

Condiciones físicas y metafísicas para un enfoque de la causación aplicable en las teorías físicas del espacio tiempo.

Autor:

Herrera Aros, Manuel Jesús

Tutor:

Lombardi, Olimpia

2021

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título de Doctor de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires en Filosofía.

Posgrado



Universidad de Buenos Aires
Facultad de Filosofía y Letras

Condiciones físicas y metafísicas para un enfoque de la causación aplicable en las teorías físicas del espacio–tiempo

Tesis de doctorado

Doctorando: Manuel Jesús Herrera Aros

Directora de tesis: Dra. Olimpia Lombardi

Co-director de tesis: Dr. Sebastian Fortin

Buenos Aires, 2021

Índice

Prefacio	1
Introducción: objetivos y estructura del trabajo	3
Capítulo 1: Teorías “difference-making” de la causación	9
Resumen	9
1.1. El problema de la causación	10
1.2. Teorías regularistas	14
1.3. Teorías contrafácticas	19
1.4. Teorías probabilistas	25
1.5. Teorías de agencia e intervención	29
1.6. Comentarios finales	35
Capítulo 2: Teorías de la causación física	37
Resumen	37
2.1. Teoría de la transferencia de Aronson	38
2.2. Teoría de la transferencia de Fair	41
2.3. Algunos comentarios sobre las teorías de Aronson y Fair	44
2.4. Teoría de procesos de Salmon (y Dowe)	45
2.5. Comentarios finales	51
Capítulo 3: Teoría de las Cantidades Conservadas	52
Resumen	52
3.1. El programa de Dowe	53
3.2. ¿Qué son los procesos e interacciones causales?	55
3.3. ¿Cuál es la conexión entre causas y efectos?	60
3.4. ¿Qué hace a una causa diferente de su efecto?	63
3.5. Comentarios finales	66
Capítulo 4: Críticas a la Teoría de las Cantidades Conservadas	67
Resumen	67
4.1. Transmitir en lugar de poseer	68
4.2. Identidad de los objetos	69
4.3. Identidad de las cantidades conservadas	71

4.4. Superposición de interacciones causales	73
4.5. Prevención y omisión	75
4.6. Críticas a la TCC desde el ámbito de la filosofía de la física	77
4.7. Leyes de conservación en relatividad general	79
Capítulo 5: Física de la causación: el problema de las leyes de conservación en la RG	83
Resumen	83
5.1. Introducción	85
5.2. Nociones básicas de RG	87
5.3. El problema de las leyes de conservación en la RG	102
5.4. Diferentes interpretaciones	112
5.5. El problema de las leyes de conservación y la TCC	128
Capítulo 6: Metafísica de la causación I: el problema de la posibilidad física	139
Resumen	139
6.1. El problema de la posibilidad física	140
6.2. El reino de la modalidad	143
6.3. La posición tradicional: el Enfoque Estándar de la modalidad física	148
6.4. Factores extra-nómicos en la posibilidad física	154
6.5. Dos casos de estudio	157
6.6. Leyes de conservación y posibilidad física	169
Capítulo 7: Metafísica de la causación II: el problema de la definición	176
Resumen	176
7.1. El problema general de la definición	177
7.2. Las tesis de Dowe	179
7.3. ¿Análisis empírico?	187
7.4. Definición de causación: ¿real o nominal?	193
7.5. La TCC: ¿una teoría refutable?	197
7.6. Una propuesta de solución al problema de la definición	199
Capítulo 8: Revisión general, conclusiones y perspectivas	202
Referencias	213

Prefacio

A lo largo de nuestra historia reciente, la causación ha sido, y sigue siendo, uno de los pilares fundamentales para el progreso conceptual de la física. El alcance de la causación en la filosofía de la ciencia y, en particular, en la filosofía de la física se ha acentuado aún más desde inicios del siglo XX con la aparición de las teorías de la relatividad general y de la mecánica cuántica. A partir de ese entonces, los aportes del concepto metafísico de la causación se han vuelto centrales a la hora de avanzar hacia nuevas explicaciones o desarrollos científico-teóricos. Y esto seguirá siendo de esta forma, ya que teorías como las de gravedad cuántica o la cosmología moderna, solo por nombrar algunas, necesitan de la causación, entre otros importantes conceptos provenientes de la filosofía, para dar explicaciones completas y satisfactorias a los problemas que abordan.

Por lo tanto, es importante estudiar, repensar y analizar los fundamentos de la causación de manera que esta línea de investigación siga siendo útil y adecuada para las explicaciones y fundamentos de las ciencias. En este sentido, esta tesis se pretende posicionar como un estudio de la causación física bajo la perspectiva de –al menos hasta ahora– nuestra mejor teoría física del espacio-tiempo, la relatividad general. En particular, se pretenden aportar argumentos que resulten esclarecedores para una adecuada comprensión de la actual noción de causación en ciencias físicas. Adicionalmente, se pretende aportar con algunas claves que permitan visualizar el camino hacia nuevos desarrollos en causación física, área del conocimiento filosófico que, como hemos mencionado, tiene una indubitable utilidad para la física teórica actual y para las ciencias en general.

Debo gratitud a muchas personas que de una u otra manera han hecho posible que esta investigación llegue a término. En primer lugar, quisiera agradecer profundamente a la Dra. Olimpia Lombardi, quien no solo ha dirigido esta tesis con gran sabiduría e inigualable compromiso, sino que además ha sido un apoyo constante en mi carrera profesional desde que tuve la fortuna de conocerla. Agradezco también al Dr. Sebastian Fortin por su apoyo académico como co-director de esta tesis. Varios comentarios y sugerencias que he incorporado en esta investigación han surgido en el marco de las discusiones con mis compañeros y compañeras del Grupo de Filosofía de las Ciencias de Buenos Aires, a todos

ellos y ellas, muchísimas gracias; en particular, quisiera agradecer a Hernán Accorinti y Cristián López, quienes amablemente me prestaron ayuda y fueron una guía para varios temas que abordo en este trabajo. No puedo dejar de mencionar aquí al Dr. Wilfredo Quezada, su erudición y pasión por la filosofía fueron una gran motivación para mí cuando daba mis primeros pasos en esta disciplina. Quiero agradecer también a Diego Romero, Ignacio Rojas, Camilo Silva y Omar Vásquez, muy buenos amigos con quienes he tenido la fortuna de compartir en diferentes instancias y discutir algunas partes de esta investigación. Por último, en el plano de lo personal, quiero agradecer a Emilia, mi compañera de vida. Iniciamos juntos, hace ya varios años, un camino común que ha estado repleto de sorpresas, un camino que ha sido maravilloso recorrer. Muchas gracias por compartir conmigo todos esos momentos.

Esta investigación ha sido llevada a cabo con el respaldo de una beca doctoral del Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET.

Buenos Aires, octubre de 2021.

Introducción: objetivos y estructura del trabajo

Indiscutiblemente, el problema de la causación ha ocupado un lugar preponderante en la investigación filosófica desde los albores de esta disciplina. Desde los antiguos filósofos griegos y aún hasta en las investigaciones filosóficas de nuestros días, muchos de los esfuerzos intelectuales se encuentran enfocados en develar la naturaleza de la causación. La vigencia de este problema metafísico no debería ser motivo de extrañeza. La causación en la actualidad no solo tiene un rol primordial dentro la filosofía misma, con las preguntas clásicas que comprende su estudio y los alcances que tiene dentro de esta disciplina, sino que también se ha vuelto relevante para un gran número de otras disciplinas científicas (por ejemplo, física, biología, farmacología, ciencias sociales y leyes, solo por nombrar algunas), donde los conceptos y elucidaciones asociados a la causación han encontrado un lugar preponderante. Esto último ha mantenido vivo el interés por el estudio de la causación y ha conservado este campo de investigación como uno de los más fructíferos y de mayor desarrollo dentro de la filosofía.

Debido, probablemente, a la misma incursión de la causación en otras áreas disciplinares, durante el siglo XX comenzaron a emerger diferentes teorías filosóficas, de naturaleza muy variada, que intentaron dar respuesta a la pregunta central acerca del problema de la causación, es decir, *qué es la causación*. Algunos de los enfoques más relevantes desde los cuales se aborda este problema son los regularistas, contrafácticos, probabilistas, de agencia e intervención y físicos de la causación. El progreso de las investigaciones en torno a este problema metafísico ha permitido que, en el marco de cada uno de estos enfoques, se puedan encontrar acalorados e interesantes debates, cada uno de los cuales es digno de estudio y análisis por derecho propio. En el caso de esta investigación, las teorías de la causación física conforman el marco general en el que se desenvuelve la discusión.

Los inicios de las teorías de la causación física datan de la década de 1970, cuando Jerrold Aronson (1971) y David Fair (1979) propusieron caracterizar las relaciones causales como la transferencia de ciertas cantidades físicas. Más tarde, Wesley Salmon (1984, 1994, 1997, 1998) y Phil Dowe (1992a, 1992b, 1992c, 1995a, 1995b, 1999, 2000, 2000a, 2001,

2009), en consonancia con los trabajos de Aronson y Fair, pero introduciendo una ontología de procesos, propusieron que la naturaleza de los procesos causales, a diferencia de los pseudo-procesos, estaría dada por la posesión y/o transferencia de cantidades físicas conservadas, esto es, cantidades gobernadas por leyes de conservación. El debate entre Salmon y Dowe respecto de la mejor manera de explicar las relaciones causales desde la física tuvo su punto culminante en la formulación de la *Teoría de las Cantidades Conservadas* (en adelante, TCC) por parte de Dowe en el año 2000. En la actualidad, la TCC sigue siendo la teoría de la causación física más robusta y desarrollada con la que se cuenta.

El análisis de Dowe de la causación no pretende ser de tipo conceptual sino de tipo empírico: no se trata de comprender el significado de la palabra “causa” tal como se la utiliza en el lenguaje, sino de establecer qué es la causación en el mundo. A su vez, el autor declara que del carácter “empírico” de su análisis se infiere que este procede a posteriori y da lugar a verdades contingentes. La TCC es también no-Humeana en un doble sentido: por una parte, no es un enfoque meramente regularista y, por otra parte, establece que la causación es una característica fundamental de nuestro mundo. Esta segunda característica de la teoría de Dowe pone en evidencia que su autor comparte una visión realista respecto de la causación. Es decir, considera que la causación es una característica objetiva de nuestro mundo, independiente de la mente de los individuos. Si Dowe concibe la causación como un rasgo objetivo con existencia independiente y, a su vez, considera que las ciencias y la física en particular pueden brindar las herramientas suficientes para caracterizar las relaciones causales, queda claro que Dowe es también un realista científico. Es decir, hay una confianza por parte del autor en que las ciencias, y específicamente la física, permiten en algún sentido, aunque sea mínimamente, conocer la “realidad” tal como es. Por último, es importante mencionar que la TCC es una teoría de la causación de naturaleza reduccionista, pues analiza las relaciones causales en términos no causales, en particular, en términos de la posesión y el intercambio de energía y momentum.

La TCC, en una versión simplificada, puede ser enunciada por las siguientes proposiciones (Dowe, 2000, p. 90):

CC1: Un proceso causal (*causal process*) es una línea de mundo de un objeto que posee una cantidad conservada.

CC2: Una interacción causal (*causal interaction*) es una intersección de líneas de mundo que involucra el intercambio de una cantidad conservada.

En relación a las proposiciones anteriores, se debe entender un *proceso* como una línea de mundo de un objeto (2000, p. 90). Una *línea de mundo* es la colección de puntos sobre un diagrama del espacio-tiempo de Minkowski que representa la historia del objeto (2000, p. 90). Un *objeto* es cualquier cosa encontrada en la ontología de la ciencia, tales como partículas, ondas o campos, o del sentido común, como sillas o mesas (2000, p. 91). Una *cantidad conservada* es cualquier cantidad gobernada por una ley de conservación (2000, p. 91). Y, por último, una *intersección* es la superposición de dos o más procesos (2000, p. 91).

Por supuesto, la TCC no ha permanecido exenta de críticas. Estas han sido múltiples y de naturaleza muy variada. Sin embargo, una corriente de críticas es de especial interés para esta investigación. Estas sostienen, en general, que la TCC no sería compatible con los espacio-tiempos relativistas generales. En efecto, las proposiciones *CC1* y *CC2* se encuentran fundamentadas en expresiones del tipo “poseer una cantidad conservada” e “intercambiar una cantidad conservada”; por lo tanto, resulta necesario contar con genuinas leyes de conservación que permitan definir estas “cantidades conservadas”. Sin embargo, existe una vasta discusión, tanto en física como en filosofía de la física, acerca de si realmente la teoría general de relatividad (en adelante, RG) cuenta con genuinas leyes de conservación. Este hecho, sostienen los críticos, es profundamente problemático para la teoría de Dowe. Algunos de los autores que critican a la TCC en este sentido son Alexander Rueger (1998), Erik Curiel (2000, 2018), Agustín Vicente (2002), Vincent Lam (2005, 2010), Carl Hoefer (2009), Tracy Lupher (2009).

