos o abstractos. Pero no todos los objetos son referentes. A menos que adoptemos una posición idealista o constructivista extrema, que ni Frege ni yo estamos dispuestos a adoptar, parece plausible admitir que hay objetos que no se hallan en la relación de referencia con ninguna expresión lingüística. Puede ser que el lenguaje sea «el límite de *mi* mundo», pero ciertamente no es el límite *del* mundo.

Los sentidos constituyen la categoría de entidades que más quebraderos de cabeza han proporcionado a los comentaristas e intérpretes de Frege. El problema no es tanto que no dispongamos de una buena definición de «sentido»; tampoco se pueden definir las otras categorías básicas de la ontosemántica de Frege y, sin embargo, no nos producen esa sensación de incomodidad. El problema con la noción de sentido es que es más novedosa, mucho menos conectada a la ontosemántica tradicional, y para ella disponemos de poca intuición y entrenamiento en el manejo. No obstante, creo que, con su noción de sentido, Frege aprehendió una idea esencial, imprescindible para cualquier análisis ontosemántico. Lo curioso es que, a pesar de su oscuridad intrínseca, la idea frageana de sentido permite aclarar y detectar muchas cosas. Espero que el análisis subsiguiente de las teorías substancie esta aseveración.

Frege rehuyó comprometerse con una descripción demasiado estricta y universalmente válida de la noción de sentido –aparte de dar caracterizaciones sumamente vagas y poco manejables, como las de que el sentido de una expresión es el modo de darse su referencia (cf. Frege, *Estudios sobre semántica*, pp. 50-51). Prefirió aclarar su noción de sentido a base de innumerables ejemplos, tanto del discurso cotidiano como del matemático. En realidad, este método de aclaración de conceptos por medio de ejemplos paradigmáticos de aplicación es, como sabemos en nuestra era postwittgensteineana, perfectamente admisible. No obstante, dada la importancia de la noción para

el presente análisis, podemos tratar de aproximarnos un poco más a una caracterización general.

Dummett se ha ocupado con mucho detenimiento de elucidar la noción de sentido. Quisiera recordar algunas de sus acotaciones, que son especialmente valiosas para el propósito presente de indagar sobre el sentido de las formulaciones de teoría. El valor esencial del sentido, según Dummett, es su valor *cognitivo*: el sentido determina el modo como podemos averiguar el referente de una expresión: «*... la noción de sentido, tal como Frege la entendió, se relaciona con nuestro reconocimiento de la referencia*» (cf. Dummett, *op. cit.*, p. 589). Más exactamente, la consideración del sentido, la fijación precisa del sentido de las expresiones, cobra su máxima importancia cuando se trata del problema de *justificar* un conjunto de enunciados, cuando queremos demostrar o al menos hacer plausible que el referente de cierta(s) expresión(es) es efectivamente el que decimos que es (cf. Dummett, *op. cit.*, p. 104). Esta comprensión del sentido aquí nos interesa especialmente cuando se trata de enunciados científicos (teóricos); Dummett apoya la idea de que, en este caso, el sentido de los enunciados viene dado por sus interconexiones.

En su contracrítica al holismo extensionalista de Quine, por ejemplo, comenta:

> «La imagen quineana del lenguaje como red articulada, por tanto, sólo es aceptable si se interpreta, no holísticamente, negando un sentido a las oraciones individuales, sino de la manera natural, reconociendo la existencia de conexiones inferenciales definidas entre las oraciones, *de modo que el sentido de una oración no-periférica consiste precisamente en sus conexiones con otras oraciones*» (Dummett, *op. cit.*, p. 608, subrayado mío).

Esta idea jugará un papel importante en el análisis de las teorías a continuación.

2. Aplicación del aparato a la concepción enunciativa de las teorías

Las tres preguntas centrales que nos proponemos atacar dentro del esquema general de Frege son:

1. ¿Qué es una teoría? ¿Es un objeto lingüístico, un objeto no-lingüístico o una función?
2. ¿Cuál es la referencia de una formulación de teoría y la de sus partes componentes?
3. ¿Cuál es su sentido?

Las tres preguntas no son independientes entre sí, de modo que puedan ser tratadas por separado. En realidad, una respuesta adecuada a cada una de ellas depende de respuestas adecuadas a las demás.

En lo que sigue, consideraremos tres concepciones distintas y actualmente vigentes de las teorías científicas: la concepción estándar, «clásica» o «enunciativa»; la concepción de Suppes y sus colaboradores, y la concepción de Suppes revisada por Adams y Sneed al introducir ciertos elementos nuevos. No examinaré concepciones más complejas y elaboradas, como la estructural de Sneed-Stegmüller en su versión madura, ni la de Ludwig. El motivo de esta omisión es simplemente la falta de tiempo y espacio. No obstante, creo que la omisión no es grave, dado que gran parte de lo que podría decirse al considerar estas concepciones más elaboradas estaría ya contenido en lo dicho para las primeras.

La aplicación del aparato ontosemántico de Frege a los tres enfoques a consideración nos permitirá no sólo dar respuestas precisas a las tres preguntas anteriores dentro de cada uno de los enfoques, sino además aprehender las diferencias esenciales entre ellos. Veremos que, gracias a las clarificaciones que permite Frege, estas diferencias son más significativas de lo que a primera vista pudiera parecer.

La concepción estándar, que podríamos llamar también «concepción enunciativa», es la más simple. Según ella, las teorías son enunciados. (En el presente contexto es indiferente que tomemos cada teoría como un conjunto de enunciados o bien como una conjunción de enunciados.) En tal caso las formulaciones canónicas (es decir, axiomáticas) de las teorías coinciden con las teorías mismas. Las teorías son así objetos lingüísticos y, por tanto, objetos concretos. Son objetos que están escritos en alguna parte o que al menos se pueden escribir en alguna parte.

Dado que las teorías son enunciados, sus referentes serán, según una tesis famosa de Frege, valores veritativos. Las teorías son verdaderas o falsas.

Los enunciados que constituyen una teoría según la concepción estándar son axiomas, puesto que hemos convenido en que la formulación canónica de teorías es la axiomática. Ahora bien, los axiomas de cualquier teoría científica propiamente tal son enunciados al menos parcialmente teóricos, no directamente observacionales, o, dicho en la terminología de Quine, enunciados que no pertenecen a la «periferia» del *corpus* científico. Si aplicamos a este caso el criterio de Dummett arriba mencionado de que «el sentido de un enunciado no-periférico consiste precisamente en sus conexiones con otros enunciados», es fácil ver cuál es el sentido de una teoría: el sentido de cada uno de sus axiomas no puede determinarse aisladamente, sino que depende del modo como está conectado con el resto de enunciados de la teoría, o sea, de sus conexiones axiomáticas; en otras palabras, el sentido de una teoría es precisamente el modo concreto en que aparece axiomatizada. La axiomatización determina el sentido de una teoría. Por eso podemos decir que en una formulación pre-axiomática

de una teoría el sentido de la misma es vago, no bien definido. Otra consecuencia interesante que podemos sacar de esta idea es que, prescindiendo del caso de formulaciones sintácticamente equivalentes, axiomatizaciones distintas de «la misma» teoría expresan sentidos distintos y por tanto *no* son, estrictamente hablando, de *la misma* teoría. Así, podemos decir que las axiomatizaciones de Pasch y de Hilbert de la geometría euclídea le dan sentidos distintos a ésta. Lo mismo ocurre con las distintas axiomatizaciones existentes de la mecánica clásica, de la relatividad especial o de la termodinámica. Ciertamente, tenemos la intuición de que en todos estos casos se trata de distintas axiomatizaciones de «lo mismo». Pero esta intuición no está justificada por la noción de sentido, ni tampoco, por cierto, por la de referencia. Pues el referente de una gran cantidad de teorías puede ser el mismo, por ejemplo lo falso, y, sin embargo, no diremos que por este solo hecho se trata de axiomatizaciones de «lo mismo» en todas ellas. Parece claro que para poder decir que dos axiomatizaciones no-sintácticamente equivalentes lo son de lo mismo, por ejemplo de la mecánica clásica o de la geometría euclídea, debemos acudir a nociones de la teoría de modelos. Ahora bien, no parece haber un modo natural de combinar estas nociones con la concepción enunciativa dentro del aparato de Frege. Parece que éste es un síntoma de que cuando queremos aplicar dicho aparato a la consideración de los modelos asociados a una teoría debemos romper el marco de la concepción estándar y pasar a una idea más estructural de las teorías, por la que los constituyentes básicos de una teoría no son enunciados, sino estructuras enteras. Sin pretensiones de dejar esta cuestión definitivamente asentada, prepararemos el terreno para su posible resolución en las secciones siguientes.

¿Qué decir de los componentes de una teoría? En la concepción que estamos examinando éstos son, aparte de las constantes lógicas y prescindiendo de los términos puramente matemáticos, los términos primitivos específicos de la teoría en cuestión y que se admite que tienen una interpretación (es decir, referencia) empírica. Ahora bien, según la axiomática moderna, no tiene sentido preguntarse por la referencia de los términos primitivos tomados éstos aisladamente; ésta sólo puede determinarse interpretando el sistema axiomático en bloque sobre algún dominio. En tal caso, el método de interpretación para fijar la referencia será igual o paralelo al utilizado en la moderna semántica estándar. Será igual si nos desviamos de la intención original de Frege y admitimos que los referentes de las expresiones funcionales no sean funciones, sino las extensiones conjuntistas de las mismas; será paralelo en caso de apegarnos estrictamente a Frege.

Si nos preguntamos ahora por el sentido de los términos primitivos de la teoría, la respuesta sería fácil si se tratara de una teoría de la matemática pura, más concretamente, de una teoría axiomatizada a lo Hilbert. En tal caso lo natural sería admitir sin más argumento que el

sentido de los términos primitivos, al igual que los axiomas mismos, vendrá determinado en bloque y exclusivamente dentro del sistema axiomático por las interconexiones entre los términos expresadas en los axiomas; es decir, el sentido del término primitivo t_1 vendrá dado por las relaciones postuladas axiomáticamente entre t_1 y los demás términos primitivos de la teoría, t_2, t_3, etc.

Sin embargo, esta determinación del sentido de los términos en teorías axiomatizadas, que parece obvia en el caso de la matemática pura, ya no lo es tanto en el caso de teorías empíricas, aunque éstas se presenten como estructuras matemáticas. Naturalmente, parece claro que en ellas las interconexiones axiomáticas entre los términos tendrán algo que ver con el sentido de los mismos; podemos decir que lo determinarán *en parte*. Pero no parece plausible decir que el sentido de los términos primitivos de teorías empíricas queda *completamente* determinado por sus interconexiones axiomáticas. Hay más cosas que decir al respecto.

En primer lugar, está claro que el sentido de los términos de una teoría empírica (y esto es justamente lo que la hace *empírica*) no queda agotado al considerar los axiomas de la teoría. Para una determinación completa del mismo hay que salir «fuera» de la teoría. Recuérdese la receta de Dummett: el sentido es lo que nos permite *reconocer* la referencia; y está claro que en las teorías empíricas este reconocimiento no puede darse sólo echando una ojeada a las interconexiones axiomáticas de los términos. El modo como reconocemos la referencia es, en parte al menos, a través de la «experiencia». De algún modo que aún hay que precisar, la experiencia interviene en la determinación del sentido de los términos empíricos.

Para dar cuenta de esta situación, la concepción estándar de las teorías empíricas introdujo la distinción entre dos niveles: el observacional y el teórico. En las ciencias empíricas habría así dos clases de términos: los observacionales y los teóricos. Y la determinación del sentido no puede ser la misma en uno y otro caso. El sentido de los términos observacionales es distinto del sentido de los términos teóricos. ¿Cuál es uno y otro?