Los problemas que enfrenta la TCC en los escenarios relativistas generales son particularmente relevantes para esta investigación por dos razones. En primer lugar, estas críticas constituyen un argumento bastante sólido que hace tambalear los cimientos de la TCC; de no ser superadas, podrían dilapidar cualquier proyecto de la causación que pretendiera asociar las relaciones causales con cantidades físicas conservadas, en particular,

si este proyecto se encuentra en la misma dirección (como “análisis empíricos”) que históricamente han tomado los programas fisicalistas desde Aronson y Fair hasta Salmon y Dowe. En segundo lugar, y aún más importante que la razón anterior, esta corriente de críticas abre un camino de análisis con una riqueza particular que permite evaluar la compatibilidad de la TCC con los “universos” relativistas generales. En efecto, cuando se analizan en detalle tanto las críticas a la TCC como la propia respuesta de Dowe a ellas, salen a la luz tres problemas filosóficos independientes desde los cuales pueden evaluarse tanto el alcance como las limitaciones de la teoría de Dowe en la teoría de la gravitación de Einstein. Estos tres problemas son los siguientes:

- (1) *El problema de las leyes de conservación.* Este problema consiste en determinar en qué sentido preciso es posible hablar de la existencia (o no) de ecuaciones de conservación en la RG.
- (2) *El problema de la posibilidad física.* Este problema corresponde a la elucidación de la noción de “mundo físicamente posible”.
- (3) *El problema de la definición.* En este caso se intenta determinar si Dowe ofrece una caracterización o una definición (y de qué tipo) de las nociones de “proceso causal” e “interacción causal” en las proposiciones *CCI* y *CC2*.

En el presente trabajo se ofrecen argumentos independientes para cada uno de estos problemas y, a partir de estos argumentos, se establecen tres condiciones que se utilizarán como referencia para evaluar la eventual compatibilidad o incompatibilidad de la TCC con la RG.

Objetivos

Esta investigación tiene como *objetivo general*, establecer las condiciones físicas y metafísicas con las que debe contar la TCC, y cualquier teoría de la causación física, para que esta se muestre compatible con las características propias de las teorías físicas del espacio-tiempo, en particular, con la RG.

Asimismo, los *objetivos específicos* son los siguientes:

- Analizar críticamente las teorías de la causación física y, en particular, la TCC de Phil Dowe, de manera de realizar una evaluación general de ellas y de su alcance en el caso de nuestras mejores teorías físicas del espacio-tiempo.
- Poner de manifiesto que, desde la perspectiva de la filosofía de la ciencia y, en particular, desde la perspectiva de la filosofía de la física, se presentan algunas ambigüedades en el análisis reductivista-fisicalista de la causación realizado por Dowe.
- Mostrar cómo tales ambigüedades afectan el alcance y los límites de la TCC en el caso de los procesos descritos por la RG.
- Analizar nuestras actuales teorías físicas del espacio-tiempo con el fin de establecer las condiciones tanto físicas como metafísicas que de ellas se derivan y que, por tanto, toda teoría de la causación física debe cumplir.

Estructura del trabajo

Esta tesis se encuentra estructurada del siguiente modo. Los Capítulos 1 y 2 brindan el marco general en el que se desenvuelve esta investigación. Mientras en el Capítulo 1 examinaremos las teorías denominadas *difference-making*, en el Capítulo 2 nos enfocaremos en las teorías de la causación física. A continuación, en el Capítulo 3, se expondrá un análisis detallado de la teoría de la causación física que es objeto de estudio en esta investigación: la TCC de Dowe. Luego, en el Capítulo 4, se brindará una extensa exposición de las principales críticas que se han formulado a la TCC, poniendo especial énfasis en la corriente de críticas que han motivado esta investigación: el problema de las leyes de conservación en la RG. El Capítulo 5 se ocupará del *problema de las leyes de conservación*, argumentando en favor de la presencia de leyes de conservación integrales en algunos modelos o soluciones de las ecuaciones de campo de Einstein. En el Capítulo 6 se abordará el *problema de la posibilidad física*, donde se argumentará que la noción de posibilidad física no solo se encuentra determinada por aquello que dictan las ecuaciones dinámicas de una teoría, como dictamina cierto enfoque estándar, sino que también intervienen algunos factores extra-nómicos en su formulación. Luego, en el Capítulo 7, se

examinará el *problema de la definición*, argumentando que las afirmaciones nucleares de una teoría de la causación deben presentarse como definiciones reales. Finalmente, en el Capítulo 8, se llevará a cabo una revisión general de los argumentos presentados a lo largo de esta tesis y se expondrán las principales conclusiones.

Capítulo 1. Teorías “*difference-making*” de la causación

Resumen

Si bien el estudio de la causación es de larga data, a partir de la segunda mitad del siglo XX se observa un renovado interés por el estudio de este importante tópico metafísico. En particular, en ese momento comenzaron a desarrollarse variados enfoques filosóficos con el objetivo de explicar la naturaleza de las relaciones causales aludiendo a –o haciendo uso de– los más diversos conceptos tomados no solo del campo que comprende la filosofía, sino que también de otras disciplinas científicas. La motivación principal detrás de estas teorías es intentar dilucidar qué es la causación, cómo podemos capturar aquello que es esencial de este tipo de relaciones o, dicho de otro modo, a qué hacemos referencia exactamente en el mundo cuando afirmamos que una cosa causó otra.

Una de las maneras de agrupar las teorías filosóficas contemporáneas de la causación es, por un lado, en aquellas que se conocen como teorías *difference-making* y, por otro lado, en las teorías fisicalistas de la causación. Las del primer grupo están guiadas por la idea de que las causas “hacen una diferencia” en sus efectos o, en otras palabras, que el efecto no hubiera sido el mismo sin la presencia de la causa. En cambio, las del segundo grupo se basan en la intuición de que la causa y el efecto están conectados por algún tipo de mecanismo, proceso o cantidad física.

Haciendo uso de la distinción anterior, en este capítulo examinaremos las teorías filosóficas de la causación etiquetadas como teorías *difference-making*. Los principales análisis de las relaciones causales que se enmarcan en esta categoría son, en primer lugar, las teorías *regularistas* que, de modo muy general, sostienen que las relaciones causales no son más que la instanciación de ciertas regularidades o patrones. En segundo lugar, pertenecen a este grupo las teorías *contrafácticas* de la causación, que afirman que el “hacedor de diferencia” (*difference-maker*) presente en las relaciones causales puede ser caracterizado en términos de dependencia contrafáctica, haciendo uso de ciertos condicionales contrafácticos. El tercer tipo de teoría de la causación que puede ser clasificada como *difference-making* es la *probabilista*. En este caso, la diferencia que la causa produce en su efecto quedaría capturada, muy aproximadamente, bajo la premisa de

que las causas aumentan la probabilidad de sus efectos. Por último, las teorías *de agencia e intervención* son el cuarto tipo de teoría de la causación que puede ser encasillada bajo el rótulo de teoría *difference-making*. En estos análisis, la intuición central es que las relaciones causales son un tipo de relación que admite algún tipo de manipulación en términos de agencia humana o intervención. De este modo, algo es causa de un efecto si mediante su manipulación se puede producir un cambio o una diferencia en el efecto. Estas cuatro teorías filosóficas acerca de la naturaleza de las relaciones causales serán el foco del análisis que se llevará a cabo en este capítulo. Las teorías de la causación física serán motivo de examen en el Capítulo 2.

1.1.El problema de la causación

La causación¹ es uno de los tópicos de la filosofía que ha sido objeto de debate desde los mismos inicios de esta disciplina hasta nuestros días. La importancia de la causación como problema filosófico no solo remite al valor que este tiene en sí mismo, sino que también reviste una importancia significativa para un gran número de otros problemas filosóficos, como el problema de las leyes de la naturaleza, el problema de la explicación, el problema del libre albedrío, el problema del tiempo, solo por nombrar algunos. La causación también juega un papel sustancial en las más diversas áreas de la filosofía, como filosofía de la ciencia, filosofía de la mente, ética, epistemología, entre otras. Adicionalmente, las investigaciones filosóficas acerca de la causación también encuentran un nicho en otras disciplinas diferentes a la filosofía, como economía, biología, ciencias médicas, química, física, entre otras. La amplia gama de problemas filosóficos, áreas de la filosofía y diferentes disciplinas científicas en las cuales la causación tiene algo para decir es, quizás, una de las razones por las cuales este concepto filosófico ha sido tan ampliamente estudiado e investigado. Probablemente, también sea este uno de los motivos por los cuales en la

¹ Si bien los términos “causación” y “causalidad” a veces son usados indistintamente, en esta tesis usaremos el término “causación” para dar cuenta de la conexión causal o, dicho de otro modo, para dar cuenta de cómo, exactamente, se encuentran conectados la causa y el efecto, y reservaremos el término “causalidad” para indicar la doctrina causal en general o para designar el punto de vista que defiende la validez del “principio causal”, esto es, que la misma causa siempre produce el mismo efecto. ver Bunge (2008, Cap. 1, Sección 1.1), para más detalles acerca de estas distinciones.

bibliografía podemos encontrar un sinnúmero de debates desde los más variados enfoques acerca de este concepto metafísico.

Esta tesis se enfoca en las teorías contemporáneas de la causación. Sin embargo, como se ha mencionado en el párrafo anterior, la naturaleza de las relaciones causales ha sido estudiada desde los inicios de la filosofía misma, por lo que gran parte de las teorías contemporáneas de la causación encuentran sus raíces –o están inspiradas en– los principales análisis históricos acerca de este concepto metafísico. Sin lugar a dudas, los primeros antecedentes se encuentran en Aristóteles –y luego en los escolásticos, quienes afianzaron la doctrina causal aristotélica– a través de su análisis de los fenómenos naturales sobre la base de cuatro tipos de causas: material, formal, final y eficiente. Con la entrada en escena de los filósofos y filósofas (o científicos y científicas) de los siglos XVI y XVII, la doctrina aristotélica de las cuatro causas comenzó a cuestionarse y lentamente las explicaciones basadas en las causas eficientes comenzaron a prevalecer por sobre las demás. Un poco después de este cambio de paradigma en relación a la corriente aristotélica del estudio de la causalidad, filósofos como David Hume e Immanuel Kant vinieron a revolucionar la escena filosófica en general y también, junto con ello, nuestras intuiciones acerca de la causación. Finalmente, los últimos antecedentes significativos antes de arribar a las teorías contemporáneas de la causación se deben a los empiristas lógicos, quienes comenzaron a estudiar la causación desde una mirada “científica” que intentaba ser compatible con los notables avances teóricos y experimentales de la ciencia, en particular de la física, propios de la época.

Las teorías contemporáneas de la causación, en general, intentan echar luz sobre la comprensión más adecuada de la naturaleza de las relaciones causales. En otras palabras, estas teorías pretenden responder a la pregunta acerca de *qué es* la causación en el mundo o qué queremos exactamente decir cuando hacemos afirmaciones del tipo “*C* causa *E*”, si es que hay algo en el mundo que pueda ser llamado “causa” o “efecto”, por supuesto. Dicho de otro modo, en esta investigación nos ocuparemos de cuál es la base o qué fundamenta una conexión causal². Por ejemplo, cuando decimos que la causa del movimiento de una

² Un análisis metafísico completo de la causación involucra básicamente dos aspectos: el problema de la relación causal y el problema de los *relata* causales. A su vez, el problema de la relación causal comprende tres sub-problemas: (i) la conexión causal, es decir, qué diferencia aquello que se

carga eléctrica es la presencia de una segunda carga, ¿qué significa “causa” en esta afirmación? Se han ofrecido algunas respuestas a este tipo de preguntas, muy diferentes entre sí, constituyendo cada una de ellas una teoría diferente acerca de la naturaleza de las relaciones causales. Es así como a partir de la segunda mitad del siglo XX comenzaron a proliferar diferentes intentos por capturar aquellas características distintivas de las relaciones causales. Algunos de los análisis más notables en esta dirección son las teorías regularistas, contrafácticas, probabilistas, de agencia e intervención y físicas³.

Ahora bien, es importante puntualizar que una visión general acerca de la naturaleza de la causación y desde la cual teorizar acerca de este concepto metafísico puede ser asumida desde diferentes enfoques. Un primer enfoque ampliamente extendido es el *enfoque reduccionista*, esto es, el que subyace a aquellas teorías de la causación que utilizan términos no causales para caracterizar la causación. El análisis contrafáctico de David Lewis y el análisis fisicalista de Salmon-Dowe, en los que las relaciones causales se reducen a términos no causales como dependencia contrafáctica y procesos, respectivamente, son ejemplos de este tipo de enfoque. La antítesis del reduccionismo es el *enfoque anti-reduccionista*, es decir, el punto de vista que sostiene que la causación no puede ser analizada en términos de conceptos no causales. También puede mencionarse el *primitivismo causal*, una visión ligeramente diferente del anti-reduccionismo, según el cual la causación es un concepto básico (primitivo) y no habría conceptos más fundamentales que este en términos de los cuales analizar las relaciones causales. En general, un o una primitivista causal afirmaría que la causación es un concepto inanalizable. Por otro lado, el

encuentra causalmente relacionado de lo que no se encuentra causalmente relacionado; (ii) la dirección causal, esto es, si la relación causal es simétrica o asimétrica y (iii) la selección causal, o cómo diferenciar las secuencias causales genuinas de las meras condiciones de fondo (ver Schaffer, 2016). En esta investigación nos enfocaremos en el problema de la conexión causal, dejando de lado el examen acerca de la dirección causal, la selección causal y el problema de los *relata* causales.

³ Otros análisis alternativos a los análisis estándar de la causación mencionados aquí son la teoría de fuerzas (debida principalmente a Bigelow, Ellis y Pargetter, 1988), los enfoques disposicionales (ver, por ejemplo, Ellis, 2001; Mumford y Anjum, 2011; Bird, 2010) y los enfoques causales provenientes del denominado “nuevo mecanicismo” (por ejemplo, Glennan, 1996). Para un estudio que aborda el problema de la causación analizando ciertas “restricciones causales” (como por ejemplo, localidad y estabilidad) utilizadas en ciencia –más que un análisis reductivo de relaciones entre *relatas*–, ver Ben-Menahem (2018). Finalmente, para un estudio novedoso que asocia las relaciones causales con los estados físicos de un sistema dinámico, ver Fazekas et.al. (2021).

pluralismo causal sostiene que no existe un único tipo de relación causal o, dicho de otro modo, que la causación puede ser analizada de diferentes maneras dependiendo del tipo de fenómeno que se pretenda explicar causalmente. Por último, un enfoque general acerca de la causación que atraviesa los enfoques anteriores es el *realismo causal*, es decir, el punto de vista que afirma que la causación es una característica distintiva del mundo. La contraparte del realismo causal es el *anti-realismo causal*. Un tipo de anti-realismo causal es el *eliminativismo* (debido principalmente a Bertrand Russell), que afirma que la causación no juega ningún papel significativo en la ciencia, pues este concepto no existiría en el marco de este cuerpo de conocimientos, y tampoco es tarea de la ciencia buscar causas en los fenómenos naturales. Gran parte de las teorías de la causación que serán examinadas en esta investigación pueden ser calificadas como reduccionistas, algunas o algunos de sus autores estarán comprometidos en mayor o menor grado con algún tipo de realismo causal y gran parte de ellos o ellas son contrarios a un punto de vista eliminativista.

Por último, se debe recordar que las teorías contemporáneas de la causación pueden ser agrupadas de diferentes maneras dependiendo del criterio que se utilice para ello. Una primera clasificación de las teorías filosóficas de la causación las divide entre teorías Humeanas, por un lado, y no Humeanas, por otro lado. Esto es, si los análisis o caracterizaciones de la causación se reducen a afirmaciones no causales, libres de todo requerimiento de necesidad o de apelaciones a poderes o disposiciones, estaremos hablando de una teoría Humeana de la causación; de lo contrario tendremos una teoría no Humeana (Schrenk, 2016, p. 212). En segundo lugar, los análisis de la causación pueden ser divididos en reductivos versus no reductivos, esto es, si la caracterización de la causación utiliza un vocabulario causal o no, respectivamente, para definir este concepto metafísico (Schrenk, 2016, p. 213). También se distinguen aquellas teorías de la causación intrínsecas en contraposición con las extrínsecas. Las del primer tipo, a diferencia de las del segundo tipo, son aquellas que caracterizan a la causación como dependiente de la relación causal misma, sin apelar a nada externo a ella (Rueger, 1998; Schrenk, 2016, p. 213). En cuarto lugar, es posible clasificar los análisis causales dependiendo de si son compatibles con casos de causación indeterminista o no (ver, por ejemplo, Dowe, 2000). Por último, los enfoques contemporáneos de la causación pueden ser divididos en teorías *difference-making*, por una

parte, y teorías fisicalistas, por otra parte⁴. La idea subyacente a las teorías *difference-making* es que una causa “marca una diferencia” para la obtención del efecto, ya que sin tal diferencia este último no se habría obtenido. En cuanto a las teorías fisicalistas, la intuición guía detrás de ellas es que existiría algún tipo de mecanismo, proceso o cantidad física que conecta la causa con el efecto.

Como se ha señalado más arriba, el marco teórico general de esta investigación corresponde a las teorías contemporáneas de la causación. Este marco teórico se expondrá en este capítulo y en los dos siguientes. Con el objetivo de presentar con la mayor claridad posible los distintos análisis o caracterizaciones de la causación, se ha elegido utilizar la última de las clasificaciones mencionadas en el párrafo anterior, esto es, dividir las teorías de la causación en aquellas denominadas como *difference-making*, por un lado, y aquellas teorías clasificadas como fisicalistas, por otro lado. En particular, en este capítulo se presentan las teorías *difference-making* de la causación, que comprenden las teorías regularistas, contrafácticas, probabilistas y de agencia e intervención. En el Capítulo 2 se expondrán las teorías fisicalistas de la causación, que involucran las teorías de transferencia y las teorías de proceso. Por último, en el Capítulo 3, profundizaremos con todo detalle en la teoría de la causación física que es el objeto de estudio de esta tesis: la Teoría de las Cantidades Conservadas de Dowe.

1.2. Teorías regularistas

Sin lugar a dudas, los trabajos de Hume (1978 [1739], 2007 [1748]) fueron un punto de inflexión para las discusiones filosóficas acerca de la causación. Hume no solo reavivó el debate en torno a estos asuntos, sino que también desarrolló un influyente análisis acerca de la naturaleza de la causación en términos del concepto de “regularidad” que se mantiene vigente hasta nuestros días y a partir del cual se han desarrollado variadas propuestas alternativas. En términos generales, bajo el enfoque regularista estándar de la causación⁵ diremos que c causa e ⁶ si y solo si:

⁴ Para un estudio de la relación entre ambos enfoques, ver Ney (2009).

⁵ Presentamos aquí la respuesta positiva de Hume al problema de la causación, en la cual el filósofo escocés intenta dar respuesta a la pregunta acerca de la naturaleza de este concepto metafísico. Ver

1. c es contiguo a e (contigüidad espacio-temporal).
2. c precede temporalmente a e (prioridad temporal).
3. Los eventos del tipo c son regularmente seguidos por eventos del tipo e (conjunción constante).

Las tres condiciones anteriores implican tres características que forman el núcleo de la propuesta regularista Humeana. En primer lugar, (1) implica que la causa y el efecto deben ser cercanos el uno del otro; o más específicamente, estos deben ser próximos en términos espacio-temporales. En segundo lugar, a partir de (2) la causa se considera como temporalmente anterior al efecto. De acuerdo con este punto de vista, la dirección de la causación se encuentra fundamentada en la dirección temporal. Por último, (3) involucra el aspecto más importante de las teorías de tipo regularista: en nuestra vida cotidiana vemos que los eventos del tipo c son *regularmente* seguidos por eventos del tipo e . Por ejemplo, si damos un golpe de puño a una caja de cartón podremos constatar que esta se hunde y, en general, podremos verificar que los golpes de puño sobre cajas de cartón (C) van seguidos por un hundimiento de ellas (E). Hume llama a esto *conjunción constante*. Por lo tanto, desde el punto de vista de las teorías regularistas de la causación, aquello que fundamenta las relaciones causales es justamente la noción de conjunción constante que se evidencia en las regularidades que experimentamos a diario.

La motivación metafísica detrás de las teorías regularistas es intentar eliminar la necesidad y la productividad que generalmente se asocian a las afirmaciones causales. De

Nota al Pie 6 para la respuesta negativa de Hume a este problema. Es importante mencionar, sin embargo, que ha sido objeto de debate si efectivamente Hume ofrece una análisis acerca de qué es la causación en el mundo. Existen otras interpretaciones que sostienen que Hume, más bien, habría desarrollado un análisis conceptual (Ducasse, 1976) de la noción de causa. También hay quienes afirman que Hume tenía cierta creencia en la existencia de poderes en la naturaleza (Wright, 1983; Craig, 1987; Strawson, 2014). Ver Dowe (2000, p. 21) para más detalles acerca de esta discusión.

⁶ A lo largo de esta tesis, a menos que se indique lo contrario, usaremos la notación estándar para designar la causa y el efecto en una relación causal. Las letras c y e minúsculas designan eventos *token* y las letras C y E mayúsculas se reservan para designar eventos *type*. De este modo, para los enfoques regularistas de la causación, si una sucesión de eventos es causal dependerá de si los eventos *token* c y e que forman esa relación pertenecen a eventos *type* C y E , respectivamente (donde estos últimos cumplen con la condición (3), por supuesto). Esto también se relaciona con la discusión entre aquellos que defienden la causación singular y aquellos que abogan por la causación general. Las visiones regularistas de la causación iniciadas por Hume defienden que la verdad de las afirmaciones causales singulares depende de las regularidades existentes en el mundo, donde las afirmaciones causales singulares pueden ser vistas como instancias de estas regularidades generales.

acuerdo con este punto de vista, no existe una conexión necesaria entre la causa y el efecto, sino que únicamente se encuentra presente nuestra creencia, fundamentada en la experiencia de ciertos patrones regulares, de que a ciertas causas deben seguirle ciertos efectos. Recordemos que, desde un punto de vista empirista, la validez de una idea se encuentra determinada por nuestra experiencia sensible. Por lo tanto, dado que no podemos percibir la idea de una conexión necesaria entre causa y efecto, esta debe ser desechada⁷. De igual manera, la noción de productividad (claramente presente, por ejemplo, en la noción de poder causal) que normalmente asociamos a la causación no tiene asidero en un enfoque regularista. La visión metafísica de que la causación debe reducirse a patrones regulares junto con la ausencia o eliminación de ciertos elementos metafísicos adicionales para fundamentar las regularidades, en particular la necesidad y productividad comúnmente atribuidas a la causación, conforman el núcleo metafísico de las visiones regularistas.

Grandes filósofos, en su búsqueda de una teoría de la causación, han seguido este credo metafísico. Uno de ellos es John Stuart Mill (2011 [1843]), quien afirma que las causas pueden ser vistas como condiciones suficientes para la ocurrencia del efecto. En particular, el autor argumenta que una causa se encuentra compuesta por una serie de factores que, en conjunto, son condiciones suficientes para el efecto. Para Mill existen dos tipos de condiciones suficientes: las positivas y las negativas. La diferencia esencial entre las condiciones positivas y negativas es que, en el primer caso, las condiciones o factores que componen la causa deben estar presentes; en cambio, con el segundo término Mill hace referencia a la ausencia de ciertos factores. Por lo tanto, en el enfoque Milleano de la causación, una causa corresponde a la conjunción de las condiciones suficientes, positivas y negativas que, instanciadas en conjunto, permiten la ocurrencia del efecto⁸.

⁷ Esto corresponde a la doctrina negativa de Hume acerca de la causación.

⁸ Es importante mencionar dos cuestiones respecto del enfoque Milleano de la causación. En primer lugar, se debe notar que, debido a que la causa se compone de múltiples factores donde algunos de ellos sean quizás irrelevantes, solo se consideran aquellos factores que se encuentran invariablemente conectados con el efecto. Por ejemplo, si se quiere investigar la causa de un accidente automovilístico, es en principio irrelevante considerar el día de la semana en que este ocurrió como un factor que sea parte de la causa. En segundo lugar, también debemos considerar que para evitar factores triviales como constitutivos de una causa, se debe tener en mente que estos factores deben ser vistos más bien como una serie de elementos que conforman el contexto de fondo sobre el cual ocurre la relación causal. Esta es la base de la noción de “campo causal” de Mackie, la cual revisaremos a continuación. Ver Psillos (2009) para más detalles acerca de estas cuestiones.

En cuanto a otro aspecto acerca del estudio de la causación, Mill argumenta que la sucesión regular Humeana es insuficiente para capturar la esencia de las relaciones causales. Para el autor es necesario incorporar la idea de “incondicionalidad”, sin la cual no sería posible distinguir entre regularidades genuinas y meros accidentes. De este modo, para Mill las causas son las condiciones suficientes a partir de las cuales se sigue el efecto bajo cualquier circunstancia. A su vez, según el autor, la incondicionalidad de una secuencia de eventos se encontrará garantizada por una ley de la naturaleza⁹.

Siguiendo el camino trazado Mill, John L. Mackie (1980 [1974]) defiende una versión diferente acerca de aquello que cuenta como causal, pero manteniendo los principios metafísicos que sustentan a los puntos de vista regularistas. Recordemos que Mill había afirmado que generalmente una causa corresponde a un conjunto de eventos (o factores). Por ejemplo, decimos que la causa del vaso quebrado en el piso se debe a la fragilidad del material con que está hecho el vaso, la dureza del piso sobre el que cae, un movimiento involuntario de la superficie en la que se encontraba el vaso, etc. Todos estos factores conforman la causa del vaso quebrado. En el lenguaje de Mackie, este tipo de causas se denominan *causas complejas*.

Es importante resaltar que las causas complejas, aunque suficientes, casi nunca son necesarias para el efecto. Es decir, una causa compleja particular es *suficiente* para el efecto, esto es, dada esta causa compleja el efecto se seguirá invariablemente; pero a su vez, podríamos obtener otra causa compleja que produzca el mismo efecto, por lo que también es *innecesaria*. Por otro lado, los factores causales singulares que conforman la causa compleja son *insuficientes*, pero *necesarios* para el efecto. La conjunción de estas condiciones tanto para las causas complejas como para los factores singulares que constituyen esta causa compleja se denominan *condiciones INUS* (por las iniciales en inglés de las cuatro condiciones). De este modo, bajo la concepción de Mackie de la causación, un evento *C* es causa de un evento *E* si *C* es una condición INUS para *E*.

⁹ Desde el punto de vista empirista Humeano, aquella explicación de ley de la naturaleza que intervenga en las explicaciones causales debe ser metafísicamente compatible con la visión de causación Humeana, es decir, una explicación que no apele ni a poderes ni a ningún tipo de necesidad.