Para tratar con detalle esta cuestión ahora deberíamos embarcarnos en una discusión exhaustiva de la famosa dicotomía «teórico-observacional», de todos los problemas que están involucrados en ella y de las distintas versiones de la misma que se han propuesto. La respuesta a la última pregunta formulada evidentemente no puede ser la misma cuando nos basamos en una interpretación absoluta de la dicotomía (como en el Carnap «medio» y en Braithwaite) que cuando aceptamos la versión más «liberal» del último Carnap y de Bar-Hillel (según la cual la distinción entre términos observacionales y teóricos no es sino un «corte en un continuo»), o finalmente reinterpretamos la dicotomía al estilo «histórico-cultural» del último Hempel: que un término sea considerado teórico u observacional depende del grado de

entrenamiento y hábito en su uso por parte de la comunidad científica. Discutir el modo como estas distintas versiones de la dicotomía influyen concretamente en la concepción del sentido de los términos empíricos nos llevaría ahora demasiado lejos. Baste de momento constatar lo siguiente. A todas estas versiones es común la idea de que el sentido de los términos observacionales se determina por la experiencia cotidiana preteórica (o por lo que, *relativamente*, puede considerarse como tal). En cambio, el sentido de los términos teóricos vendría en parte determinado por sus interconexiones dentro de las leyes de la teoría, ciertamente, pero también por las *reglas de correspondencia* que los unen a los términos observacionales. (A través de las reglas de correspondencia «fluiría» parte del sentido de los términos observacionales hacia los teóricos.) En ningún caso, sin embargo, ni en los términos teóricos, ni menos aún en los observacionales, el sentido viene determinado exclusivamente por los axiomas de la teoría.

3. Aplicación del aparato a la concepción suppesiana de las teorías

Creo que no es forzado interpretar los trabajos de Suppes y sus colaboradores en la axiomatización de las teorías empíricas como una desviación sustancial de la concepción estándar –desviación de la que el propio Suppes quizá no fue plenamente consciente en un principio, pero que se manifiesta claramente al aplicarle la ontosemántica de Frege. A partir de los trabajos de Suppes, ya no es natural concebir una teoría como un enunciado, ni siquiera como una expresión lingüística cualquiera. A diferencia del caso de la concepción estándar, aquí hay que distinguir claramente la formulación de teoría de la teoría misma. La formulación de teoría es ahora de la forma «... es un *P*», donde *P* es un predicado conjuntista; por ejemplo: «... es una mecánica clásica de partículas». Su referente, la teoría, es un concepto y, en consecuencia, una función. Así, pues, frente a la concepción estándar, según la cual una teoría es un objeto, en Suppes una teoría es una función. Claro que se trata de una función, o concepto, de tipo especial –de un tipo en el que no parece haber pensado Frege. Esto se ve al preguntarnos cuáles son los objetos que caen bajo ese concepto que llamamos «teoría»: no son ni objetos físicos, ni números, ni valores veritativos. Los objetos que completan la entidad funcional que llamamos, por ejemplo, «mecánica clásica de partículas», son cosas complejas y heterogéneas a las que en la literatura se las suele llamar de varias maneras: «estructuras», «modelos» o «sistemas». Mosterín ha enfatizado la conveniencia de usar el término «sistema» en vez del de «estructura» para tales objetos, aunque el último sea más corriente. Aquí me atengo a esta propuesta. Claro que el nombre es lo de menos. Lo importante es saber de qué se trata. Los sistemas

de las teorías empíricas de los que aquí se trata son tuplos que constan típicamente de uno o varios conjuntos de objetos más cierto número de funciones. Los conjuntos suelen ser conjuntos de objetos físicos como partículas y estados, y de objetos matemáticos como números y vectores. Las funciones pueden ser propiedades, relaciones y sobre todo funciones métricas. Los sistemas son, pues, por principio, objetos heterogéneos. Con ello obtenemos la configuración ontológica básica de una teoría en el enfoque de Suppes. Simplificando la jerarquía ontológica de este esquema, podríamos decir que una formulación de teoría se refiere a un concepto bajo el cual caen a la vez objetos (conjuntos físicos y matemáticos) y funciones.

Más difícil es determinar el sentido de una formulación de teoría a lo Suppes. ¿Cuál es el sentido de una expresión conceptual del tipo «...es una mecánica clásica de partículas»? Es sabido que la idea de que las expresiones predicativas y, en general, funcionales, tienen un sentido propio ha sido muy controvertida. Algunos comentaristas de Frege lo han interpretado como si él estuviera dispuesto a identificar sentido y referencia para ese tipo de expresiones. Creo que Dummett ha mostrado convincentemente que esta interpretación es incoherente con el aparato general de Frege. Por otro lado, lo cierto es que Frege no nos proporciona un criterio general para el sentido de las expresiones funcionales. Frege dice (cf. «Sobre el sentido y la referencia», *Estudios sobre semántica*, p. 51) que el sentido en general es el modo de darse la referencia a través de la expresión lingüística. Pero este criterio aquí no nos sirve, porque el modo de darse de las teorías es siempre el mismo: a través de un predicado conjuntista. A primera vista, pues, el sentido de las formulaciones de teoría de Suppes es opaco; lo que realmente marca la diferencia de sentido entre dos formulaciones de teoría distintas está implícito en ellas: está contenido en las condiciones de definición de los predicados conjuntistas respectivos. Ellas no se hallan presentes en la formulación misma de teoría. Ahora bien, si aceptamos la interpretación de Dummett (cf. *op. cit.*, p. 153) acerca del sentido de las expresiones definidas, según la cual el sentido del *definiens* se traslada íntegramente al sentido del *definiendum*, y tenemos en cuenta que el predicado conjuntista que formula una teoría viene determinado por las condiciones de definición (que son los axiomas en la concepción estándar), entonces parece claro que las interconexiones axiomáticas postuladas en las condiciones de definición constituyen el sentido del predicado conjuntista de la teoría.

En cuanto al sentido y a la referencia de los términos que aparecen en los tuplos que expresan los sistemas que son modelos de la teoría, a ellos se puede aplicar en principio lo mismo que ya hemos dicho para los términos primitivos en la concepción enunciativa. Es decir, la referencia viene dada por una interpretación sobre algún universo empírico de modo igual o similar a lo que se hace en la semántica

formal estándar, mientras que su sentido viene determinado en parte por las interconexiones axiomáticas del predicado conjuntista y en parte por la experiencia preteórica, ya sea de forma inmediata para los términos que aparecen en enunciados observacionales, o bien de forma mediata, a través de reglas de correspondencia, para los términos teóricos. Ahora bien, la formulación de teoría a lo Suppes nos hace ver un aspecto interesante de la relación entre sentido y referencia de los términos primitivos que no se manifiesta claramente en la concepción enunciativa. Al formular una teoría mediante un predicado conjuntista, resulta de inmediato claro que ese predicado puede aplicarse igualmente a multitud de cosas diversas; en general, pues, habrá un número indefinido de sistemas distintos, todos los cuales son modelos de la teoría. Para emplear un neologismo, podríamos decir que la concepción suppesiana nos lleva directamente a un «plurimodelismo» de las teorías empíricas. Como en general las teorías empíricas no son categóricas, esa multitud de modelos es realmente heterogénea. Y todos son, en principio, igualmente interesantes para el teórico empírico.

Pues bien, partiendo de esta idea plurimodelista, llegamos a una visión algo sorprendente de la compleja relación que existe, en el caso de teorías empíricas axiomatizadas conjuntistamente, entre referencia, sentido e identidad de los términos de una teoría. Por un lado, la referencia de los términos primitivos varía, evidentemente, de un modelo a otro de la teoría (pues se refieren a distintos sistemas físicos). Por ejemplo, el término «partícula», que aparece como primitivo en una axiomatización de la mecánica clásica de partículas, puede referirse en un modelo a una galaxia, en otro a un planeta, en el tercero a un péndulo y en el último a una molécula de gas. Los referentes en todos estos casos son muy distintos entre sí desde el punto de vista ontológico. Incluso, según la interpretación radical de Suppes de su propio proceder, pueden pertenecer a categorías ontológicas «inconmensurables»; en efecto, según reconoce Suppes explícitamente, un modelo efectivo de su axiomatización de la mecánica clásica de partículas puede ser tal que las «partículas» sean números naturales (tres modelos de este tipo se usan en la axiomatización original de McKinsey, Sugar, Suppes). Quizá también podría haber modelos de esta teoría cuyas «partículas» fueran personas, ángeles o quién sabe qué. Cualquier cosa puede ser, en principio, el referente de «partícula», siempre y cuando satisfaga las condiciones de definición del predicado «es una mecánica clásica de partículas». Nada debe tener en común, ontológicamente, la referencia en los diversos modelos.

En cambio, el caso del sentido es distinto. Ciertamente, *una parte* del sentido del término «partícula» en los ejemplos indicados será diversa; es la parte «experiencial» del sentido. El modo de experimentar el referente de «partícula» cuando se trata de un planeta o cuando

se trata de una molécula es manifiestamente distinto, para no hablar de cómo «experimentamos» números o ángeles. Pero ya sabemos que el sentido no se agota en este aspecto. Sabemos que el sentido también viene determinado en parte por las interconexiones axiomáticas entre los términos. Y éstas son *comunes* a todos los modelos de la teoría. Siempre son las mismas, justamente por tratarse de modelos de la *misma* teoría (=del mismo predicado conjuntista). Puede decirse que son precisamente los axiomas (=las condiciones de definición del predicado) lo que da *unidad de sentido* a esa multitud heterogénea de modelos, con sus diversas referencias. Aunque el término «partícula» pueda referirse a cosas muy diversas en la mecánica, tiene parcialmente el mismo sentido en todos los casos, y esto es precisamente lo que nos justifica en usar el mismo término en todos los modelos (lo cual, de lo contrario, si no hubiera ese sentido parcialmente-unitario, aparecería como una arbitrariedad terminológica). Lo mismo ocurre con los términos funcionales de la teoría. La referencia de «masa», pongamos por caso, es distinta en cada modelo, pero su sentido coincide en parte en todos los modelos. Por eso hablamos simplemente de «masa» en todos los casos y no de «masa-en-el-modelo-x»; y por eso es legítimo escribir tratados con títulos como «El concepto de masa».

4. Aplicación del aparato a la concepción estructural sneediana

Como se ha visto en capítulos anteriores, el programa sneediano de reconstrucción de teorías empíricas adopta el método de Suppes de axiomatizar teorías empíricas definiendo predicados conjuntistas, pero también sabemos que ese programa introduce muchos más elementos de análisis de la estructura de las teorías que la concepción original de Suppes. Algunas de estas complejidades adicionales son claramente relevantes para el problema de la determinación ontosemántica de las teorías y sus componentes, e implican una aplicación mucho más matizada del aparato de Frege que en los casos anteriores. En particular, hay dos ideas sneedianas que son especialmente importantes en el contexto que nos ocupa: primera, que la identidad de una teoría queda determinada no sólo por su predicado conjuntista correspondiente, sino también por la acotación del dominio de aplicaciones propuestas; segunda, la distinción funcional entre términos T-teóricos y T-no-teóricos dentro de cada teoría T. La primera idea abre una nueva perspectiva acerca de la referencia y el sentido de las formulaciones de teoría; la segunda es pertinente para el sentido de los términos primitivos de las formulaciones. Trataremos de ambas cuestiones sucesivamente.

Para tratar la primera cuestión en su especificidad vamos a olvidarnos de la distinción entre T-teórico y T-no-teórico (y, por

tanto, entre M_p y M_{pp}) y considerar solamente en qué medida las categorías ontosemánticas quedan afectadas por la introducción de un dominio de aplicaciones propuestas. Recordemos dos características de este dominio: a) que no corresponde a la totalidad de los modelos efectivos de la teoría (muchos sistemas que son formalmente modelos no son aceptados como aplicaciones «auténticas»), sino que corresponde, en todo caso, a un subconjunto del conjunto de los modelos, y b) que en la delimitación del dominio y, por tanto, en la identificación misma de la teoría, entran en juego consideraciones pragmáticas e intuitivas, que no tienen correlato en las teorías de la matemática pura.

Pues bien, al restringir de este modo el campo de aplicación del predicado conjuntista por medios externos al mismo, es decir, medios que no están contenidos ni explícita ni implícitamente en él, se modifica sustancialmente tanto la referencia como el sentido de la formulación de teoría correspondiente. En efecto, si adoptamos el criterio heurístico de Frege de que para identificar conceptos basta con identificar las extensiones correspondientes (aunque no sean lo mismo), entonces, al introducir la idea de aplicaciones propuestas, estamos introduciendo un elemento completamente externo al predicado de la teoría para determinar la extensión del concepto de la teoría y, por tanto, la referencia misma de la formulación de teoría. Dentro del esquema general de Frege, esto significa que la delimitación de las aplicaciones propuestas constituye parte del *sentido* de la formulación de la teoría. Pues es a través de la delimitación de las aplicaciones propuestas que llegamos a la determinación de la referencia de la formulación de teoría. Y digo que constituye solamente *parte* del sentido porque naturalmente las condiciones de definición del predicado conjuntista seguirán interviniendo en la constitución del sentido, como hemos visto al hablar de la concepción suppesiana original.

Si aceptamos esta idea acerca del sentido, podemos comprender mejor algunos puntos de la concepción sneediana de las teorías que han sido objeto de largas y complicadas discusiones. En primer lugar, se nos hace natural que el dominio de aplicaciones propuestas sea una entidad inherentemente vaga. Pues es típico de los sentidos, como objetos, que sus límites sean imprecisos y que no dispongamos de criterios tajantes de identidad para ellos. Esto es precisamente lo que ocurre con el dominio de aplicaciones propuestas. Pero esta vaguedad no debe ser valorada como algo negativo, sino que, como ha subrayado Stegmüller, es precisamente a través de un grado razonable de vaguedad que funcionan las teorías empíricas. Lo mismo podríamos decir en general de los sentidos en la mayoría de contextos. Un grado razonable de vaguedad en el sentido hace al lenguaje más flexible.

En segundo lugar, ahora aparece más plausible el criterio sneedia-

no de identidad de teorías. Una idea controvertida de Sneed es que dos teorías físicas pueden estar expresadas mediante el mismo predicado conjuntista, definido por los mismos axiomas, y, sin embargo, ser dos teorías distintas, por ser distintos su dominios de aplicaciones propuestas. Por ejemplo, podemos tener dos teorías, una de hidrodinámica y otra de electrodinámica, con exactamente el mismo predicado conjuntista (los mismos axiomas) y no obstante admitir intuitivamente que se trata de teorías diferentes. Podemos decir en tal caso que la misma formulación de teoría se refiere a conceptos (teorías) distintos, pues es por medio de las aplicaciones propuestas que determinamos la referencia, y ellas son distintas en uno y otro caso.

Finalmente, si admitimos, como creo que hay que hacer, que el dominio de aplicaciones propuestas no es un objeto estático, dado de una vez por todas, sino una entidad genidéntica, entonces podemos dar un contenido preciso y plausible a una afirmación muy difundida en la reciente filosofía de la ciencia: la de que el sentido de una teoría cambia con su evolución histórica. En efecto, si el sentido está en parte constituido por el dominio de aplicaciones propuestas y si este dominio varía normalmente porque le añadimos o quitamos elementos, entonces está claro cómo y por qué el sentido de las teorías, y por tanto también su referencia, varía con el tiempo.

Pasemos ahora a la consideración del sentido de los términos primitivos. La dicotomía sneediana entre términos T-teóricos y T-no-teóricos, que tiene un carácter muy distinto al de la dicotomía teórico-observacional, aporta un elemento nuevo en la determinación del sentido de los términos. Este elemento nuevo son las relaciones interteóricas, fundamentalmente la relación de *presuposición* (y quizá de teorización), a la que se ha aludido al final del capítulo 2.8. Hemos visto en la discusión precedente del sentido de los términos que éste viene determinado en parte por los axiomas y en parte por su relación con la experiencia preteórica. Dejemos a esta última de momento a un lado, porque la metateoría estructural no tiene nada que decir sobre cuestiones epistémicas. Entonces quedarían solamente las interconexiones axiomáticas como determinantes del sentido de los términos. Pero esto es debido a que en las concepciones clásica y suppesiana de las teorías se analizan estas últimas como unidades aisladas. Esto ya no es posible en la metateoría estructural. En efecto, los términos T-no-teóricos son, por definición, los que *provienen* de teorías subyacentes a T, presupuestas por T. La determinación de los términos T-teóricos, por el criterio de teoricidad de Sneed, presupone sólo la teoría T en cuestión. Por ello puede decirse que la determinación de esos parámetros es «asunto interno» de T; su sentido viene dado sólo por los axiomas de T. En cambio, la situación es esencialmente distinta para los términos T-no-teóricos. Su determinación, al menos en algunas aplicaciones, presupone teorías subyacentes a T. Su

sentido depende, por tanto, no sólo de las interconexiones axiomáticas dentro de la propia *T*, sino de los axiomas de alguna teoría *T'* subyacente y previa a *T*. Incluso puede decirse que ésta es la forma primordial de determinación del sentido de los términos *T*-no-teóricos: a través de una instancia *externa* a *T*. (Esta instancia también es una teoría; pero es *otra* teoría.) De ello se desprende que, partiendo de la concepción estructural sneediana, no se puede plantear la determinación del sentido de los términos sin mayores cualificaciones, sino que hay que especificar si se trata de términos *T*-teóricos o *T*-no-teóricos. El sentido de los términos *T*-teóricos es un sentido «*T*-eico», puesto que son términos *específicos* de *T*, mientras que el sentido de los términos *T*-no-teóricos no es «*T*-eico» (o sólo lo es de manera subsidiaria). Así, en la termodinámica, el sentido de los términos «entropía» y «energía interna» es un «sentido termodinámico», mientras que el sentido de «presión» es primordialmente hidrodinámico y el de «volumen» es geométrico. Ante este ejemplo concreto, quizá se piense que hemos dado un gran rodeo formal para llegar a una conclusión intuitivamente banal; pero el mérito de un aparato filosófico general, como es la ontosemántica de Frege, puede consistir precisamente en que nos hace ver de manera formal, sistemática, por qué es efectivamente correcta una intuición aparentemente banal.

4.2. LOS CONCEPTOS DE MATERIA Y SUS DIFICULTADES

Una tendencia corriente dentro del pensamiento contemporáneo ha sido la de establecer una dicotomía tajante entre las actividades científicas y las filosóficas. Un número importante de filósofos y científicos han sostenido que ciencia y filosofía no tienen nada que ver una con otra, que los problemas y objetos de estudio de ambas disciplinas pertenecen a categorías enteramente distintas y que tratar de establecer puentes o modos de influencia entre una y otra es sólo una forma de hacer mala ciencia o mala filosofía. Esta actitud de mutua ignorancia entre ciencia y filosofía puede que tenga la virtud de hacerles la vida más fácil a ciertos científicos y a ciertos filósofos que tienen graves dificultades para salir de su propia estrecha área de especialización y para asimilar el lenguaje de otras áreas. Sin embargo, hacerse la vida fácil no debe ser el lema ni de la buena ciencia ni de la buena filosofía, y la política de mutua ignorancia sólo lleva al oscurecimiento o negligencia de una serie de problemas que conciernen tanto a científicos como a filósofos. Es simplemente un hecho que el conjunto de conceptos y problemas tratados por las ciencias y el conjunto de conceptos y problemas tratados por la filosofía tienen una intersección no vacía. Un caso típico que, a mi entender, muestra claramente que hay una zona de empalme entre filosofía y ciencia, es el análisis del concepto de materia. Este concepto y la problemática ligada a él pertenecen a ese ámbito de ideas y cuestiones que, por su generalidad y por el papel que han jugado en nuestra evolución científica y cultural, no es posible catalogar ni como estrictamente filosóficos ni como estrictamente científicos; y, por otra parte, me parece claro que hay que tomarse en serio ese ámbito

conceptual y no anatematizarlo como conjunto de sin-sentidos, como algunos filósofos estarían tentados de hacer.

Esa zona conexión entre ciencia y filosofía produce algunos de los problemas más interesantes para cualquier intento de comprensión global del mundo y de nuestro conocimiento del mundo. Muchos de esos problemas, aunque no todos, son de la clase que podríamos denominar «ontológicos», es decir, referentes a nuestra concepción de «lo que hay» en general. El análisis del concepto de materia pertenece a esta clase.

En lo que sigue me propongo analizar el concepto de materia o, mejor dicho, distintos conceptos posibles de materia, sus características básicas y sus dificultades, y expondré este análisis desde el punto de vista que he denominado ontológico, con lo cual me propongo no sólo examinar ese caso concreto, sino también ejemplificar una determinada concepción de lo que es o debe ser la ontología.

El dilema metodológico básico que se plantea en cualquier elucidación ontológica y, en este caso concreto, en la elucidación del concepto de materia es el siguiente. Por un lado, queremos obtener una caracterización general de la materia (o de cualquier otra noción ontológica fundamental) que sea lo más resistente posible ante los resultados empíricos concretos, pues no queremos construir ontologías que se vengan abajo cada par de años ante nuevos experimentos o cálculos; queremos una ontología y, en particular, un concepto de materia que resista lo mejor posible los embates del desarrollo científico. Por otro lado, naturalmente, tampoco queremos proponer un concepto de materia que sea una pura elucubración sin contenido empírico, ajena a las teorías científicas en vigencia; esto sería refugiarse en un discurso puramente especulativo sin conexión con la ciencia, lo cual, a mi entender, es hacer mala ontología. Seguir demasiado fielmente las teorías científicas del momento es suicida (además de acrítico); no seguirlas en absoluto es estéril. Este es el dilema del ontólogo. Por supuesto, la resolución de esta tensión estaría en seguir alguna vía media. El problema consiste en averiguar cuál es esa vía en términos concretos.

Creo que, por lo menos en el caso que nos ocupa, la estrategia adecuada no es dar una *definición* de la materia con pretensiones de validez absoluta, sino considerar varias alternativas posibles de elucidación del concepto de materia. Es decir, la pregunta que se trata de plantear y responder no es tanto la de «¿Qué es la materia?», sino más bien la de «¿Qué puede ser la materia?» o, dicho de otro modo, «¿Qué aplicaciones o usos del predicado "es material" son plausibles?» Con esta forma de plantear la pregunta adoptamos de entrada un punto de vista abierto a varias concepciones posibles de la materia.

Nuestra tarea consistirá entonces no en dar una definición única y tajante, sino en precisar las distintas alternativas tanto como sea posible y examinar su grado de plausibilidad ateniéndonos a nuestras

intuiciones previas, a los resultados científicos aceptados y a consideraciones de coherencia lógica. Podemos adelantar que, ante ese examen, veremos que todas las alternativas consideradas se enfrentan a problemas conceptuales de ninguna manera triviales. El hecho de que cualquier elucidación del concepto de materia posea aspectos problemáticos no significa que ese concepto carezca de sentido o que sea irremediablemente confuso. Pero el ontólogo no-dogmático (por desgracia, una «rara avis») debe reconocer y encarar esos problemas antes de tomar partido. Esa es la tarea que se emprende a continuación; en resumidas cuentas: precisar las distintas alternativas posibles de elucidación del concepto de materia y examinar sus dificultades.