Se han formulado, por supuesto, variadas críticas a la visión de Mackie acerca de la causación y su conexión con condiciones necesarias y suficientes¹⁰. Una de ellas tiene relación con lo siguiente. Supongamos que *ABCD* es la causa compleja del efecto *E*. Es posible verificar que cada uno de los factores causales singulares *A*, *B*, *C* y *D* son condiciones INUS para el efecto *E*. Sin embargo, este hecho podría traer ciertas consecuencias indeseables, pues permitiría la incorporación de causas que en la práctica son triviales o irrelevantes. Por ejemplo, no diríamos que la presencia de oxígeno es un factor causal relevante cuando el efecto es la rotura de una ventana producto del golpe con un balón de fútbol. La solución a estas dificultades por parte de Mackie fue incorporar la noción de *campo causal*. El campo causal corresponde a una serie de suposiciones relacionadas con ciertos factores causales que estarían presentes o no de acuerdo al contexto analizado. En relación al ejemplo anterior, la presencia de oxígeno en la rotura de una ventana o en una colisión de dos automóviles no debe ser considerado como un factor causal, sino más bien como una suposición implícita acerca de las condiciones atmosféricas del lugar del accidente.

Antes de finalizar esta breve revisión acerca de los enfoques regularistas de la causación deben notarse dos asuntos relevantes. En primer lugar, como se ha mencionado más arriba, estos enfoques de la causación iniciados con el influyente trabajo de Hume han servido como base para otros enfoques acerca de la naturaleza de las relaciones causales que mantienen la esencia metafísica del análisis Humeano pero que, sin embargo, se distancian en varios aspectos de él (ejemplos de ello son los análisis de Mill y Mackie vistos anteriormente). Este hecho no implica que los enfoques regularistas estándar se encuentren “pasados de moda”; muy por el contrario, estos se encuentran lejos de pasar al olvido. Si bien ya no con la misma fuerza que en épocas pasadas, aún existe una amplia línea de investigación en relación a este tipo de análisis de las relaciones causales (ver, por ejemplo, May y Grabhoff, 2001; Baumgartner, 2008). En segundo lugar, es importante puntualizar que las críticas a los enfoques regularistas de la causación, ya sea en sus versiones más cercanas a la propuesta de Hume o en aquellas versiones alternativas que conservan la propuesta metafísica regularista, han sido múltiples y variadas, por lo que

¹⁰ Ver, por ejemplo, Psillos (2002, Cap. 3, Sección 3.2) y Psillos (2009) para más detalles acerca de las críticas a la teoría INUS de Mackie.

sería imposible ofrecer una revisión aunque sea mínimamente detallada de ellas aquí. Para un examen de estas críticas se sugiere ver David Lewis (1973b), Michael Baumgartner (2008) y Christopher Hitchcock (2018), entre otros.

Una teoría de la causación de inspiración Humeana que enfrenta adecuadamente las críticas de las que son objeto los enfoques regularistas es la teoría contrafáctica de Lewis. Este será el tema de análisis de la siguiente sección.

1.3. Teorías contrafácticas

En general, las teorías contrafácticas de la causación se basan en la idea de que las relaciones causales se encuentran fundamentadas en la verdad de ciertos enunciados contrafácticos. Es decir, bajo el alero de una teoría contrafáctica de la causación diríamos que c es causa de e , si el enunciado “si c no hubiera ocurrido, e no habría ocurrido” es verdadero. Las teorías de este tipo, en consecuencia, son teorías que basan la causación en la noción de dependencia, en este caso contrafáctica, entre la causa y el efecto. Uno de los análisis contrafácticos de la causación más influyentes se debe a David Lewis, quien formuló la primera versión de su teoría en 1973. Esta primera versión fue dando paso a versiones posteriores más refinadas, producto del impacto que tuvo su análisis en la comunidad filosófica y las consecuentes críticas que debió enfrentar. Por lo tanto, basaremos esta sección principalmente, aunque no exclusivamente, en el artículo de Lewis “*Causation*” del año 1973.

Como se ha mencionado en el párrafo anterior, la teoría de la causación de Lewis basa su análisis en la noción de dependencia, en particular, en la noción de dependencia contrafáctica, la cual a su vez depende de una semántica de mundos posibles que permite a su autor establecer las condiciones de verdad de los contrafácticos en cuestión en términos de la noción de similaridad entre mundos posibles. En consecuencia, el primer paso de Lewis en la elaboración de su teoría de la causación es especificar el uso de una semántica de mundos posibles, que le permitirá evaluar los enunciados contrafácticos sobre los cuales se basa su teoría.

Considérense dos proposiciones A y C a partir de las cuales se forma el condicional contrafáctico “si A no sucediera, entonces C no ocurriría”. De acuerdo con Lewis, el

condicional contrafáctico formado por A y C es verdadero en un mundo w si y solo si, o bien (i) no hay A -mundos posibles, en cuyo caso el condicional contrafáctico es vacuo, o bien (ii) algún A -mundo donde C ocurre es más cercano (a w) que cualquier A -mundo donde C no ocurre (1973b, p. 560). Ciertamente, la regla anterior para evaluar la verdad de un enunciado contrafáctico en un mundo w depende de qué entendamos por mundo posible “más cercano” a otro. Conciente de esto, Lewis propone la noción de *similaridad comparativa global entre mundos posibles*. Para el autor, un mundo posible es “más cercano” a lo actual que otro si el primero se asemeja más al mundo actual que el segundo (1973b, p. 559). Esta relación de similaridad defendida por Lewis es considerada como primitiva en sus análisis contrafácticos; sin embargo, el autor menciona que es un tipo de relación muy familiar para nosotros, puesto que constantemente hacemos juicios de comparación en nuestra vida cotidiana¹¹.

Una vez definida la noción de similaridad comparativa global entre mundos posibles, que permite la evaluación de los enunciados contrafácticos, el segundo paso de Lewis es definir la idea de *dependencia contrafáctica*. Supongamos que contamos con dos familias de proposiciones. La primera familia está compuesta por las proposiciones A_1, A_2, \dots, A_n y la segunda familia por las proposiciones C_1, C_2, \dots, C_n . Si todos los contrafácticos correspondientes entre las proposiciones que componen ambas familias son verdaderos, entonces diremos que las proposiciones C_1, C_2, \dots, C_n *dependen contrafácticamente* de las proposiciones A_1, A_2, \dots, A_n , respectivamente (1973b, p. 561).

Esta noción de dependencia contrafáctica permite a Lewis dar el siguiente paso, ahora hacia la noción de *dependencia causal*. Para el autor, la dependencia contrafáctica es al mismo tiempo dependencia causal. Para ser más precisos, si una familia de proposiciones C_1, C_2, \dots, C_n depende contrafácticamente de otra familia de proposiciones A_1, A_2, \dots, A_n , estamos también habilitados para decir que las proposiciones C_1, C_2, \dots, C_n *dependen causalmente* de las proposiciones A_1, A_2, \dots, A_n (1973b, p. 561). Ahora bien, debe notarse que la teoría contrafáctica de la causación de Lewis es una teoría cuyos *relata* causales son eventos, mientras que hasta ahora hemos definido las nociones centrales de su teoría en

¹¹ Ver Lewis (1973a; 1973b, pp. 559-560) para más detalles acerca de la relación de similaridad comparativa global y de los análisis contrafácticos.

términos de proposiciones. Lewis convierte sus definiciones en términos de proposiciones en definiciones en función de eventos mediante una regla muy simple: “la dependencia contrafáctica entre eventos es simplemente dependencia contrafáctica entre las correspondientes proposiciones” (1973b, p. 562).

De este modo, Lewis arriba a la noción de dependencia causal entre eventos simples¹². Supongamos dos eventos particulares c y e , y sus correspondientes proposiciones asociadas $O(c)$ y $O(e)$. La dependencia causal entre eventos particulares se define del siguiente modo: “ e depende causalmente de c si y solo si la familia $O(e), \neg O(e)$ depende contrafácticamente de la familia $O(c), \neg O(c)$ ” (p. 563). Dicho de otro modo, tal como señala Lewis, “si e ocurre o no depende de si c ocurre o no” (1973b, p. 563).

Se debe notar el siguiente hecho relevante en la definición de dependencia causal. Supongamos que Ana golpea un balón de fútbol que viaja directo a un vaso de cristal que se encuentra sobre una mesa, y el vaso cae y se rompe. En esta situación intervienen dos eventos: el golpe de Ana al balón (c) y la rotura del vaso de cristal (e). Para evaluar si existe una dependencia causal entre c y e , debemos determinar si los siguientes contrafácticos son verdaderos: (i) “si Ana hubiera golpeado el balón, el vaso se habría quebrado”, y (ii) “si Ana no hubiera golpeado el balón, el vaso no se habría quebrado”. Ante la situación planteada, el contrafáctico (i) es trivialmente verdadero, pues tanto c como e en efecto ocurrieron; por lo tanto, si existe dependencia causal entre estos eventos dependerá de la verdad del contrafactual (ii). Este contrafáctico parece ser verdadero debido a que los mundos posibles donde Ana no golpea el balón y el vaso se quiebra son mundos “más alejados” del mundo actual que aquellos donde Ana no golpea el balón y el vaso no se quiebra. Dado que tanto (i) como (ii) son verdaderos, entonces efectivamente existe una dependencia causal entre los eventos c y e . Ahora bien, el lector podría preguntarse qué sentido tiene evaluar el contrafáctico (i) si los eventos c y e de hecho ocurrieron. La respuesta a esta inquietud se encuentra en el hecho de que la noción de dependencia causal de Lewis pretender capturar este tipo de dependencia independientemente de lo que ocurre en el mundo actual. De hecho, si los eventos c y e no

¹² Lewis también ofrece una definición de dependencia causal en términos de familias de eventos. Ver Lewis (1973b, p. 562).

hubieran ocurrido, ahora sería el contrafáctico (ii) el trivialmente verdadero y la dependencia causal entre estos eventos se dirimiría por la verdad del contrafáctico (i). Sin embargo, como se mencionará a continuación, para Lewis la causación se define entre eventos actuales.

Volvamos nuestra atención hacia el cuarto paso de Lewis en su formulación de una teoría contrafáctica de la causación. Una vez definida la noción de dependencia causal, el siguiente paso de Lewis es establecer la conexión entre esta noción, pero considerando eventos actuales, y la noción de causación. En palabras de Lewis, “si c y e son dos eventos actuales tales que e no habría ocurrido sin c , entonces c es causa de e ” (1973b, p. 563). Es decir, dos eventos actuales que dependen causalmente implican causación. Sin embargo, la inversa no es cierta: la causación no implica dependencia causal. En otras palabras, la dependencia causal entre eventos actuales es suficiente, pero no necesaria para la causación.

Lewis es consciente de la problemática recién señalada, esto es, que puedan existir casos de causación sin dependencia causal. La solución que ofrece el autor a este problema es reformular la noción de dependencia causal de manera que esta posea la característica de transitividad, pues Lewis considera que la causación efectivamente manifiesta esta característica. Antes de enunciar la reformulación de dependencia causal ofrecida por Lewis, mencionaremos un breve ejemplo donde podemos encontrar casos de causación sin dependencia contrafáctica. Estos ejemplos se conocen como casos de *anticipación temprana* (*early preemption*).

Volvamos a la situación de Ana golpeando un balón y quebrando un vaso de cristal con su golpe. Supongamos ahora que Ana se encuentra acompañada de su amiga Berta. Supongamos, además, que Ana golpea el balón en dirección al vaso de cristal, y un poco después (muy poco) Berta también golpea su balón en la misma dirección que Ana. Sin embargo, Berta golpeará su balón solo en el caso de que Ana no lo haga. Vamos a suponer, adicionalmente, que ambas tienen muy buena puntería y que el golpe de Berta es ligeramente más potente que el de Ana, de manera que el balón golpeado por Berta llegará exactamente en el mismo instante en que el balón de Ana lo hubiera hecho. De esta forma, si el vaso de cristal no es quebrado por el lanzamiento de Ana será, de seguro, quebrado por

el lanzamiento de Berta. Ahora bien, no hay mayores dudas de que el lanzamiento de Ana es la causa de la rotura del vaso; sin embargo, parece no ser verdadero el enunciado contrafáctico “si Ana no hubiera golpeado su balón, el vaso no se habría quebrado”. En efecto, el enunciado contrafactual anterior es falso, pues si Ana no hubiera golpeado su balón, lo habría hecho Berta con su lanzamiento. Por lo tanto, en esta situación, tenemos un caso de causación sin dependencia contrafáctica. El rótulo de *anticipación temprana* para este tipo de casos se debe a que Berta fue *anticipada* por Ana.

Ahora entra en escena el quinto y último paso de Lewis en la formulación de su teoría contrafáctica de la causación. En primer lugar, de manera de prevenir los contraejemplos del tipo “anticipación temprana”, Lewis reformula la noción de dependencia causal de modo que esta exhiba la característica de transitividad entendida en el sentido usual. Lewis afirma, “sea *c, d, e* ... una secuencia finita de eventos particulares actuales, tales que *d* depende causalmente de *c*, *e* de *d*, y así sucesivamente. Entonces esta secuencia es una cadena causal” (1973b, p. 563). En otras palabras, Lewis extiende la noción de dependencia causal de modo de convertirla en una noción transitiva, y denomina esta secuencia de eventos que dependen causalmente *cadena causal*. Finalmente, está todo listo para enunciar la definición final que Lewis busca. Haciendo uso del concepto de cadena causal recién introducido, el autor afirma que “...un evento es una *causa* de otro si y solo si existe una cadena causal que va desde el primer al segundo evento” (1973b, p. 563). Bajo el análisis contrafáctico de Lewis, la dependencia causal (en su versión transitiva) es suficiente y necesaria para la causación, por lo tanto su análisis debería evitar exitosamente los casos de anticipación temprana. Y, en efecto, lo hace. Sin mayores dificultades es posible mostrar que tanto el contrafáctico (i) como el (ii) son verdaderos en la situación de Ana y Berta presentado más arriba.

Para completar el análisis contrafáctico de la causación de Lewis es importante mencionar algunas cuestiones adicionales a lo dicho hasta ahora. En primer lugar, como ya ha sido mencionado más arriba, la teoría de Lewis es un análisis de la relación causal entre eventos, es decir, los *relata* de la relación causal son eventos entendidos en un sentido amplio; tal como menciona el autor, estos podrían ser relámpagos, batallas, conversaciones, impactos, etc. (1973b, p. 558). En segundo lugar, la teoría contrafáctica de Lewis es una teoría de la causación singular y no de generalizaciones causales. Es decir, el análisis de

Lewis aplica para afirmaciones del tipo “la bebida en exceso causó un malestar estomacal a Juan” y no para afirmaciones como “beber en exceso causa malestar estomacal”. En tercer lugar, Lewis afirma que los contrafácticos involucrados en su teoría deben ser interpretados de manera estándar. En particular, el autor quiere evitar las situaciones en que ciertos contrafácticos puedan ser entendidos como contrafácticos en retroceso (*backtracking counterfactual*) –llamados así porque el orden temporal del antecedente y el consecuente se encuentran invertidos– puesto que, en estos casos, su análisis puede calificar, erróneamente, algunas relaciones que son intuitivamente no causales como causales. Por último, también es importante destacar que la teoría contrafáctica de la causación elaborada por Lewis en su trabajo de 1973 no se muestra especialmente apta para casos de indeterminismo. Sin embargo, en trabajos posteriores (ver Lewis, 1986a) el autor realiza algunos ajustes a su teoría con la intención de que esta pueda explicar los casos de causación indeterminista.

Antes de finalizar esta sección, revisemos brevemente algunas consecuencias o dificultades de la teoría de la causación de Lewis. Un primer hecho que se deriva de la definición de causación en términos de cadenas causales es que la relación causal sería transitiva, es decir, si *c* causa *a* y *a* causa *e*, entonces *c* causa *e*. Este hecho se ajusta bien con muchas de nuestras intuiciones causales; sin embargo, parece haber ciertas situaciones en que nuestras intuiciones no responden a lo dictaminado por la teoría contrafáctica de Lewis. Por ejemplo, supongamos que una persona se enferma de un resfriado común (*c*) y como consecuencia de ello toma la medicación indicada para el resfriado que la afecta (*a*), esto último a su vez, causa la mejoría de la persona (*e*). Por transitividad deberíamos concluir que la causa de la mejoría de la persona fue haber enfermado, lo cual no parece ser correcto. Existe un amplio debate en torno al problema de la transitividad de las relaciones causales (ver Paul y Hall, 2013, para un análisis detallado); incluso el propio Lewis (2000) presenta varios de estos casos.

El segundo problema con la teoría contrafáctica de la causación de Lewis se conoce como *anticipación tardía (late preemption)*¹³. A diferencia de los casos de anticipación temprana que hemos analizado más arriba, los casos de anticipación tardía sobreviven a la solución de Lewis en términos de la noción de cadena causal. Como ejemplo de ello,

¹³ Ver Lewis (1986a) para varios casos de anticipación temprana y tardía.

volvamos al caso de Ana y Berta. Supongamos ahora que, a diferencia de la anticipación temprana, tanto Ana como Berta golpean sus respectivos balones al mismo tiempo, y dado que el lanzamiento de Ana es ligeramente más fuerte que el de Berta, llega primero al objetivo (el vaso de cristal), quebrando el vaso. ¿Es verdadero el último contrafáctico de la cadena causal “si el balón de Ana no hubiera estado en el aire a punto de golpear el vaso este no se habría quebrado”? Pues no. Lo habría quebrado el lanzamiento de Berta. En conclusión, los casos de anticipación tardía siguen siendo una piedra en el zapato para la teoría contrafáctica de Lewis, al menos en su versión inicial. Existe una variante a la anticipación tardía que se conoce como *anticipación triunfadora* (*trumping preemption*) (ver Schaffer, 2000a). En este tipo de situaciones tenemos dos cadenas causales que llevan a la ocurrencia del efecto, pero solo una de ellas es la triunfadora¹⁴.

Los casos de anticipación tardía y anticipación triunfadora llevaron a Lewis (2000) a modificar sustancialmente su enfoque contrafáctico de la causación, desarrollando una teoría basada en la idea de influencia causal en lugar de dependencia causal, la cual lidia mejor con los casos problemáticos antes mencionados, pero que dista bastante del enfoque inicial de Lewis del año 1973.

Una vez presentadas las teorías de la causación en términos de condicionales contrafácticos, en la siguiente sección volveremos nuestra atención hacia los enfoques probabilistas de la causación, los cuales utilizan elementos de la teoría de probabilidades para caracterizar las relaciones causales.

1.4. Teorías probabilistas

En general, las teorías probabilistas de la causación se basan en la intuición general de que algún cambio en las causas hará una diferencia en sus efectos, y que esta diferencia puede ser capturada mediante la noción de dependencia probabilista utilizando la teoría de probabilidades. Más específicamente, este tipo de enfoques de la causación afirman que las causas cambian la probabilidad de sus efectos. Estas intuiciones acerca de la forma de capturar la esencia de las relaciones causales comenzaron a desarrollarse fuertemente en la

¹⁴ Ver Psillos (2002), Miguel (2006) y Beebe, Hitchcock y Menzies (2009, Cap. 8) para un completo análisis tanto de la teoría contrafáctica de Lewis como de los problemas que enfrenta.

segunda mitad de siglo XX, logrando gran influencia en su momento. Sin embargo, más recientemente, estos enfoques basados en dependencia probabilista han sido paulatinamente reemplazados por enfoques que se conocen con el nombre de “modelado causal”, los cuales desarrollan métodos de inferencia causal utilizando las nociones de probabilidad. El modelado causal se ha convertido en un campo enormemente fructífero en las más variadas disciplinas. En esta sección se realiza una breve presentación histórica de las ideas más importantes asociadas a los enfoques probabilistas de la causación.

Los enfoques probabilistas fueron desarrollados inicialmente por autores y autoras como Hans Reichenbach (1971 [1956]), Irving Good (1959, 1961a, 1961b), Patrick Suppes (1970), Nancy Cartwright (1979), Brian Skyrms (1980), Ellery Eells (1991), entre otros. Muy aproximadamente, la idea central detrás de estos análisis puede ser capturada del siguiente modo (Hitchcock, 2018)¹⁵:

(CP) C es una causa de E cuando se cumple que $P(E/C) > P(E)$

En la formulación de la causación probabilista (CP) anterior, $P(E/C)$ indica una probabilidad condicional. Es decir, la probabilidad de que ocurra E dado C . Por lo tanto, (CP) indica que estaremos habilitados para decir que C es una causa de E cuando la probabilidad de que ocurra E dado que C ya ocurrió es mayor que la probabilidad de que E ocurra sin que haya ocurrido C . En esta formulación se ve muy claramente el “eslogan” detrás de los enfoques probabilistas: “las causas cambian (aumentan) la probabilidad de sus efectos”. Por supuesto, (CP) es una versión muy general que intenta dar cuenta de la esencia de los enfoques probabilistas de la causación y, por lo tanto, inevitablemente no considera las distintas variantes o visiones específicas con las que cada autor o autora ha contribuido a este programa de la causación. Con el fin de, en alguna medida, profundizar en estas visiones específicas y así ampliar nuestra comprensión de los enfoques probabilistas de la causación, brindaremos a continuación un breve examen de los principales aportes a esta línea de pensamiento.

¹⁵ Tal como muestra Hitchcock (2018), una versión alternativa y equivalente a la presentada aquí es la siguiente: C es una causa de E cuando se cumple que $P(E/C) > P(E/\neg C)$.

Reichenbach (1971 [1956]) fue el primero en presentar una teoría probabilista de la causación. En su trabajo *The Direction of Time* del año 1956, el autor intenta explicar, utilizando nociones probabilistas, la dirección del tiempo en términos de la dirección de la causación. Para ello, el autor formula algunas condiciones mediante la teoría de probabilidades con el objetivo de dar cuenta de la eventual asimetría presente en las relaciones causales¹⁶. Con este objetivo en mente, el autor presenta el “principio de causa común” y el “principio de intermediación causal”. El primer principio corresponde a una condición que permite determinar cuándo dos efectos tienen una causa común, mientras que el segundo principio permite determinar cuándo una causa puede encontrarse entre dos efectos. Estas dos condiciones, en particular el principio de causa común, determinan una red causal ordenada que es utilizada para explicar la dirección de la causación y, por lo tanto, la dirección del tiempo.

Poco después, Good (1959, 1961a, 1961b) realiza un análisis de la causación en términos probabilistas interpretando las probabilidades como expresión de propensiones, esto es, un tipo de probabilidad física. Sin embargo, el análisis en términos de dependencia probabilista de mayor influencia fue el de Suppes (1970)¹⁷, que se basa en dos conceptos centrales: *causas prima facie* y *causas espurias*. Con estos dos conceptos en mano, Suppes sostiene que una causa genuina es una causa prima facie no espuria. Formalmente, las definiciones de Suppes son las siguientes¹⁸ (Suppes, 1970, Cap. 2):

*Definición 1*¹⁹. El evento $B_{t'}$ es una *causa prima facie* del evento A_t si y solo si
(i) $t' < t$, (ii) $P(B_{t'}) > 0$ y (iii) $P(A_t|B_{t'}) > P(A_t)$.

¹⁶ Es importante notar que el análisis de Reichenbach (1971 [1956]) corresponde a un análisis metafísico de las relaciones causales, es decir, el autor intenta explicar, en términos de probabilidades, qué es la causación en sí misma. Sin embargo, en Reichenbach (1959) también podemos encontrar algunas explicaciones para dar cuenta de la causación desde un punto de vista epistemológico (más detalles en Williamson, 2009).

¹⁷ Ver Salmon (1980) para una revisión crítica de las propuestas de Reichenbach, Good y Suppes presentadas aquí. Sobre causación probabilista, ver también Salmon (1984, Cap. 7).

¹⁸ En estas definiciones, Suppes interpreta un evento como un subconjunto del espacio muestral.

¹⁹ En esta definición y en las siguientes, el subíndice t en cada uno de los eventos A , B , C representa el instante en que ocurre ese evento.

Definición 2. Un evento $B_{t'}$ es una *causa espuria* en un primer sentido de A_t si y solo si $B_{t'}$ es una causa prima facie de A_t y existe un $t'' < t'$ y un evento $C_{t''}$, tales que (i) $P(B_{t'}C_{t''}) > 0$, (ii) $P(A_t|B_{t'}C_{t''}) = P(A_t|C_{t''})$ y (iii) $P(A_t|B_{t'}C_{t''}) \geq P(A_t|B_{t'})$.

Definición 3. Un evento $B_{t'}$ es una *causa espuria* de A_t en un segundo sentido si y solo si $B_{t'}$ es una causa prima facie de A_t y existe un $t'' < t'$ y una partición $\pi_{t''}$ tal que para todos los elementos $C_{t''}$ de $\pi_{t''}$ (i) $P(B_{t'}C_{t''}) > 0$, $P(A_t|B_{t'}C_{t''}) = P(A_t|C_{t''})$ y (iii) $P(A_t|B_{t'}C_{t''}) \geq P(A_t|C_{t''})$ ²⁰.

Como se mencionó más arriba, el autor concluye su análisis sosteniendo que una causa genuina (o directa, en términos de Suppes) es una causa prima facie que, a su vez, no cumple las condiciones para ser espuria²¹ (ver Suppes, 1970, p. 25).

Más tarde, Cartwright (1979) y Skyrms (1980) señalaron un importante problema asociado a los análisis probabilistas de la causación, conocido como “paradoja de Simpson”, que se encuentra relacionada con las inversiones de desigualdades probabilistas. Para intentar solucionar este problema, Cartwright (1979) incluyó en el análisis de la causación probabilista la noción de *contexto de fondo*. De este modo, diremos que un evento es una causa debido a que aumenta la probabilidad de su efecto en cada contexto de fondo. Esta condición se conoce como *unanimidad del contexto*. Skyrms (1980) relaja la condición anterior sosteniendo que es suficiente con exigir que la causa aumente la probabilidad de su efecto en *al menos un contexto de fondo* (disminuyéndolo en ningún otro). Por último, la idea de “unanimidad del contexto” también es criticada por John Dupré (1984), quien afirma que esta idea es susceptible de contraejemplos. Una respuesta a las críticas de Dupré se encuentra en Eells (1988)^{22,23}.

²⁰ El autor ofrece una tercera definición para el concepto de causa espuria para casos asintóticos (ver Suppes, 1970, p. 27) que hemos omitido aquí debido a su escasa relevancia respecto de las ideas que se presentan en esta sección.

²¹ El autor también define otros tipos de causas además de la causa directa, como la causa suplementaria, suficiente y negativa. Ver Suppes (1970, Cap. 2).

²² Eells (1991) también ofrece un enfoque probabilista de la causación que prescinde de la noción de contexto de fondo.