Sin embargo, antes de entrar en la problemática propiamente dicha, es conveniente decir dos palabras acerca de por qué estimo importante el problema de elucidar el concepto de materia en la forma en que lo haremos aquí, es decir, desde el punto de vista ontológico. Existen dos motivaciones básicas para esta empresa:

a) *Motivación ideológica general.* La elucidación del concepto de materia subyace (o debería subyacer) a cualquier discusión sobre la ideología filosófica que llamamos «materialismo», que, por mi parte, entiendo fundamentalmente como una doctrina ontológica monista. El hecho de que en la mayoría de discusiones y estudios sobre el materialismo no se examine primero el concepto de materia de manera adecuada sólo habla en contra del método seguido en tales discusiones, no en contra de la necesidad, a mi entender obvia, de empezar con un examen del concepto de materia. Un materialista que no nos ofrezca un concepto razonablemente claro y preciso de materia antes de abogar por su doctrina ontológica es un materialista que no sabe lo que está defendiendo. (Este será uno de los puntos esenciales expuestos en el capítulo siguiente.)

b) *Motivación metodológica particular.* Una de las tareas básicas de la filosofía de la ciencia es analizar los supuestos semánticos y ontológicos de las teorías científicas. Esto se ha hecho poco hasta la fecha; la mayoría de los filósofos de la ciencia se han concentrado en el análisis estructural de las teorías. Pero esto puede y debe cambiar en el futuro. Sin olvidar los importantes resultados obtenidos en la reconstrucción de las estructuras de las teorías científicas, actualmente hemos alcanzado una fase en la que conviene complementar el estudio puramente estructural con el análisis de los supuestos ontosemánticos de las teorías científicas. Ahora bien, aunque el concepto de materia no aparece explícitamente en las formulaciones estándar de la gran mayoría de teorías científicas, muchos científicos y filósofos piensan que «materia» es una de las nociones básicas que intervienen en la caracterización del dominio al que se refieren la mayoría, si no la totalidad, de las teorías científicas, especialmente las de la física. Para

indagar hasta qué punto esta opinión es correcta y en qué sentidos se presupone alguna noción de materia en las teorías físicas, es preciso comenzar con un recuento sistemático de las formas posibles en que esa noción puede entrar en la base ontológica de dichas teorías.

Hasta donde alcanzo a ver, el concepto de materia ha aparecido esencialmente dentro de cuatro esquemas conceptuales distintos, que han ejercido y aún ejercen una gran influencia en el desarrollo de las discusiones científicas y filosóficas. Los cuatro esquemas son otras tantas formas posibles de usar ese concepto o, mejor dicho, ese *término* en la construcción de una ontología general. Naturalmente, estos cuatro esquemas no siempre han aparecido en forma clara y bien delimitada, a veces se han mezclado entre sí o han adoptado subformas especiales. Sin embargo, al menos como primera aproximación, me parece válido y útil identificar esas cuatro formas posibles de hablar sobre la materia.

Primer esquema: materia como substrato común universal

Esta es la forma de entender la materia más elemental y menos dependiente del desarrollo científico. Dentro de este esquema la materia aparece como el denominador común de todas las cosas reales o existentes. Y dado que, a primera vista al menos, las cosas que aparecen a nuestro alrededor son muy distintas entre sí y poseen propiedades no sólo diferentes, sino hasta mutuamente incompatibles, este esquema supone que, a pesar de las apariencias en contra, hay un *substrato* común a todas las cosas, aunque no lo percibamos directamente. Es decir, se supone que hay una realidad «*detrás*» o «*por debajo*» de lo que percibimos, y que aunque no la percibamos podemos estar seguros de su existencia (no siempre se dice claramente cómo y por qué podemos estar seguros de ella). En resumen, se trata de lo que podríamos denominar la teoría del *substrato universal recóndito*.

Como ha puesto de relieve Stephan Toulmin en su artículo sobre la materia en la *Encyclopedia of Philosophy*, este concepto, por lo demás muy inespecífico, se retrotrae a la «hylé» de los griegos y es la noción en la que piensa vagamente el hombre común no especializado en ciencia o filosofía cuando habla de materia.

Esta concepción de la materia es también la más estrechamente ligada tradicionalmente al materialismo como doctrina ontológica monista, ya que, en general, para esta doctrina se considera relativamente irrelevante qué especificaciones del concepto de materia se adopten: lo único que importa es que se admita la existencia de ese substrato recóndito y universal.

Tratemos de precisar este esquema para facilitar su discusión. Para ello hemos de presuponer que entendemos lo que quiere decir que algo

es real, o sea, que conocemos los usos del predicado «es real». El término «materia» dentro de este esquema es entonces un predicado que admitimos con referencia fáctica, o sea no puramente formal, aunque tal referencia suele dejarse inespecificada. Este predicado, llamémosle *M*, se refiere a un substrato universal en el siguiente sentido. *M* es el único predicado tal que, para cualquier teoría *T* cuyo dominio de definición *D* contenga objetos reales, se cumple que *M* se aplica a todos los objetos de *D* –aunque el predicado *M* normalmente no aparece explícito en la formulación de *T*: por eso es *M* un predicado «recóndito».

Las dos dificultades básicas que veo en esta comprensión del concepto de materia son las siguientes:

1) Si no diferenciamos claramente los criterios de aplicación de *M* de los criterios de aplicación del predicado «es real», cosa que no se hace dentro de este esquema, entonces es difícil ver que los predicados «es material» y «es real» no son analíticamente sinónimos. Es decir, entonces es difícil ver que al afirmarse «todo lo real es material», como afirma el materialismo, se afirme algo más que «todo lo que hay es real», o sea «todo lo real es real», lo cual, evidentemente, no es afirmar gran cosa.

2) Concebir la materia simplemente como un substrato universal es una determinación ontológica pobre de contenido, puesto que no agota nuestras intuiciones presistemáticas de lo que significa ser materia. En efecto, podríamos idear doctrinas de substrato universal recóndito que interpretaran el predicado *M* de modo intuitivamente incompatible con la interpretación anterior. Por ejemplo, podríamos interpretar *M* como «es espiritual» (como hizo Leibniz), o bien como «es una sensación» (como hicieron Berkeley y los fenomenalistas), o bien como «es energía» (como hicieron Ostwald, Helm y otros adalides de la energética), o finalmente como «es un campo electromagnético» (como propugnaron Abraham y Lorentz). Intuitivamente quisiéramos distinguir la referencia de todos estos predicados de la de «es material». Pero el esquema que examinamos no nos proporciona la menor indicación en este sentido.

Segundo esquema: la materia como lo opuesto al espíritu

Este esquema funciona como una oposición entre dos formas ontológicas básicas, de las cuales una se suele describir como lo espiritual o lo mental y se caracteriza por no estar localizada espacio-temporalmente y por ser en algún sentido activa o dinámica. Generalmente se supone dentro de este esquema que lo mental es específico de los seres humanos, por oposición a otras formas de vida, y en muchos casos también de seres «superiores» no perceptibles,

llamados dioses, demonios, espíritus, etc. La materia se define entonces por oposición a lo espiritual o mental y en este enfoque suele caracterizarse como algo pasivo.

Una subforma menos espiritualista de este esquema, en vez de hacer uso del concepto de espíritu o mente, se limita a presuponer que entendemos las nociones de «Yo» y «voluntad», y define la materia como lo que se opone, o se resiste, a «mi voluntad».

Está claro que este esquema se adecúa a la doctrina ontológica conocida como *dualismo*, según la cual puede establecerse dentro de lo que hay una dicotomía tajante e irreductible entre una realidad material y una realidad espiritual o mental. Sin embargo, mediante una argumentación de lógica curiosa, a veces también se usa este esquema para apoyar una concepción monista, concretamente materialista. En efecto, primero se define la materia como lo que se opone al espíritu o a la voluntad del Yo y posteriormente se pasa a la negación de la existencia del espíritu o de un Yo independiente de la materia, con lo cual sólo quedaría la materia como realidad. Es decir, se intenta caracterizar lo real mediante la negación de lo que se supone que son las características de lo irreal.

Con esta última observación apuntamos ya a la dificultad central de este esquema, a saber: cómo caracterizar los predicados «mental» y «material» de modo que por un lado correspondan a nuestras intuiciones y por otro resulten efectivamente predicados mutuamente excluyentes. Si definimos lo material por oposición a lo espiritual, pero no damos criterios claros de aplicación del predicado «es espiritual», entonces no estamos definiendo ni distinguiendo nada. Definir lo material como «lo que se resiste a mi voluntad», aunque a primera vista parece más concreto, tampoco ayuda mucho, porque, aparte de que el significado del término «mi voluntad» no está nada claro, sea como sea que lo entendamos, parece intuitivamente obvio que hay cosas que no se resisten a mi voluntad, como ciertas partes de mi cuerpo, y que, sin embargo, consideramos materiales, mientras que puede haber cosas como dioses y demonios que, aun siendo consideradas espirituales, se resistan a mi voluntad.

Tercer esquema: la materia como «plenum»

En este esquema, que puede ser aplicado bien a una doctrina *monista* como la de la física estoica, bien a una *dualista* como la cartesiana, la materia se entiende como un continuo que llena el espacio, se la identifica con la extensión espacial o al menos se postula que en toda extensión espacial hay algo que siempre es lo mismo y que denominamos materia. Esto es la «res extensa» de Descartes. Puede considerarse que la teoría de la relatividad generalizada, al menos bajo cierta interpretación de esta teoría, también apoya una

identificación de la materia con el espacio. Las propiedades geométricas del espacio serían simplemente lo que llamamos propiedades de la materia.

Otra teoría física que toma la materia como un *plenum* continuo, aunque de una forma distinta a la relatividad general, es la dinámica de Boscović, en la que se admite un continuo de partículas materiales de diversas clases. Los puntos de la recta real se identifican aquí con las partículas materiales.

La nota común a estas concepciones continuistas de la materia es, pues, la siguiente: se admite que el predicado único «es material» conviene a cualquier punto espacial.

La dificultad básica en este esquema estriba en que no está claro cómo distinguimos el predicado «es material» del predicado «es espacial» o, dicho de otro modo, cómo distinguimos las propiedades puramente materiales de las puramente geométricas de las cosas. Si no disponemos de criterios que nos permitan distinguir una cosa de la otra, la identificación de la materia con un continuo espacial no nos dice nada, es una tautología. No creo que hasta la fecha se haya conseguido dar con tales criterios. Esta dificultad, naturalmente, va asociada a la dificultad en construir una ontología científicamente aceptable de lo que llamamos «espacio» o «espacio-tiempo», problema que sigue en pie, a pesar de los esfuerzos de los filósofos de la relatividad.

Cuarto esquema: la materia como lo indivisible

La idea de que la materia es una realidad numéricamente distinta y fácticamente indivisible se retrotrae a los atomistas griegos. Ha sido y es probablemente la concepción de materia más articulada y más influyente en el desarrollo de la filosofía y de la ciencia. Aún subyace a las intuiciones de la mayoría de los físicos actuales. La materia consistiría en un número muy grande, quizá infinito, de entidades invisibles por sus dimensiones ultramicroscópicas, pero que ocupan un cierto espacio; no son por tanto puntuales. La materia en su totalidad sería, pues, algo particulado, discreto.

El nombre de estas entidades últimas ha cambiado con la evolución de la ciencia. Primero fueron los átomos. A partir de fines del siglo pasado se vio que los supuestos átomos no eran tales átomos en el sentido genuino de la palabra, ya que estaban compuestos de otras entidades aún más minúsculas, a las que se llamó partículas elementales. Durante unos setenta años se consideró que no había nada más allá de las partículas elementales. Hace una década aproximadamente se planteó la posibilidad de que las partículas elementales a su vez estuvieran compuestas de los llamados «*quarks*». La teoría de los *quarks* ha ganado mucha popularidad entre los físicos

de altas energías en los últimos tiempos, si bien algunos físicos teóricos se sienten molestos por ese proceso de sucesivas atomizaciones que no parece tener fin, como un juego de cajas chinas. Sea como sea, para la discusión presente es relativamente irrelevante que tomemos como objeto de análisis conceptual los átomos clásicos, las partículas elementales o los *quarks* de moda. En el contexto presente podemos hablar indiferentemente de átomos o partículas.

Aunque esta concepción atomista de la materia se ha solido asociar a las formas más populares del materialismo, y por tanto a un monismo ontológico, lo cierto es que si la analizamos detenidamente nos lleva más bien a un *dualismo* por un lado y a un *pluralismo* por otro. El dualismo ontológico inherente a la concepción atomista de la materia proviene del hecho de que esos átomos deben estar espacialmente separados entre sí (de lo contrario obtendríamos el tercer esquema de la materia como continuo a lo Boscović). Debe haber algo real que separa los átomos. Esta segunda realidad, que es el sostén ontológico de los átomos, es el vacío. El vacío de la física moderna no es una mera no-entidad, como hubiera sido para los antiguos, sino algo real provisto de propiedades reales: fundamentalmente las propiedades de campo de interacciones. O sea que, según este esquema, hay dos clases de cosas radicalmente distintas: átomos y vacío (o campos). Es decir, tenemos un dualismo átomos-vacío o, si se prefiere, en lenguaje más moderno, partículas-campos. Por otro lado, para explicar gran parte de las propiedades macroscópicas de las cosas que observamos, especialmente las propiedades químicas, no sólo debemos admitir que los átomos son numéricamente distintos, sino que tienen propiedades distintas, por lo que el conjunto total de átomos se divide en varias clases de equivalencia: hay diversos tipos de átomos o partículas. Es decir, el predicado «es material» no se elucida completamente como «es partícula», sino más bien como «es una *k*-partícula», donde *k* es una variable para los distintos tipos de partículas. En este sentido el presente esquema es pluralista.

Vale la pena detenerse un poco en las características básicas de los átomos que están presupuestas en este esquema.

1) Los átomos están localizados en el espacio, con lo cual se presupone que tienen límites y formas espaciales bien determinables. Una expresión paradigmática de esta noción de materia discreta la dio ya Galileo: «Al concebir algo como material lo concibo limitado y teniendo una forma determinada; teniendo tamaño relativo respecto a otras cosas; en algún lugar en determinado momento; tocando o no otros cuerpos; siendo uno en número o varios» (cf. *Discoveries and Opinions of Galileo*, comp. y trad. por Stillman Drake, p. 274).

2) Los átomos son móviles. A ellos se aplica genuinamente el concepto de velocidad, y por tanto también los de aceleración y fuerza aplicada. La afirmación más decidida que conozco de que la movili-

dad espacial es una característica esencial de la materia se encuentra en los *Metaphysische Anfangsgründe* de Kant.

3) Los átomos se mueven *en* un medio, el vacío. Ahora bien, para la física actual este medio no es algo completamente inerte y amorfo, sino que posee importantes propiedades físicas: propiedades de campo de fuerzas o interacciones.

4) Otras dos características fundamentales que han jugado un gran papel en el desarrollo de la concepción atomista de la materia, pero que en la actualidad son cuestionadas, son las de eterna duración o indestructibilidad y masa de los átomos. El postulado de indestructibilidad, formulado ya explícitamente por Lucrecio, se vino abajo ante los experimentos de creación y aniquilación de partículas elementales. En cuanto a la masa, atomistas clásicos como Torricelli consideraban que el atributo de masa era necesariamente parte de la idea misma de átomo o partícula. Todavía a fines del siglo pasado, Charles de Freycinet afirmaba en una obra de filosofía de la ciencia bastante popular entonces: «Si yo tuviera que definir la materia diría: la materia es todo lo que tiene masa» (cf. *Essais sur la philosophie des sciences*, p. 168). Pero ya unos años antes algunos autores críticos, sobre todo Mach en su artículo «Über die Definition der Masse», habían cuestionado la supuesta necesidad del concepto de masa. Para Mach, «masa» es simplemente el nombre de ciertas relaciones cinemáticas de los objetos materiales[1]. Aunque en la actualidad este tipo de crítica de la noción de masa no es comúnmente aceptada ni por filósofos ni por científicos, la masa ha perdido su carácter fundamental para la materia por otras razones. En la física de partículas se suele hablar de partículas de masa cero, como el fotón y el neutrino, y aunque esta noción de partícula hubiera parecido un absurdo a los atomistas clásicos, ya nos hemos acostumbrado a ella. (Así es la historia.)

En esta situación las únicas propiedades fundamentales de las partículas con las que nos quedamos son su localización espacial y su movilidad en el vacío, sujeta a la acción de campos.

Veamos cuáles son las dificultades inherentes a este esquema. En primer lugar está el problema de su localización espacial. Sabemos por mecánica cuántica que los límites espaciales de las partículas en cualquier instante dado no están, por principio, bien definidos, sino que se trata de una distribución probabilística. La idea de la materia como algo precisamente delimitado en el espacio ha perdido su sentido. Esto, sin embargo, parece chocar con una idea intuitiva básica de que la materia es algo en principio localizable. Si abandonamos esta propiedad, abandonamos algo que parece esencial a la noción misma de materia. Por otra parte, al aceptar límites espaciales

[1] Cf. la discusión de Mach en el Cap. 3.5.

borrosos para las partículas se hace más difícil también distinguir entre cosas que son partículas y cosas que no son partículas, es decir, entre materia y no-materia.

Otro problema conceptual en la idea de partícula está relacionado con la teoría de la relatividad. En este contexto ha sido enfatizado por Mendel Sachs en *The Field Concept in Contemporary Science*. Si se incorpora el concepto de partícula a la teoría de la relatividad y además se supone que una partícula, por ser indivisible, es algo que por definición reacciona como una unidad a cualquier acción externa, entonces las partículas deberían ser entidades puntuales. En efecto, la velocidad de cualquier interacción, según la relatividad, es finita, y, por tanto, si la partícula tuviera una extensión finita, una porción de la misma reaccionaría antes que otra y la partícula en su totalidad ya no sería una unidad indivisible. Por razones emparentadas, Margenau, en su *Physics and Philosophy*, combatió la idea de que las partículas estén «llenas» de algo; para él son sólo singularidades del espacio y por eso mismo, según él, tampoco se les puede aplicar propiamente la noción de movimiento, en contra de la idea básica kantiana. Ahora bien, por otro lado, la idea de las partículas elementales como puntos en el espacio vacío choca no sólo con nuestras intuiciones corrientes acerca de lo que debe ser la materia, sino con la adscripción de dimensiones finitas para la mayoría de partículas que se estudian en la física de altas energías, adscripción que parece estar bien fundamentada.

Una tercera dificultad está conectada con la pluralidad ontológica de las partículas en la física actual. En la actualidad se conocen más de 200 clases distintas de partículas elementales, algunas de ellas, como las famosas «resonancias», tan alejadas de cualquier intuición razonable de lo que puede ser una partícula material que se hace difícil catalogarlas como tales. Si tomáramos los *quarks* como los verdaderos átomos del Universo, esta pluralidad ontológica sólo disminuiría en número, pero no en cualidad, puesto que, a parte de que los *quarks* sólo sirven para explicar la constitución de los hadrones y no de todas las partículas, también hay que distinguir por lo menos seis tipos distintos de *quarks*. El problema que se plantea ante tal pluralidad es simplemente si podemos seguir hablando de *la materia*, como si se tratara de una realidad única. ¿No habremos de aceptar más bien que, en definitiva, hay distintos tipos de constituyentes últimos de la realidad, mutuamente irreducibles, y que seguir hablando de *la* materia es una piadosa reliquia del pasado? (Recuérdese que en los varios milenios de uso que lleva el concepto de materia ésta ha solido presentarse como una realidad única y homogénea que pretende explicar unitariamente las diferencias y variaciones en el mundo percibible. Si al concepto de materia se le despoja de su carácter unitario, pierde mucho de su atractivo, al menos para el filósofo.)

He expuesto hasta aquí diversas formas posibles e históricamente dadas de entender el concepto de materia, así como las dificultades que plantean. El hecho de que la noción de materia sea multívoca y problemática no significa, empero, que haya que considerarla inútil o carente de sentido. Lo único que significa es que, antes de pronunciarse en forma decidida por alguna ontología universal en la que el concepto de materia intervenga esencialmente, debemos proseguir nuestros esfuerzos por aclarar el concepto y ver qué da de sí. En esta tarea es obvio que deben colaborar estrechamente científicos y filósofos.

4.3. LOS PROBLEMAS DEL MATERIALISMO

El materialismo es una doctrina confusa. Si se cree que el materialismo es una doctrina clara es porque afirma que sólo existe la materia y porque se supone que todo el mundo sabe lo que es la materia. Pero este supuesto es falso. Nadie *sabe* hoy día a ciencia cierta lo que es la materia. (Otra cuestión es la de que muchos *crean* saberlo.) Tampoco lo saben los físicos de partículas, los especialistas a quienes el resto de los mortales deberíamos preguntar qué es la materia. Ciertamente, si se les hacen preguntas ontológicas de este tipo, algunos físicos (los más osados) darán ciertas respuestas esotéricas acerca de «ondas de probabilidad» o de «puntos de singularidad del espacio-tiempo» o algo por el estilo. Se trata de respuestas que la mayoría de las personas que se autotitulan materialistas no entienden; por lo demás, tales respuestas cambian de sentido cada cinco o diez años, y en ellas ni siquiera los propios especialistas están de acuerdo.

Sin embargo, el materialista lego, aunque no entiende las respuestas de los físicos, suele quedar muy satisfecho con ellas: piensa que hay al menos algunas personas en el mundo que saben lo que es la materia; y ello le basta para sostener la afirmación rotunda: «Sólo existe la materia», la cual equivale en estas circunstancias a la afirmación (ya no tan rotunda): «Sólo existe algo que una pequeña minoría de especialistas sabe lo que es.»

El caso, no obstante, es que si el materialista lego quiere indagar más y le sigue preguntando al físico osado en qué consiste exactamente todo eso de las ondas de probabilidad, puntos de singularidad, etc., lo llevará pronto a una situación embarazosa, en la que el físico, si es honesto, deberá reconocer que las cosas están todavía muy embrolla-

das, que aún está por llegar «la» teoría plenamente satisfactoria, etc.; y al final seguramente le recomendará al materialista lego que se espere todavía unos años para saber qué es la materia.

Creo que ésta es realmente la situación en la física de fundamentos hoy día. Y también creo que, dada esta situación, es por lo menos poco cauteloso ser materialista. En efecto, si uno afirma «todo es materia», pero no tiene una idea razonablemente clara de qué es la materia, entonces se halla en una posición tan incómoda como la de alguien que afirmara «todos los que viven enfrente son ugrofineses» y no supiera qué son los ugrofineses. Una afirmación semejante en tales circunstancias sólo puede justificarse si se apela a una autoridad indiscutible («Fulano, que puede reconocer bien a los ugrofineses, me ha dicho que enfrente sólo viven ugrofineses»). Pero si, como en el caso de la materia, no hay tal autoridad indiscutible, o si la supuesta autoridad (los físicos) anda más bien desorientada y no se decide a pronunciarse, entonces realmente lo mejor es «callarse sobre lo que no se puede hablar».