Como es de esperar, los enfoques de la causación que basan sus análisis en la dependencia probabilista han sido objeto de variadas críticas. Algunas de ellas afirman que estos admitirían ciertos contraejemplos (Williamson, 2005, 2009). Otras sostienen que este tipo de enfoques de la causación no se adaptan adecuadamente a la explicación de ciertos mecanismos físicos cruciales para la causación (Williamson, 2009). Finalmente, otras críticas argumentan que este tipo de teorías de la causación fallan al intentar dar cuenta de casos de causación indeterminista (Dowe, 2000).

Ahora bien, como se ha mencionado al inicio de esta sección, los enfoques probabilistas de la causación, basados en la idea de dependencia probabilista, han sido progresivamente reemplazados por los enfoques de modelado causal. A partir de la década de 1980 comenzaron a desarrollarse diversas técnicas probabilistas-matemáticas con el objetivo de inferir relaciones causales a partir de ciertas variables aleatorias consideradas en el análisis. Estas técnicas se conoce con el nombre de *métodos de inferencia causal*. Los trabajos pioneros en este campo se deben a Judea Pearl (1988, 2000), Richard Neapolitan (1990), y Peter Spirtes, Clark Glymour y Richard Scheines (1993). Muchas de las investigaciones actuales en el campo de la causación probabilista se deben a este tipo de métodos de inferencia.

1.5. Teorías de agencia e intervención

Detrás de los enfoques de modelado causal recién mencionados se encuentra una idea inspiradora para un nuevo enfoque filosófico de la causación: la noción de *manipulación*. En efecto, los métodos de inferencia causal se basan en el hecho de que ciertas manipulaciones en una variable deberían producir algún cambio en alguna otra variable que se encuentre relacionada causalmente con la primera. Las teorías de agencia e intervención basan su análisis en esta idea, defendiendo que las relaciones causales son relaciones que

²³ Hasta aquí hemos hecho un breve repaso por las principales contribuciones de diversos filósofos y filósofas a los análisis probabilistas de la causación. Sin embargo, todos estos análisis también presentan algunas características comunes. En primer lugar, debemos mencionar que, en general, los enfoques de la causación en términos de dependencia probabilista que hemos presentado están pensados para capturar la noción de causación general y no para relaciones causales singulares. En segundo lugar, se debe notar que en los enfoques de la causación probabilista, los *relata* causales son eventos (o variables aleatorias) o un conjunto de ellos, muchas veces interpretados como un subconjunto del espacio muestral.

pueden ser explotadas con fines de manipulación y control (Woodward, 2003). Más específicamente, quienes defienden este grupo de teorías afirman que si C es una causa de E , entonces si C es manipulada de manera apropiada, tendremos como resultado un cambio en E (conversamente, si hubiera un cambio en E producto de la manipulación apropiada de C , entonces C causa E) (Woodward, 2009). Las teorías de agencia e intervención comenzaron a desarrollarse fundamentalmente a partir de la segunda mitad del siglo XX con los trabajos de Robin Collingwood (1940), Douglas Gasking (1955), Georg von Wright (1971), Peter Menzies y Huw Price (1993), James Woodward y Christopher Hitchcock (2003), James Woodward (2003), entre otros.

De acuerdo con Woodward (2009), los enfoques que basan sus análisis de la causación en la noción de manipulación pueden dividirse en dos grandes grupos. En primer lugar, se encuentran los *enfoques de agencia*, según los cuales la conexión causal se caracteriza, en particular, por agencia humana, es decir, por aquellas manipulaciones potencialmente realizables por un ser humano. En segundo lugar, se encuentran los *enfoques intervencionistas*, donde las relaciones causales se analizan utilizando la noción de “intervención”, que se define en términos causales y correlacionales, y no la noción de manipulación humana como en los enfoques basados en agencia. A continuación, examinaremos brevemente los enfoques de agencia de la causación para luego dar paso a los enfoques intervencionistas.

Los enfoques de agencia (y también los enfoques intervencionistas) tienen su origen en los trabajos clásicos sobre causación basada en manipulación. Un ejemplo de estos análisis clásicos²⁴ se encuentra en Gasking (1955), quien afirma “...uno dice A causa B en los casos donde uno podría producir un evento o estado del tipo A como un medio para producir uno del tipo B . Es decir, he explicado la relación ‘causa-efecto’ en términos de la relación ‘producir por medio de’” (p. 485). Luego, el autor añade una explicación adicional para la relación “producir por medio de” que se encuentra asociada al concepto de manipulación. En efecto, para Gasking “producir por medio de” en la cita anterior significa

²⁴ Otros trabajos en la misma línea son los de Collingwood (1940) y von Wright (1971), por ejemplo.

“producir B produciendo A ” donde el término “producir” se encuentra asociado a la realización de una manipulación (1955, p. 486).

Otro intento por caracterizar las relaciones causales en términos de manipulación – esta vez desde un enfoque probabilista– se debe a Menzies y Price (1993). En efecto, al inicio de su trabajo los autores señalan: “en este artículo defendemos la visión de que las nociones ordinarias de causa y efecto tienen una conexión directa y esencial con nuestra habilidad para intervenir en el mundo como agentes” (p. 187). La idea nuclear detrás del análisis manipulabilista de las relaciones causales del tipo de Menzies y Price es que un evento A es causa de un evento B si es el caso que provocar la ocurrencia de A es un medio a través del cual un *agente libre* provoca la ocurrencia de B (1993, p. 187). De acuerdo con los autores, la caracterización anterior evoca un tipo de relación que puede ser interpretada a través de lo que ellos denominan “probabilidades del agente”. Estas probabilidades deben considerarse como probabilidades condicionales de acuerdo con la teoría de probabilidades; por lo tanto, la probabilidad del agente que debería ser atribuida a B condicional a A , esto es $P(B/A)$, es la probabilidad que B tendría si uno eligiera realizar A (Menzies y Price, 1993, p. 190).

Los enfoques manipulabilistas mencionados anteriormente presentan, por supuesto, una serie de críticas y problemas. En primer lugar, en relación a la propuesta de Menzies y Price, parece no haber una relación directa entre “acción de un agente libre” y “manipulación” (ver Woodward, 2009). En efecto, no toda acción libre implica un proceso de manipulación y, por otro lado, una acción que no sea libre puede calificar como un tipo apropiado de manipulación. En segundo lugar, existe el problema que Woodward (2009) denomina el “problema de la proyección”. Desde una teoría manipulabilista, ¿cómo entendemos la causa del movimiento de una hoja que se desprende de un árbol producto de la acción natural del viento si no hay una acción de un agente libre? En tercer lugar, se presenta el problema con aquellas situaciones en que no hay posibilidad técnica de manipulación humana. Por ejemplo, ¿qué rol juega una teoría manipulabilista en las relaciones causa-efecto producto de la acción gravitacional? En cuarto lugar, acecha el problema de la circularidad: algunas y algunos autores sostienen que las nociones de “producir” o “provocar”, involucradas en las caracterizaciones de la relación causal recién mencionadas, ya tienen un contenido causal y, por lo tanto, se incurriría en un problema de

circularidad si las relaciones causales pretenden analizarse en estos términos. Por último, las teorías manipulabilistas de la causación han sido criticadas debido a que convertirían a la relación causal en un fenómeno indeseablemente antropocéntrico.²⁵

Los enfoques intervencionistas de la causación surgen con el propósito de solucionar las dificultades señaladas en el párrafo anterior. Suele considerarse que algunas de ellas han sido resueltas por este segundo enfoque manipulabilista; sin embargo, también es discutible si otras dificultades han sido efectivamente solucionadas por los enfoques intervencionistas. Independientemente del hecho anterior, los enfoques intervencionistas, debidos principalmente a Woodward (2003), han representado un claro avance hacia una caracterización más completa de la causación en términos de manipulación. Revisemos a continuación sus principales tesis.

En su libro *Making Things Happen: A Theory of Causal Explanation*, del año 2003, Woodward aborda el problema de la naturaleza de la causación en términos manipulabilistas, pero recurriendo a la noción novedosa de *intervención*²⁶. La estrategia general del autor consiste en establecer condiciones necesarias y suficientes para caracterizar las relaciones causales, haciendo uso de tres tipos de causas identificadas en su análisis: *causa directa*, *causal total* y *causa contribuyente*. En los siguientes párrafos seguiremos la estrategia argumentativa utilizada por el autor en su libro con el propósito de no perder claridad en la exposición de las ideas principales de su enfoque.

El primer paso de Woodward consiste en proponer una condición suficiente (CS) y una condición necesaria (CN) para la causación²⁷:

- (CS) Si (i) hay una intervención posible que cambia el valor de X tal que (ii) llevando a cabo esta intervención (y no otras intervenciones) cambiará el valor de Y , o la probabilidad de distribución de Y , entonces X causa Y (2003, p. 45).

²⁵ Ver Menzies y Price (1993) y Woodward (2009, Sección 4) para más detalles acerca de los problemas a los que se enfrentan las teorías manipulabilistas.

²⁶ Analizaremos, en primer lugar, la propuesta general de Woodward acerca de la causación, para luego pasar a especificar la noción de “intervención” utilizada por el autor en su teoría.

²⁷ En estas definiciones, los eventos X e Y son interpretados como variables, por lo que estas pueden tomar diversos valores o estados posibles.

(CN) Si X causa Y entonces (i) hay una posible intervención que cambia el valor de X tal que (ii) si esta intervención (y no otras intervenciones) fuera llevada a cabo, el valor de Y (o la probabilidad de algún valor de Y) cambiaría (2003, p. 45).

Al examinar las condiciones (CS) y (CN), el autor reconoce que (CS) funciona adecuadamente para una serie de situaciones causales clásicas, como la “doble prevención” (ver Hall, 2000) o la “causación por desconexión” (ver Schaffer, 2000b). Sin embargo, (CN) no tiene la misma suerte: es posible presentar ciertos contraejemplos que indican que es necesario reformular esta última condición (Woodward, 2003, 2009). Por lo tanto, se hace necesario extender el análisis realizado hasta aquí en términos de (CS) y (CN). Esto lleva al autor a distinguir entre dos tipos de causas: causa total y causa contribuyente.

Woodward (2003) define *causal total* del siguiente modo: “ X es una causa total de Y si y solo si tiene un efecto total no nulo en Y ; es decir, si y solo si hay alguna intervención en X solamente (y ninguna otra variable) tal que para algún valor de otra variable además de X , esta intervención en X cambiará el valor en Y ” (p. 50)²⁸. Por otro lado, el autor afirma que la noción de *causa contribuyente* “...intenta capturar la idea intuitiva de que X influye en Y a lo largo de alguna ruta incluso si, debido a una cancelación, X no tiene un efecto total en Y ” (p. 50). De este modo, una causa directa será siempre una causa contribuyente; sin embargo, puede haber causas contribuyentes que no sean causas directas²⁹ (Woodward, 2003, p. 50). Dadas las definiciones de causa total y causa contribuyente, es posible mostrar que las condiciones (CS) y (CN) son plausibles únicamente si la palabra “causa” en estas condiciones es entendida en el sentido de causa total, pero si interpretamos la noción de “causa” como causa contribuyente, entonces (CN) resulta ser falsa (no así (CS), que seguiría siendo válida)³⁰.

²⁸ De acuerdo con Woodward: “el efecto total de un cambio Δx en X sobre Y es entonces el cambio en el valor de Y o en la probabilidad de distribución de Y que resultaría de una intervención en X solamente que la modifica en la cantidad Δx , dados los valores de otras variables (2003, p. 50).

²⁹ Diremos que X es una *causa directa* de Y si y solo si la influencia de X sobre Y no es mediada por alguna otra variable, es decir, considerando todas las demás variables fijas (Woodward, 2003, p. 42).

³⁰ Ver Woodward (2003, pp. 50–51), para ejemplos ilustrativos acerca de esta última declaración.

Usando las definiciones presentadas anteriormente, Woodward arriba a la formulación final de su teoría manipulabilista (TM), que consiste en condiciones necesarias y suficientes tanto para cuando “causa” es interpretado como causa directa como para cuando “causa” es entendida como causa contribuyente:

(TM) Una condición necesaria y suficiente para que X sea una *causa directa* de Y respecto de un conjunto de variables V es que haya una posible intervención en X que cambiará Y o la distribución de probabilidad de Y cuando se mantienen fijas en algún valor todas las otras variables Z_i en V . Una condición necesaria y suficiente para que X sea una *causa contribuyente* de Y respecto de un conjunto de variables V es que (i) haya un camino dirigido de X a Y tal que cada conexión en este camino es una relación causal directa; esto es, un conjunto de variables $Z_1 \dots Z_n$ tales que X es una causa directa de Z_1 , la cual es a su vez una causa directa de Z_2 , la cual es una causa directa de $\dots Z_n$, la cual es una causa directa de Y , y que (ii) haya alguna intervención en X que cambiará Y cuando todas las otras variables en V que no están en este camino están fijas en algún valor. Si hay solo un camino P de X a Y o si el único camino alternativo de X a Y además de P no contiene variables intermedias (es decir, directas), entonces X es una causa contribuyente de Y siempre que haya alguna intervención en X que cambiará el valor de Y , para algunos valores de otras variables en Y (2003, p. 59, énfasis en el original).

(TM), a través de sus condiciones necesarias y suficientes, captura la esencia de los enfoques manipulabilistas en términos de intervenciones. Parafraseando a Woodward, estos enfoques quedan bien representados por el eslogan “no hay diferencia en las relaciones causales sin una diferencia en las relaciones de manipulabilidad, y no hay diferencia en las relaciones de manipulabilidad sin una diferencia en las relaciones causales” (2003, p. 61). Ahora bien, para que el análisis de la causación en términos de intervenciones sea completo es necesario precisar qué se entiende en las definiciones anteriores por “intervención”, algo que habíamos dejado pendiente al inicio de esta sección.

La noción de intervención usada en el análisis de la causación anterior se debe a Woodward y Hitchcock (2003) y Woodward (2003)³¹. De acuerdo a esta caracterización, *I* es una intervención para *X* respecto de *Y* si y solo si esta reúne las siguientes condiciones (Woodward, 2003, p. 98)³²:

- (C1) *I* causa *X*
- (C2) *I* actúa como un interruptor para todas las otras variables que causan *X*. Es decir, cuando *I* alcanza un cierto valor, *X* deja de depender de otras variables.
- (C3) Cualquier camino dirigido de *I* a *Y* pasa por *X*.
- (C4) *I* es estadísticamente independiente de cualquier variable *Z* que cause *Y*.

La teoría manipulabilista (o intervencionista) de Woodward representa un esfuerzo no menor por intentar capturar aquello que es distintivo de las relaciones causales genuinas, apoyándose en la noción de intervención definida anteriormente. Sin embargo, como es de esperar, esta no ha sido inmune a las críticas. Principalmente, algunos problemas relacionados con el antropomorfismo y la circularidad, mencionados brevemente más arriba, siguen acechando a este nuevo enfoque acerca de la naturaleza de la causación³³.

1.6. Comentarios finales

En este capítulo hemos examinado las principales teorías filosóficas contemporáneas de la causación calificadas bajo el rótulo de “marcadoras de una diferencia”, o teorías “difference-making”. Este examen incluyó las teorías regularistas, contrafácticas, probabilistas y de agencia e intervención. Algunas observaciones importantes para destacar de la presentación ofrecida en las secciones anteriores son las siguientes. En primer lugar,

³¹ Otros autores que han intentado caracterizar la noción de intervención son Hausman (1998), Spirtes et.al. (1993) y Pearl (2000), por nombrar a algunos.

³² Ver Woodward (2003, Cap. 3) para más detalles acerca de la noción de intervención.

³³ Ver Woodward (2016) para una breve revisión de las críticas a las teorías intervencionistas y Woodward (2009) para una revisión más o menos detalla del problema de la circularidad.

debemos mencionar que las teorías examinadas en este capítulo pueden ser vistas como análisis de tipo reduccionista acerca de la causación, pues en sus caracterizaciones o definiciones acerca de aquello que cuenta como una relación causal se usan términos o conceptos no causales, ya sea mediante la idea de instanciación de una regularidad, o a través de dependencia contrafáctica, o haciendo uso de conceptos como “probabilidad”, “agencia” o “intervención” (aunque Woodward, con su teoría intervencionista, no estaría de acuerdo con esta conclusión). En segundo lugar, es notorio el hecho de que algunos de estos enfoques contemporáneos de la causación toman prestados elementos de otras disciplinas diferentes a la filosofía para sus análisis; un claro ejemplo de ello son las teorías probabilistas y los enfoques de modelado causal. Por último, un tercer elemento a destacar de este tipo de teorías es que, en general, tratan de ajustarse a los avances y descubrimientos de otras disciplinas científicas, en particular, de las ciencias empíricas. Esto es claro, por ejemplo, en las teorías de agencia e intervención, que parecen pretender adaptarse a fenómenos propios de la biología, la economía o las ciencias sociales, entre otras disciplinas; o en las teorías probabilistas, que surgen como una manera de explicar causalmente ciertos fenómenos indeterministas propios de la física.

Capítulo 2. *Teorías de la causación física*

Resumen

Como contraparte de las teorías *difference-making* analizadas en el capítulo anterior, en este capítulo nos abocaremos al estudio de las teorías de la causación física.

En general, el estudio de las teorías de la causación física puede dividirse en dos vertientes: las teorías de transferencia, por un lado, y las teorías de proceso, por otro lado. Siguiendo este esquema, el presente capítulo se inicia con un análisis de las teorías de transferencia; en particular, se presentan las propuestas de Jerrold Aronson de 1971 y David Fair de 1979. Estas teorías son consideradas las iniciadoras de los programas reductivistas *a posteriori* de la relación causal. Veremos que, tanto en Aronson como en Fair, hay un intento por reducir las relaciones causales a la transferencia de cantidades físicas conservadas. Estos enfoques de las relaciones causales surgen como una reacción a la idea, generalizada en aquel momento, de eliminar cualquier afirmación causal de las ciencias empíricas, en particular, de la física.

A continuación, presentamos las teorías de proceso de la causación, que si bien recogen las propuestas de Aronson y Fair respecto de la transferencia de cantidades físicas, lo hacen en términos de una ontología de procesos. Las teorías de la causación fundamentadas en procesos físicos fueron desarrollándose sobre la base de los aportes de diferentes filósofos y filósofas, pero su progreso se debe principalmente a las contribuciones de Wesley Salmon y Phil Dowe a través de un fecundo debate entre ambos que tuvo lugar durante la década de 1990, y que finalizó con la publicación de la obra de Dowe del año 2000, *Physical Causation*, que contiene la versión más acabada de las teorías de proceso de la causación, la Teoría de las Cantidades Conservadas (TCC). En la presentación de las teorías de proceso de la causación seguiremos el desarrollo del histórico debate entre Salmon y Dowe. La versión de la teoría de Dowe del año 2000 será objeto de estudio del siguiente capítulo.

2.1. Teoría de la transferencia de Aronson

La primera teoría de la causación física, la teoría de Aronson, surgió como una reacción a la postura, fundamentalmente promovida por Russell, que pretendía eliminar la causación del escenario científico. A inicios de la segunda mitad del siglo XX, tal vez los análisis de la causación más influyentes eran los enfoques manipulabilistas, uno de cuyos principales exponentes fue Gasking (1955), quien afirmaba, en términos muy generales, que las relaciones causales son aquellas que son susceptibles de ser manipuladas por un agente racional (recuérdese el Capítulo 1, Sección 1.5, sobre teorías de la causación en términos de agencia e intervención). De acuerdo con Aronson (1971a), el hecho de que la causación sea reducida a la manipulación humana juega a favor del dictamen Russelliano. Como una reacción a este enfoque, en su artículo “On the Grammar of ‘Cause’”, Aronson intenta justificar la existencia de causación en las ciencias teóricas (en particular, en la física) suministrando condiciones no antropomórficas para caracterizar las relaciones causales³⁴.

En acuerdo con el punto de vista de Russell acerca de la (no) relación entre la causación y la física teórica, algunos filósofos manipulabilistas argumentan que, dado que no existe un uso legítimo del término “causa” que se encuentre libre de antropomorfismo, y dado que la tarea de la ciencia teórica no es manipular la naturaleza, la causación debería ser eliminada de la ciencia y, en particular, de la física. En este contexto, Aronson se plantea el desafío de encontrar una base para la asimetría causa-efecto (es decir, para distinguir aquello que llamamos “causa” de aquello que llamamos “efecto”), sin tener que apelar a la noción de manipulación humana. Detrás de esta búsqueda se esconde una intención aún más profunda por parte de Aronson: fundamentar las afirmaciones causales en contextos científicos. Quizás es este hecho el que motiva al autor a buscar esta fundamentación en los elementos que nos provee la física, como mostraremos a continuación.

El punto de partida del análisis de la causación de Aronson corresponde a un examen gramatical de la palabra “causa”. En efecto, el autor afirma que esta palabra se encuentra asociada a verbos transitivos mecánicos como empujar, tirar, levantar o golpear. Este hecho indicaría, de acuerdo con Aronson, que la palabra “causa” es una palabra-dimensión; es

³⁴ Algunas ideas sobre su teoría de la causación también son presentadas en Aronson (1971b).

decir, “causa” serviría como un término más general para un grupo de términos del mismo tipo y que cumplen la misma función. El autor plantea, de hecho, que en una oración el verbo transitivo puede ser reemplazado por la palabra “causa” o viceversa (1971a, p. 417). Para justificar su afirmación, Aronson usa como ejemplo la oración “John golpeó el libro en el suelo”, afirmando que esta puede ser reemplazada por “John causó que el libro estuviera en el suelo”. En razón del análisis anterior, el autor afirma: “Si ‘*A* causa *B*’ es verdadero, entonces debería haber un verbo transitivo específico (por ejemplo, forzar, tirar, etc.) disponible para describir tal relación entre *A* y *B*” (1971a, p. 419).

Sobre la base del análisis gramatical anterior, Aronson presenta su enfoque de la causación mediante tres condiciones. El cumplimiento de cada una de estas condiciones indicaría una potencial relación causal entre los eventos analizados. La primera condición se encuentra relacionada a la distinción entre “cambios naturales” y “cambios no naturales” que introduce el autor:

- (1) En “*A* causa *B*”, *B* designa un cambio en un objeto, el cual es un cambio no natural (1971a, p. 421).

De acuerdo con Aronson, en la ciencia teórica nos encontramos con dos tipos de cambios: aquellos que ocurren en un objeto de manera independiente de otros objetos y aquellos que ocurren producto de la interacción de dos o más objetos. Los primeros son llamados por Aronson “cambios naturales”, y un ejemplo de este tipo de cambio ocurre en aquellos objetos que se encuentran sometidos a la ley de inercia, pues estos permanecerán en reposo o seguirán moviéndose con velocidad constante sin la necesidad de que existan fuerzas externas que hagan este trabajo. Los cambios del segundo tipo son llamados por Aronson “cambios no naturales”. Para el autor, estos últimos representan los genuinos cambios causales.

La segunda condición de Aronson es la siguiente:

- (2) En “*A* causa *B*”, en el momento en que ocurre *B*, el objeto que causa *B* está conectado con el objeto que experimenta el cambio (1971a, p. 422).

Con esta segunda condición, Aronson se asegura que la causa y el efecto no estén separados espacialmente en el momento en que se produce el cambio, evitando los

problemas que presenta la acción a distancia. Muy probablemente al tanto de los desarrollos de la mecánica cuántica y sus desavenencias con el problema de la acción a distancia, Aronson no postula esta condición como un requisito estricto, sino más bien como una condición que reflejaría los desarrollos históricos propios de la mecánica.

Por último, la tercera condición que postula Aronson puede ser enunciada como sigue:

- (3) Antes del momento de la ocurrencia de *B*, el cuerpo que hace contacto con el objeto efecto posee una cantidad (por ejemplo, velocidad, impulso, energía cinética, calor, etc.) que se transfiere al objeto efecto (cuando se establece el contacto) y se manifiesta como *B* (1971a, p. 422).

Esta es la condición más relevante de las tres mencionadas, pues a la luz de esta condición queda de manifiesto, de acuerdo con el autor, que aquello que subyace el uso de los verbos transitivos mecánicos (forzar, empujar, tirar, etc.) es un proceso de transferencia de una cantidad física entre el objeto que denominamos “causa” y el objeto que denominamos “efecto”. Es decir, el análisis gramatical de la noción de “causa” llevado a cabo por Aronson deja entrever que el sustrato común a los verbos transitivos es que el uso de ellos en mecánica implica una transferencia de cantidades físicas entre la causa y el efecto. Además de lo anterior, la condición (3) permite determinar una dirección para la relación causal. Es decir, con esta última condición Aronson ha encontrado la base que buscaba para determinar la asimetría de la relación causa-efecto. Ahora es claro que, de acuerdo con el autor, el objeto que transfiere una cantidad física es el objeto asociado con la causa y el objeto que recibe esta cantidad es el objeto asociado con el efecto.

Por ejemplo, podemos sustituir “Juan volcó el tintero” por “Juan causó que el tintero se volcara” y explicar esta situación diciendo que la mano de Juan transfirió un movimiento rotacional al tintero, el que terminó por volcarse (Aronson, 1971a, p. 423). En la situación “un jugador de fútbol pateó el balón” también podemos decir “un jugador de fútbol causó que el balón se moviera”, y en este caso afirmamos que el jugador de fútbol transfirió movimiento al balón, el que finalmente se movió. Respecto de este punto es importante aclarar que, técnicamente, lo que se transfiere es momentum y/o energía, y no movimiento. Aronson se encuentra al tanto de esto, pero afirma que usar el término “movimiento” en

lugar de los conceptos técnicos de energía y/o momentum no altera el espíritu explicativo de sus tres condiciones.

Antes de finalizar nuestra exposición acerca de la teoría de la transferencia de Aronson, es importante resaltar un aspecto de esta teoría que será relevante para nuestras discusiones posteriores. Este aspecto se relaciona con las leyes de conservación. Si bien los posteriores desarrollos de Fair, Salmon y Dowe ubican a las leyes de conservación en un lugar mucho más central dentro de sus análisis en comparación con la teoría de Aronson, este es un aspecto que ya se insinúa en el trabajo de este último autor. En efecto, Aronson afirma que existe una relación muy clara entre la causación y las leyes (o principios) de conservación en razón de que, si una cantidad existe, debe haber alguna fuente para esa cantidad. Si decimos, por ejemplo, que existe una transferencia de momentum de un objeto *A* a un objeto *B*, es natural preguntarse por el origen de ese cambio y, de acuerdo con Aronson, preguntarse por el origen de ese cambio es otra manera de decir que las cantidades (por ejemplo, momentum y energía) dentro de un sistema se conservan (1971a, p. 426).