Claro que muchos materialistas, sobre todo si son filósofos tradicionales o activistas políticos, no se preocupan en absoluto por lo que digan o dejen de decir los físicos. Para ellos está bien claro qué es la materia: las cosas que se ven y, sobre todo, se *tocan* (y mejor si uno se da un golpe en la cabeza, como quería Lenin: entonces resulta dolorosamente claro qué es la materia)[2]. Lo más cortés que se puede decir de esta concepción es que es sumamente ingenua: apela a las formas más burdas de los prejuicios del llamado «sentido común». Y lo más interesante del materialismo puramente ingenuo es que, en último análisis, se acerca peligrosamente a su enemigo acérrimo, el idealismo subjetivo. Pues si sólo hay materia, y la materia es simplemente lo que se ve y se toca, entonces sólo hay lo que se ve y se toca –y nos pasamos al bando de Berkeley.

Esta deplorable defección parece evitable si se caracteriza la materia no sólo como «lo que se ve y se toca», sino además como «lo que está ahí afuera», o «lo que es exterior a mí», o «lo que es independiente de *mi* ver y tocar». El problema con este tipo de añadidos es darles un sentido preciso que satisfaga las intuiciones del materialista. Veo por lo menos dos callejones sin salida a estas caracterizaciones adicionales.

En primer lugar, suponen una distinción entre el «fuera» y el «dentro», entre «lo externo» y «lo interno», que, en último término, nos lleva a alguna forma de dualismo materia-espíritu. Si la materia es «lo externo», «lo independiente de *mi* voluntad», entonces, por pura lógica, es que *hay* algo «interno», hay «*mi* voluntad» (o «la concien-

[2] Es interesante que muchas doctrinas epistemológicas y ontológicas vayan ligadas esencialmente a algún sentido particular: el materialismo está vinculado al sentido táctil, como el empirismo lo está al sentido visual.

cia», o «el sujeto», o como se le quiera llamar) que no es materia. Y entonces, en vez de caer en los brazos de Berkeley, el materialista cae en los de Descartes.

En segundo lugar, aun cuando se le pudiera dar un sentido no-dualista a la caracterización de la materia como «lo exterior» o «lo que se resiste a mi voluntad», tendríamos el problema de que esa «exterioridad independiente» no tiene por qué ser «material» en ningún sentido. Mis mesas, paredes y martillos, a pesar de que puedan ocasionarme algún dolor de cabeza (o justamente por eso), podrían muy bien ser espíritus malévolos y jocosos, que existieran externa e independientemente de mí mismo. Afirmar «hay cosas externas a mí e independientes de mí» no basta para ser materialista. Podemos sustituir a Berkeley y Descartes por Leibniz.

Hay, naturalmente, un tipo de materialista más refinado que trata de evadir los problemas del materialista ingenuo dejando de lado las cosas que se ven y se tocan y elaborando una concepción más teórica del universo. Ya no se postula que la materia es lo que se ve y se toca, sino que es lo que está «detrás» (como constructo teórico) de lo que se ve y se toca. Comúnmente, este materialismo refinado acude a una concepción atomista de la materia. La materia consiste, así, en agregados de átomos, y la tesis materialista refinada sostiene entonces que lo único que existe en el universo son átomos. Todos los objetos reales son conjuntos de átomos y las relaciones entre objetos son, en último análisis, relaciones entre átomos. No hay entonces diferencia ontológica entre lo externo y lo interno. El llamado «sujeto percipiente» es un conjunto de átomos que se halla en ciertas relaciones atómicas con el «objeto percibido», el cual también es un conjunto de átomos. Y los llamados «fenómenos de la conciencia» son también, en definitiva, fenómenos atómicos. Así se sale de los callejones sin salida en que se había metido el materialista ingenuo, y además se llega a una «visión científica» del universo.

Desgraciadamente, también el materialismo atomista tiene sus problemas. Para empezar, podemos volver a hacer la pregunta insidiosa de siempre: «La materia son los átomos; muy bien, ¿y qué son los átomos?» El *filósofo* atomista no suele tener una respuesta clara a esta pregunta; en general, él mismo ni siquiera se la plantea. Y si uno se la plantea a él, dirá (si es honesto) que él no lo sabe muy bien, pero que sí lo saben bien los físicos. Que es como decir que uno no sabe qué son los ugrofineses, pero que hay alguien que sí lo sabe. Volvemos al problema del principio: hay que acudir a la autoridad de los especialistas. Los especialistas nos dirán que los átomos son agregados de partículas elementales; y si les preguntamos qué son las partículas elementales, nos responderán, como ya habíamos observado al principio, con explicaciones esotéricas que casi nadie (incluyendo a la inmensa mayoría de materialistas) entiende, o bien, si perseveramos, nos dirán que tampoco ellos lo saben bien. De lo cual

resulta que los materialistas atomistas están afirmando una tesis que no entienden bien –porque dada la situación actual nadie *puede* entenderla bien.

Pero no era ésta la dificultad básica del materialismo atomista que quería subrayar ahora. El punto que ahora interesa (y que parece ser independiente de los resultados a que pueda llegar la física de partículas en el futuro) es que si se dice que la materia consiste en agregados de átomos y no se quiere admitir la visión absurda de un mundo infinitamente denso e inmóvil, entonces hay que admitir la *existencia*, además de los átomos, de otra cosa: el vacío. El vacío es la entidad por la que circulan los átomos (o las partículas elementales, los «verdaderos» átomos de la física actual) y que permite explicar cómo los átomos hacen todas las cosas que se supone que hacen. Pero entonces a la tesis de que todo lo que *existe* es materia se le puede plantear en seguida la pregunta: ¿Es materia el vacío? Caben dos posibles respuestas a esta pregunta:

(1) Afirmar que el vacío también es materia *por definición*, lo cual, además de chocar con nuestra intuición ordinaria de lo que hay que entender por «objeto material», hace al concepto de materia sumamente heterogéneo y, por ende, *ad hoc*.

(2) Negar que el vacío sea materia, pero propugnar que un materialismo «bien entendido» es lo suficientemente liberal como para admitir la existencia de entes no «estrictamente» materiales, aunque «íntimamente relacionados» con los objetos materiales. O, dicho de otro modo, el materialismo liberal puede admitir la existencia de algunas entidades que, sin ser propiamente materia, «ayudan» a la materia a existir y a tener las propiedades que tiene.

Pero entonces, si admitimos esta última respuesta, debemos estar dispuestos en principio a admitir también cualquier tipo de entidades no-materiales que nos parezcan convenientes en un momento dado para «ayudar» a la materia (los átomos) a existir: desde dioses infinitos hasta mónadas leibnizianas, pasando por formas platónicas y almas cristianas. Un materialismo, en verdad, muy aguado.

Se puede tratar de trivializar el problema del vacío adoptando una postura «lingüística» como tercera respuesta posible. Puede decirse que, aunque los átomos circulen por el vacío, el vacío «propiamente hablando» no existe, precisamente por ser vacío. Pero esta salida, aparte de ser un mero juego de palabras, tendría la consecuencia inaceptable de calificar de absurdos todos los estudios empíricos que se han hecho sobre el vacío desde la época de Torricelli y Pascal: serían estudios *sobre nada*. El hecho es que el problema del vacío es un problema real para el materialismo atomista, y ya los fundadores de esta doctrina en la Antigüedad se preocuparon mucho por él sin llegar a una solución satisfactoria.

Por lo demás, el problema se plantea en términos similares para otras entidades de cuya existencia están convencidos los físicos modernos y de las cuales no parece tener mucho sentido predicar la materialidad: campos de fuerza, ondas electromagnéticas, acciones a distancia (al menos para la gravitación) son cosas que existen, pero que no son partículas materiales ni se reducen a agregados o movimientos de partículas materiales. Para no hablar de las oscuras entidades que pululan en la mecánica cuántica y en la física de partículas (fotones, ondas de probabilidad, dipolos de vacío, agujeros negros, etc.), y que incluso han llevado a algunos físicos (Heisenberg, por ejemplo, poco antes de su muerte) a adelantar la posibilidad de que los constituyentes últimos del universo no sean las llamadas «partículas elementales», sino *formas* matemáticas puras, algo así como las formas ideales de Platón.

Pero podemos olvidar de momento estas profundidades abismales de los últimos gritos (¿agónicos, quizá?) de la física de partículas y quedarnos con la serie de entidades de la mecánica y la electrodinámica clásicas que no son partículas materiales, ni agregados de tales, y cuya existencia es admitida sin discusión. También podríamos añadir ciertas propiedades de los sistemas termodinámicos, como la energía interna y la entropía, que no parecen definibles o explicables en términos de partículas en movimiento solamente, y cuya «realidad física» parece indubitable. Algunos sostienen que se ha demostrado la materialidad de las propiedades termodinámicas al haber «reducido» la termodinámica a la mecánica estadística. Pero aún hoy día, a cien años de la supuesta reducción, es muy dudoso que haya tenido lugar tal reducción con el rigor necesario y para todas las aplicaciones de la termodinámica. La cuestión sigue abierta.

Enfrentado a tales entidades, el materialista suele replicar en tono condescendiente que también caben dentro de la concepción materialista, puesto que pertenecen a la «visión científica» del mundo y su existencia ha sido suficientemente confirmada por los experimentos de los científicos. Ahora bien, esto, en verdad, no es argumento alguno para predicarles la materialidad, a menos que se defina la materia como aquello en cuya existencia cree una comunidad científica en una etapa histórica dada. Esta definición haría del materialismo una doctrina vacía de contenido –servilmente subordinada a los caprichos de la moda científica del momento. Con esta actitud podríamos tomar cualquier cosa dentro del materialismo. En efecto, la estructura conceptual de nociones tales como las de *transmisión telepática* o *influencia astral* no es tan distinta de la de nociones como *campo electromagnético* o *acción a distancia*, respectivamente. Veo sólo dos diferencias entre ambos grupos de ejemplos. Primera, que los segundos conceptos han sido matematizados y los primeros no; segunda, que los segundos conceptos pertenecen al bagaje conceptual del hombre culto moderno y los primeros (¿todavía?) no. Pero ambas

diferencias me parecen irrelevantes para la discusión presente. No siento ninguna simpatía especial hacia la parapsicología ni hacia la astrología; no obstante, puedo *imaginar* que parapsicólogos y astrólogos matematicen sus conceptos, y que éstos empiecen a aparecer en los libros de texto del bachillerato. (Todo es posible en la historia.) Si tal cosa ocurriese, me parecería un claro *non sequitur*, y además un *tour de force* vergonzante, concluir que transmisiones telepáticas e influencias astrales concuerdan perfectamente con la concepción materialista.

Naturalmente, dado el estado actual de la ciencia, un materialista consecuente rechazará tales cosas como «oscurantistas». Pero entonces, ¿por qué admite tranquilamente la existencia de entidades tales como campos electromagnéticos y acciones a distancia, que conceptualmente se parecen más a las entidades telepáticas y astrológicas mencionadas que a los átomos? Decir que están «mejor confirmadas» no es, obviamente, una respuesta pertinente, pues no es de confirmación empírica de lo que aquí se trata, sino de materialidad o no-materialidad de principio.

La respuesta que suele intentar el materialista en este punto es que las entidades de la física, a diferencia de las entidades parapsicológicas y astrológicas, están «íntimamente vinculadas» a procesos materiales: aunque ellas mismas no sean partículas ni agregados de tales, son sus «efectos», o «propiedades emergentes», o «propiedades concomitantes», o algo por el estilo. Con lo cual, aparte de introducir una muy dudosa ontología de causas y efectos, de propiedades emergentes, etc., no se resuelve el problema en absoluto. Puesto que el hecho de que *A* sea el efecto o la causa o la propiedad emergente de *B*, y *B* sea una entidad material, no implica la materialidad de *A*. Un espiritualista puede sostener consistentemente que el alma es un efecto o una propiedad emergente del cuerpo y, sin embargo, quedarse con la tesis de que el alma es una entidad no-material, espiritual. Y también puede admitirse sin contradicción que «el espíritu es el término final de la evolución de la materia»; estas cosas y otras aún más extrañas se han dicho. El punto no es que parezcan extrañas a una mente esclarecida, sino que no son contradictorias y que el materialista no puede mostrar su falsedad de principio. Newton, por cierto, estaba plenamente convencido de que la gravitación y las fuerzas electromagnéticas eran efectos espirituales de una realidad espiritual que actuaba sobre la materia, y también creía que esto era perfectamente compatible con su concepción atomista de la materia. Y sea como sea que evaluemos actualmente su convicción, lo que no podemos decir es que fuera incoherente.

Finalmente, seamos benevolentes con el materialista y admitamos la posibilidad de que puedan reducirse a materia todas las entidades y propiedades materialísticamente molestas de la física clásica. Admitamos, por ejemplo, que se pueda sustituir las acciones a distancia por

las interacciones de supuestas partículas elementales (los «gravitones»), que se pueda volver a una especie de teoría del éter para el campo electromagnético, y que realmente se haya mostrado sin lugar a dudas que todas las propiedades termodinámicas son reducibles estrictamente a propiedades de partículas en movimiento. Es más, olvidemos al enojoso vacío, a la inquietante mecánica cuántica y a la pavorosa física de altas energías. Y, para definitiva tranquilidad del materialista, supongamos fuera de toda duda que la parapsicología, la astrología y otras disciplinas semejantes quedarán eternamente en el estadio de basura intelectual (a menos que se decidan a introducir átomos en sus teorías).

Con todas estas generosas concesiones, sigue habiendo dos entidades, de cuya realidad están persuadidos todos los científicos y la mayoría de los filósofos, y que no parece de ninguna manera plausible suponer que sean reducibles a la materia: espacio y tiempo o, si se prefiere, la entidad compleja espacio-tiempo. Puntos espaciales e instantes temporales (o su contrapartida minkowskiana, los puntos-universo) son entidades cuya existencia conviene admitir para una visión adecuada del mundo. Por otro lado, me parece absurdo decir que los puntos y los instantes son objetos materiales o agregados de tales. Esto sólo podría decirse estirando tanto el concepto de materia que en él cupiera cualquier cosa imaginable. La única salida que se vislumbra en este punto para el materialista es una posible interpretación de la teoría de la relatividad generalizada, según la cual el espacio-tiempo y la materia se confunden en un todo indistinguible cuasi-parmenídeo. Pero en esa fusión es más bien la materia la que sale perdiendo y el espacio-tiempo (la geometría) quien sale ganando. Y por otro lado tal interpretión de la relatividad generalizada, además de chocar con el corpus generalmente reconocido de la física actual, está lejos de haber sido confirmada cabalmente, y más lejos aún está la comprensión plena de sus consecuencias ontológicas y epistemológicas. Admito que este materialismo geométrico es el más coherente que puedo imaginar en este momento (más coherente, por supuesto, que el materialismo atomista); lo que creo es que nadie o casi nadie tiene actualmente una idea clara acerca de sus consecuencias para nuestra concepción del mundo. Por lo menos yo no la tengo, y tampoco la tienen mis amigos materialistas. Y, en realidad, la materia de esta última forma de materialismo sería una entidad tan sumamente abstracta que difícilmente podría provocar el entusiasmo de nuestros materialistas cotidianos. Por supuesto, nadie piensa en esta materia puramente geométrica cuando polemiza con los anti-materialistas.

Las dificultades del materialismo expuestas hasta aquí me parecen muy graves para alquien que quiera sostener el materialismo con honestidad intelectual. Ello no implica que las considere por principio insalvables. El caso es simplemente que la doctrina me parece tan

problemática que considero filosóficamente prudente abstenerse de comulgar con ella, al menos en la situación actual.

Creo que la dificultad radical del materialismo es, en último término, la misma que la de cualquier concepción monolítica del universo que no quiera ser vacuamente verdadera. Si pretendemos que todas las cosas que existen pertenecen a una misma categoría determinada, o sea si pretendemos que todo *x* es *P*, entonces debemos dar cierta caracterización no-vacua del predicado *P*, a base de condiciones empíricas que sean al menos un poco restrictivas (que no las pueda cumplir cualquier cosa imaginable). Si no se cumple este requisito, afirmar «todo *x* es *P*» es no afirmar nada, ya que *P* puede ser entendido a gusto de cada quien. Ahora bien, es difícil imaginar que pueda determinarse satisfactoriamente un predicado tal que convenga por igual a todas las cosas que existen. No hay por qué suponer que la realidad satisface nuestros deseos de supersimplificación conceptual. Tal como están las cosas, me parece *a priori* más sensato suponer lo contrario: que la realidad es heterogénea en vez de homogénea, y que hay de todo en la viña del Señor. No hay por qué suponer que podamos meter todas las cosas en el mismo saco, a menos que se trate de un saco sin fondo, es decir, un pseudo-saco.

Permítaseme terminar con una confesión personal. Las líneas anteriores podrían hacer creer que soy un adversario enconado del materialismo filosófico. Esta creencia podría dar lugar a dos reacciones posibles (que pueden darse juntas). Si uno se adscribe al dogma filosófico según el cual si alguien no es materialista, entonces debe ser forzosamente idealista, concluirá de ahí que soy un idealista recalcitrante. No creo que valga la pena detenerse a discutir esta burda dicotomía a lo blango-o-negro, que manda al infierno idealista a todos los que no quieren (o no pueden) entrar al cielo materialista. Me limito a observar que encontrarle graves dificultades al materialismo no implica que para mí el idealismo sea la gran cosa. (El idealismo es, a mi entender –¿hace falta decirlo?–, una doctrina aún más confusa e inverosímil que el materialismo.)

Otra reacción posible, y mucho más respetable, es la que proviene de cierta actitud vital frente al mundo. Muchas personas que se autoconsideran materialistas asumen esta concepción más como una actitud vital que como una doctrina filosófica técnica. Estas personas podrían pensar que, puesto que no me adscribo al materialismo, mi actitud vital es exactamente la opuesta. Nada de eso. La actitud vital del materialista (al menos la del materialista de mente abierta) me parece más valiosa humanamente que la de la mayoría de sus contrincantes habituales. El materialista suele ser antioscurantista, anti-místico, anti-religioso, anti-romántico y, en general, anti-cuentos-de-hadas. Todos éstos son «antis» con los que mi temperamento filosófico simpatiza plenamente. Sin embargo, una cosa es el temperamento y las simpatías personales y otra muy distinta las

conclusiones que hay que sacar de un análisis conceptual correcto. Este debe ser despiadado, ante todo, con la idiosincrasia de uno mismo.

BIBLIOGRAFIA

Adams, E. W.: *Axiomatic Foundations of Rigid Body Mechanics*. Stanford, 1955.

——: «The Foundations of Rigid Body Mechanics and the Derivation of Its Laws from Those of Particle Mechanics». En *The Axiomatic Method* (comp. por L. Henkin, P. Suppes, A. Tarski). Amsterdam, 1959.

Avenarius, R.: *Philosophie als Denken der Welt gemäss dem Prinzip des kleinsten Kraftmasses*. Leipzig, 1876.

Balzer, W., y J. D. Sneed: «Generalized Net Structures of Empirical Theories». *Studia Logica*, 36 (1977) y 37 (1978).

Bar-Hillel, Y.: «Neorrealism vs. Neopositivism: A Neo-Pseudo Issue». En *Aspects of Language* (comp. por Y. Bar-Hillel). Amsterdam-Jerusalén, 1970.

Bochner, S.: *The Role of Mathematics in the Rise of Science*. Princeton, 1966.

Bourbaki, N.: *Topologie générale*. París, 1951.

Braithwaite, R. B.: *La explicación científica* (trad. del inglés de V. Sánchez de Zavala). Madrid, 1965.

Buchdahl, H. W.: *The Concepts of Classical Thermodynamics*. Cambridge, Gran Bretaña, 1966.

Cabanis, P. J. G.: *Oeuvres philosophiques*. París, 1.ª ed., 1823-1825; 2.ª ed., 1956.

Cajori, F.: «Appendix to Sir Isaac Newton's *Mathematical Principles of Natural Philosophy*». Berkeley, 1947.

Callen, H. B.: *Thermodynamics*. Nueva York, 1960.

Carathéodory, C.: «Untersuchungen über die Grundlagen der Thermodynamik». *Mathematische Annalen*, 67, 1909.

Carnap, R.: *Der Logische Aufbau der Welt*. 1.ª ed., 1928; 2.ª ed., Hamburgo, 1966.

Cohen, I. B.: *Introduction to Isaac Newton's Principia*. Harvard, 1971.

Comte, A.: *Discours sur l'esprit positif*. 1.ª ed., París, 1844; edición franco-alemana, traducida e introducida por I. Fetscher, Hamburgo, 1956.

——: *Cours de philosophie positive*. 1.ª ed., París, 1830-42; edición resumida por el Dr. Robinet, París, 1881.
D'Alembert, J. R.: *Traité de dynamique*. París, 1921 (1.ª ed., París, 1743).
Davis, M.: *Computability and Unsolvability*. Nueva York, 1958.
Duistermaat, J. J.: «Energy and Entropy as Real Morphisms for Addition and Order». *Synthese*, 18, 1968.
Dummett, M.: *Frege - Philosophy of Language*. Londres, 1973.
Einstein, A.: «Autobiography». En *Albert Einstein - Philosopher-Scientist* (comp. por A. Schilpp). La Salle, Illinois, 1949.
Falk, G.: «On the Third Law of Thermodynamics». *The Physical Review*, 115, 1959.
Falk, G., y H. Jung: «Axiomatik der Thermodynamik». En *Handbuch der Physik*, t. III/2 (comp. por S. Flügge), 1959.
Ferrater Mora, J.: *El sentido de la muerte*. Buenos Aires, 1947.
——: *El ser y el sentido*. Madrid, 1967.
Feyerabend, P.: «Explanation, Reduction and Empiricism». En *Minnesota Studies in the Philosophy of Science*, t. III (comp. por H. Feigl y G Maxwell). Minneapolis, 1962.
——: «Problems of Empiricism». En *Beyond the Edge of Certainty* (comp. por R. G. Colodny). Englewood Cliffs, 1965.
Frege, G.: *Estudios sobre semántica* (trad. del alemán de C. U. Moulines). Barcelona, 1971.
Freycinet, Ch.: *Essais sur la philosophie des sciences*. París, 1896.
Galilei, G.: *Discoveries and Opinions of Galileo* (comp. y trad. de S. Drake). Nueva York, 1957.
García Aguilar, C. L.: «Reseña de J. C. C. McKinsey, A. C. Sugar, P. Suppes, "Axiomatic Foundations of Classical Particle Mechanics"». *Crítica*, 28, 1978.
Gerhard, C.: *Der Briefwechsel von Leibniz mit Mathematikern*. Berlín, 1899.
Gibbs, W.: *Collected Works*, t. I, *Thermodynamics*. Yale, 1948.
Giles, R.: *Mathematical Foundations of Thermodynamics*. Nueva York, 1964.
Goldstein, H.: *Mecánica clásica* (trad. del inglés de C. Enríquez de Salamanca). Madrid, 1966.
González, J. L.: *Una reconstrucción conjuntista de la semántica de Frege*. México, 1979.
Hall, A. R., y M. B. Hall (comp. y coment.): *Unpublished Scientific Papers of Isaac Newton*. Cambridge, Gran Bretaña, 1962.
Helmholtz, H. V.: *Philosophische Vorträge und Aufsätze* (comp. de H. Hörz y S. Wollgast). Berlín, 1971.
——: *Vorträge und Reden*. Braunschweig, 1903.
Hermann, A., y W. Kaiser: «Der Positivismus in der Physik des 18. und 19. Jahrhunderts». *Rete*, I/2, 1972.
Hertz, H.: *Die Prinzipien der Mechanik*. Leipzig, 1894.
Hornix, W. J.: «An axiomatization of classical phenomenological thermodynamics». En *A Critical Review of Thermodynamics* (comp. por E. B. Stuart, B. Gal-Or, A. J. Brainard). Baltimore, 1970.
Hörz, H., y S. Wollgast: *Einleitung zu Hermann von Helmholtz' Vorträge und Aufsätze*. Berlín, 1971.
Jané, I.: «Observaciones sobre el concepto de aproximación empírica». *Crítica*, 35. México, 1980.
Kahl, R.: *Introduction to Selected Writings of Hermann von Helmholtz*.

Kant, I.: *Metaphysische Anfangsgründe der Naturwissenschaft*. Darmstadt, 1966.

Krajewski, W.: *The Correspondence Principle and the Growth of Knowledge*. Dordrecht, 1978.

Kirchhoff, G.: *Vorlesungen über Mechanik*. Leipzig, 1876.

Kolakowski, L.: *Die Philosophie des Positivismus* (trad. del polaco de P. Lachmann). Munich, 1971.

Kuhn, T. S.: *The Structure of Scientific Revolutions*. Chicago, 1.ª ed., 1962; 2.ª ed., 1970.

——: «Postscript 1969». En *The Structure...*, *op. cit.*, 2.ª ed.

——: «A Letter to W. Stegmüller Concerning *Theorie und Erfahrung*». Manuscrito no publicado, 1975.

——: «A Formalism for Scientific Change». Ponencia presentada ante el 5.º *Congreso Internacional de Lógica, Metodología y Filosofía de la Ciencia*, London, Ontario, 1975. Una versión revisada y ampliada de este trabajo se publicó como «Theory-Change as Structure-Change: Comments on the Sneed Formalism». *Erkenntnis*, 10/2, 1976.

Lagrange, J. L. de: *Mécanique analytique*. París, 1788.

Lakatos, I.: «La falsación y la metodología de los programas de investigación científica». En *La crítica y el desarrollo del conocimiento* (trad. del inglés de F. Hernán). Barcelona, 1975.

Lange, F. A.: *Historia del materialismo* (trad. del alemán de V. Colorado). Madrid, 1903.

Ludwig, G.: *Deutung des Begriffs 'physikalische Theorie' und axiomatische Grundlegung der Hilbertraumstruktur der Quantenmechanik durch Hauptsätze des Messens*. Berlín-Heidelberg, 1970.

——: *Einführung in die Grundlagen der Physik*, t. I. Düsseldorf. 1974.

——: *Die Grundstrukturen einer physikalischen Theorie*. Berlín-Heidelberg, 1978.

Mach, E.: «Über die Definition der Masse». En *Carls Repertorium der Experimentalphysik*, t. 4. Munich, 1868.

——: *Die Mechanik in ihrer Entwicklung historisch-kritisch dargestellt*. 1.ª ed., Praga, 1883; 9.ª ed., Leipzig, 1933.

——: *Análisis de las sensaciones* (trad. del alemán de E. Ovejero). Madrid, 1925.

Margenau, H.: *Physics and Philosophy*. Dordrecht, 1978.

Mayr, D.: «Investigations of the Concept of Reduction». *Erkenntnis*, 10/3, 1976.

McKinsey, J. J. C., y A. C. Sugar y P. Suppes: «Axiomatic Foundations of Classical Particle Mechanics». *Journal of Rational Mechanics and Analysis*, t. 2, núm. 2, 1953. (Trad. castellana de A. G. de la Sierra, Universidad Michoacana, Morelia, 1978.)

Merz, J. Th.: *A History of European Thought in the Nineteenth Century*. Dover, 1965.

Mittelstaedt, P.: *Klassische Mechanik*. Mannheim, 1970.

Mosterín, J.: «Sobre el concepto de modelo». *Teorema*, VIII/2. Valencia, 1978.

Moulines, C. U.: *La estructura del mundo sensible. Sistemas fenomenalistas*. Barcelona, 1973.

——: «A Logical Reconstruction of Simple Equilibrium Thermodynamics», *Erkenntnis*, 9/1, 1975.

——: «Approximate Application of Empirical Theories: A General Explication». *Erkenntnis*, 10/2, 1976.
——: «Reply to John D. North's 'On Making History'», en *Probabilistic Thinking, Thermodynamics, and the Interaction of the History and Philosophy of Science* (comp. por J. Hintikka, D. Gruender y E. Agazzi). Dordrecht, 1981.
—— y J. D. Sneed: «Suppes' Philosophy of Physics». En *Patrick Suppes* (comp. por R. J. Bogdan). Dordrecht, 1979.
Nagel, E.: *La estructura de la ciencia* (trad. del inglés de N. Míguez). Buenos Aires, 1968.
Newton, I.: *Philosophiae Naturalis Principia Mathematica*. Ed. latina de A. Koyré e I. B. Cohen, Cambridge, Massachusetts, 1972; versión inglesa de F. Cajori, con el título *Newton's Principia-Motte's Translation Revised*. Berkeley, 1947.
——: *Opticks*. En *Opera quae extant omnia*, t. IV, 1.ª ed., Londres, 1779-1785; 2.ª ed., Stuttgart, 1964.
Planck, M.: *Vorlesungen über Thermodynamik*. Berlín, 1964.
Prokhovnik, S. J.: *The Logic of Special Relativity*. Cambridge, Gran Bretaña, 1967.
Putnam, H.: «What Theories Are Not». En *Logic, Methodology and Philosophy of Science* (comp. por E. Nagel, P. Suppes, A. Tarski). Stanford, 1962.
Rattansi, P. M.: «Some Evaluations of Reason in Sixteenth- and Seventeenth-Century Natural Philosophy». En *Changing Perspectives in the History of Science* (comp. de M. Teich y R. Young). Londres, 1973.
Roberts, F. S., y R. D. Luce: «Axiomatic Thermodynamics and Extensive Measurement». *Synthese*, 18, 1968.
Sachs, M.: *The Field Concept in Contemporary Science*. Springfield, 1973.
Scheibe, E.: «Die Erklärung der Keplerschen Gesetze durch Newtons Gravitationsgesetz». En *Einheit und Vielheit. Festschrift für Carl Friedrich von Weizsäcker* (comp. por E. Scheibe y G. Süssmann). Gotinga, 1973.
——: «Vergleichbarkeit, Widerspruch und Erklärung». En *Philosophie und Physik* (comp. por R. Haller). Braunschweig, 1975.
——: «Eine Fallstudie zur Grenzfallbeziehung in der Quantenmechanik». Manuscrito, 1979.
Shwayder, D. S.: «Uses of Language and Uses of Words». En *The Theory of Meaning* (comp. por G. H. R. Parkinson). Oxford, 1968.
Sneed, J. D.: *The Logical Structure of Mathematical Physics*. Dordrecht, 1971.
——: «Philosophical Problems in the Empirical Science of Science: A Formal Approach». *Erkenntnis*, 10/2, 1976.
Stegmüller, W.: *Theorie und Erfahrung*. Berlín-Heidelberg, 1970.
——: *Theorienstrukturen und Theoriendynamik*. Berlín-Heidelberg, 1973.
——: «Estructura y dinámica de las teorías. Algunas reflexiones sobre J. D. Sneed y T. S. Kuhn» (trad. de A. Morones). *Dianoia*, México, 1975.
——: «Accidental ('Non-Substantial') Theory-Change and Theory-Dislodgement». *Erkenntnis*, 10/2, 1976.
——: «A Combined Approach to the Dynamics of Theories». *Theory and Decision*, 9, 1978.
——: *The Structuralist View of Theories*. Nueva York, 1979.
Suppes, P.: *Introduction to Logic*. Nueva York, 1957.

——: «Models of Data». En *Logic, Methodology and Philosophy of Science* (comp. por E. Nagel, P. Suppes, A. Tarski). Stanford, 1962.
——: *Set-theoretic Structures in Science*. Stanford, 1970.
Symon, K. R.: *Mechanics*. Reading, 1963.
Tisza, L.: *Generalized Thermodynamics*. Cambridge, Massachusetts, 1966.
Toulmin, S.: «Matter». En *The Encyclopedia of Philosophy* (comp. por P. Edwards), 1967.
Truesdell, C. A.: «Rückwirkungen der Geschichte der Mechanik auf die moderne Forschung». *Humanismus und Technik*, 13, 1969.
Tuomela, R.: *Theoretical Concepts*. Viena, 1973.
Wells, R. S.: «Frege's Ontology». En *Essays on Frege* (comp. por E. D. Klembe). Chicago, 1968.
Wittgenstein, L.: *Tractatus logico-philosophicus*. 1.ª ed., 1921; 2.ª ed., Francfort, 1960.

Exploraciones metacientíficas
C. Ulises Moulines

EXPLORACIONES METACIENTIFICAS es básicamente una obra de filosofía de la física, pero también se ocupa de aspectos relacionados con historia de la filosofía y de las restantes ciencias. C. ULISES MOULINES, profesor de filosofía de la ciencia en diversas universidades de Alemania y Estados Unidos y actualmente miembro del Instituto de Investigaciones Filosóficas de México, es —en opinión de Jesús Mosterín, prologuista del volumen— «el filósofo de la ciencia más descollante de la joven generación latinoamericana e hispánica». Tras dos capítulos introductorios en torno a la naturaleza de la epistemología, el resto de la obra —subtitulada significativamente ESTRUCTURA, DESARROLLO Y CONTENIDO DE LA CIENCIA— se divide en tres grandes partes. La primera sección —«Análisis estructural de la ciencia»— se ocupa de la reconstrucción formal de las estructuras sincrónicas subyacentes a las teorías científicas y de las relaciones entre ellas. El enfoque básico para estos análisis es la metodología de la llamada concepción estructuralista de las teorías científicas, pero no se limita a formulaciones de carácter general, sino que ilustra con ejemplos concretos, obtenidos de la ciencia real, el desarrollo conceptual del razonamiento; la mecánica clásica, la termodinámica y la relación entre la teoría de Kepl la teoría de la gravitación son algunos de los casos reconstrui con todo detalle. La segunda parte —«Génesis y desarrollo de programas científicos»— contiene estudios sobre diversos aspe diacrónicos de las teorías y metodologías concretas que han surgido en la historia de la ciencia: la revolución newtoniana, desarrollo de la mecánica clásica durante la Ilustración, la fisiología de Helmholtz y el positivismo decimonónico como metodología de las ciencias naturales. La última parte —«Hacia una ontología de la ciencia»— se ocupa de algunos temas pertenecientes a la ontosemántica de la ciencia, en especial las ideas de Frege aplicadas en las ciencias empíricas y las dificultades que enfrentan el concepto de materia y la doctrina materialista.