2.2. Teoría de la transferencia de Fair

Mientras que los trabajos de Aronson (1971a, 1971b) fueron el puntapié inicial de los enfoques fisicalistas de la causación, el trabajo de Fair de 1979 vino a consolidar esta línea de investigación. Siguiendo la idea básica detrás de la teoría de Aronson, en su artículo titulado “Causation and the Flow of Energy” Fair sostiene que la física puede iluminarnos acerca de la naturaleza de las relaciones causales y que, de hecho, tales relaciones se encontrarían asociadas con un flujo de energía-momentum de la causa al efecto (p. 220).

La propuesta de Fair pretende posicionarse tanto en contra de la visión regularista como en contra de la visión contrafáctica de la causación. De acuerdo con el autor, la identidad entre “causación” y “flujo de energía-momentum” es una relación hipotética que se ha descubierto empíricamente mediante las herramientas que nos provee la física (1979, p. 231). Las relaciones causales caracterizadas de este modo son, por lo tanto, relaciones objetivas entre eventos del mundo físico (1979, p. 224). Como veremos en el siguiente capítulo, esta son las mismas ideas que subyacen a la TCC de Dowe.

En particular, Fair afirma que, en una relación causal, se produce una transferencia de cantidades físicas, específicamente de energía y/o momentum de la causa al efecto, las cuales son cantidades conservadas a través del tiempo (1979, p. 228). En términos técnicos, la física nos dice que el comportamiento de estas cantidades puede ser especificado a través de las derivadas respecto a la posición y el momentum de la energía total de un sistema cerrado, en el marco de la función Hamiltoniana (1979, p. 228).

Fair ofrece algunos ejemplos con el objetivo de mostrar las virtudes de su análisis causal. La situación en la que afirmamos que el lanzamiento de una pelota fue la causa de la rotura de una ventana puede ser explicada, de acuerdo con la hipótesis de Fair, afirmando que la energía cinética de los trozos de vidrio provenientes de la ventana tiene su fuente en la energía cinética que poseía la pelota antes del impacto en ella. La energía cinética de un automóvil tiene su origen en la energía química de la combustión en los pistones. Por último, cuando en un juego de billar decimos que la bola *A* en movimiento ha causado el posterior movimiento de la bola *B* inicialmente en reposo, lo que en realidad queremos decir, de acuerdo con el análisis de Fair, es que la bola *A* ha transferido una parte o la totalidad de su energía y su momentum a la bola *B*, la que luego de la colisión comienza a moverse como consecuencia de la ganancia de estas cantidades transferidas.

Fair también hace frente a algunos aspectos ontológicos de su análisis reduccionista de la causación. El autor afirma que, en general, las relaciones causales toman como argumentos (como causa y efecto) objetos muy variados, como eventos, objetos físicos, propiedades, acciones, estados mentales, etc. Para ganar en simplicidad, Fair propone llamar “*A*-objetos” y “*B*-objetos”, para la causa y el efecto respectivamente, a esta plétora de objetos presentes en nuestro lenguaje causal. Sin embargo, el autor reconoce que esta reducción ontológica es una de las debilidades de su propuesta, siendo una de las razones por las que Fair afirma que su análisis debe ser pensado no como una teoría robusta, sino como un programa filosófico (1979, p. 234).

La debilidad de este aspecto de la propuesta de Fair se puede comprobar, en primer lugar, en que para establecer las condiciones físicas de la transferencia de energía-momentum, es necesario disponer de una descripción física de las propiedades de la energía-momentum de los *A*- y *B*-objetos (1979, p. 233). En segundo lugar, existe el

problema de la reducción de las afirmaciones causales del lenguaje ordinario al lenguaje de la física. En general, en el lenguaje ordinario usamos expresiones más generales que las usadas en el lenguaje físico para indicar ciertas características o propiedades de los objetos. Así por ejemplo, Fair sostiene que no tendríamos problemas en afirmar que una primera bola fue la causa del movimiento de una segunda bola en un juego de billar, pero no estaríamos dispuestos a aceptar que el movimiento de la primera bola fue la causante de la forma o color de la segunda bola. Este hecho requiere, de acuerdo al autor, una teoría de reducción de propiedades (1979, p. 234).

Un segundo aspecto que es abordado por Fair en su análisis reductivista de la causación, tiene relación con la identidad a través del tiempo de las cantidades físicas transferidas³⁵. El autor argumenta que la identificación de la energía y/o momentum a través del tiempo puede caracterizarse adecuadamente haciendo uso del concepto de “sistema cerrado”³⁶. Recordemos que un sistema físico se encuentra cerrado en un intervalo de tiempo si no hay flujo, en este caso de energía y/o momentum, a través de la superficie de este sistema en ese intervalo temporal. En el contexto de un sistema cerrado, serán las leyes de conservación las que nos permitirán reconocer las cantidades que se conservan, logrando así su identificación a través del tiempo³⁷. A esto Fair agrega que la relación causal sería una relación transitiva al igual que la relación de transferencia de cantidades en la que se encuentra fundamentada. El autor señala que, si consideramos las dos precisiones anteriores, podemos estar seguros de que la energía (momentum) transferida al objeto denominado “efecto” es la misma energía (momentum) poseída por el objeto denominado “causa” (1979, p. 235).

Estos han sido algunos de los aspectos más importantes del análisis de Fair de la causación. En la siguiente sección presentaremos algunas reflexiones generales respecto de las teorías de transferencia, tanto en la versión de Aronson como en la versión de Fair.

³⁵ Ver Capítulo 4, Sección 4.3, para una exposición detalla del problema de la identidad a través del tiempo de las cantidades conservadas en las teorías de la causación física.

³⁶ Ver Capítulo 4, Sección 4.6, para más detalles acerca del concepto de “sistema cerrado” y su relación con las teorías de la causación física.

³⁷ Como nota Fair, esta forma de identificar las cantidades conservadas implica que los objetos que poseen estas cantidades también puedan ser identificados a través del tiempo (1979, p. 234).

2.3. Algunos comentarios sobre las teorías de Aronson y Fair

Las propuestas de Aronson y Fair aportan una de las ideas centrales detrás de los análisis fisicalistas de la causación: la transferencia de cantidades físicas. Como señalamos en las secciones precedentes, la imagen que nos dejan estos análisis de la causación es que la naturaleza de las relaciones causales puede ser capturada por la transferencia de cantidades físicas como energía, momentum, calor, etc., desde la causa al efecto.

Los análisis de Aronson y Fair no solo comparten la noción de transferencia para caracterizar a las relaciones causales, sino que también comparten la motivación detrás del desarrollo de sus respectivas teorías. Esta motivación se basa en el propósito de recomponer el uso de la noción de causa en las ciencias modernas y, en particular, en la física. Ya en la exposición anterior acerca de la teoría de Aronson se mencionaron algunos pasajes de su artículo “On the Grammar of Cause”, donde intenta defender la tesis de que la causalidad sí tiene cabida en la ciencia moderna, en contraposición a la visión ampliamente extendida en ese momento, que pretendía eliminar la causación de la ciencia, defendida por filósofos como Russell (1913) o Willard Quine (1966). La misma postura de Aronson puede encontrarse en la propuesta de Fair, quien afirma que “uno de los principales objetivos de la ciencia física puede ser descrito como el descubrimiento de las causas de los fenómenos físicos” (1979, p. 220). A continuación, el autor agrega lo siguiente acerca del rol de la causación en la explicación de los fenómenos físicos: “al menos típicamente, tales explicaciones [las explicaciones causales] proceden especificando las *causas* del evento a ser explicado [...] es mi creencia que podemos tener un sentido más amplio de explicación en términos causales que en términos de leyes físicas” (1979, pp. 220-221, énfasis en el original). Estos pasajes del artículo de Fair, así como también los de Aronson presentados en la Sección 2.1, ponen de manifiesto, como hemos señalado, que existe una tesis transversal a ambas propuestas que se relaciona con defender las explicaciones causales en las ciencias físicas.

Otra característica compartida por los análisis reductivistas de Aronson y Fair tiene relación con el carácter supuestamente local de estas teorías de la causación. En efecto, desde el punto de vista de las teorías de la transferencia, si dos eventos se encuentran conectados causalmente dependerá exclusivamente de la cantidad transferida en la

interacción y no de regularidades o leyes que se encuentren fuera de la región local del espacio-tiempo donde ocurre esa interacción. Las teorías locales se oponen a las teorías globales de la causación en las que la relación causal sí depende de elementos (leyes, por ejemplo) que se encuentran fuera de la región espacio-temporal en que esta ocurre, como sucede, por ejemplo, con las teorías contrafácticas. La característica de localidad de la causación también es asumida por los posteriores desarrollos de la causación en términos de procesos, como los de Salmon y Dowe que veremos más adelante³⁸.

A pesar de los esfuerzos de Aronson y Fair por desarrollar enfoques de la causación lo más completos posible, de todos modos se vieron enfrentados a variadas dificultades. Algunas de estas críticas se relacionan con cómo explicar la causación como persistencia de un objeto, cómo abordar el problema de la identidad a través del tiempo de las cantidades físicas y cómo explicar la asimetría causal. En Dowe (2000, Cap. 3) se puede encontrar una breve presentación de cada una de estas críticas.

2.4. Teoría de procesos de Salmon (y Dowe)

En 1984, Salmon publica su libro *Scientific Explanation and the Causal Structure of the World*, donde presenta su teoría de procesos de la causación después de varios años de trabajo en torno al tema³⁹. El influyente libro de Salmon vino a revitalizar la discusión acerca de los problemas de la explicación y de la causación, recibiendo críticas de diverso tipo y suscitando nuevas investigaciones respecto de estos tópicos filosóficos. Uno de los críticos de la teoría de la causación de Salmon fue Dowe. En efecto, durante la década de 1990, Salmon y Dowe participaron de un fructífero debate mediante la publicación de reiterados trabajos académicos en los cuales argumentaban en favor de sus puntos de vista respecto de diversas cuestiones involucradas en las teorías de proceso de la causación. El resultado de este debate, por un lado, fue la propuesta por parte de Salmon de una segunda versión de su teoría y, por otro lado, el desarrollo de Dowe de su propia teoría de la causación, inspirada en la de Salmon. En la presentación de esta sección nos guiaremos en

³⁸ Ver Capítulo 4, Sección 4.7, para una crítica a la característica de localidad de las teorías de la causación física.

³⁹ Algunas ideas acerca de la causación también se pueden encontrar en Salmon (1998).

nuestro relato por el desarrollo histórico del debate Salmon-Dowe. La teoría de la causación de Dowe será abordada en detalle en el siguiente capítulo.

El objetivo general de Salmon con su teoría de la causación es proveer un enfoque de las relaciones causales que se muestre adecuado para una teoría causal de la explicación científica (1984, p. 135). La teoría de procesos de la causación de Salmon descansa, como muy bien resume Dowe, en cinco supuestos: (a) la causalidad es un aspecto objetivo del mundo, (b) la causalidad es un aspecto contingente del mundo, (c) una teoría de la causación debe ser consistente con la posibilidad del indeterminismo, (d) la teoría debe ser independiente del tiempo, de modo que sea coherente con una teoría causal del tiempo y (e) la teoría no debe violar las restricciones de Hume en relación a “poderes ocultos” (Dowe, 2000, p. 66).

Salmon presenta su teoría sobre la base de la explicación de dos conceptos que serían fundamentales para nuestra comprensión de la causación: la propagación (o transmisión) y la producción. Ambos conceptos, de acuerdo con Salmon, son muy familiares para el sentido común, pues forman parte de nuestras afirmaciones causales cotidianas (1984, p. 139).

En cuanto a la noción de propagación, el concepto de *proceso causal* juega un papel central en la propuesta de Salmon, ya que, para el autor, las entidades básicas son los procesos y no, por ejemplo, hechos o eventos. Para clarificar la noción de proceso, Salmon establece algunas comparaciones con estos últimos. Por un lado, afirma que una de las características sobresalientes de los procesos respecto de los eventos es que los primeros tienen una duración espacio-temporal mayor y una extensión espacial mayor que los eventos (1984, p. 139). Así, en un diagrama espacio-temporal, un evento aparece representado por un punto, mientras que un proceso aparece representado por una línea⁴⁰. De este modo, como ejemplifica el autor, un evento puede ser una pelota de béisbol golpeando una ventana y, en esta misma situación, un proceso sería la pelota viajando desde el bate a la ventana (1984, p. 139).

⁴⁰ Es importante mencionar que la noción de proceso causal de Salmon, como el mismo autor señala, se encuentra inspirada en el concepto de “línea causal” de Russell. Para detalles acerca del concepto de línea causal, ver Russell (1948).

En este punto se hace necesario distinguir entre procesos causales y pseudo-procesos. Para establecer esta distinción, Salmon afirma, en primer lugar, que los procesos (pseudo o no) presentan cierto grado de uniformidad, exhiben una estructura (1984, p. 144). De este modo, para el autor la diferencia entre los procesos causales y los pseudo-procesos se encuentra en el hecho de que los primeros pueden transmitir su propia estructura, mientras que los segundos no pueden hacerlo (1984, p. 144). Esta es una característica fundamental de los procesos causales, pues es la característica que le otorga el papel central que cumple esta noción en la teoría de Salmon. En efecto, al autor afirma que los procesos que transmiten su propia estructura son también capaces de transmitir marcas, señales, información, energía e influencia causal (1984, p. 146). Así, los procesos causales son los que se ocupan de propagar la influencia causal en nuestro mundo (1984, p. 146).

De acuerdo con Salmon, la distinción entre procesos causales y pseudo-procesos en términos de transmisión de estructura es adecuadamente capturado por el criterio de transmisión de marca de Reichenbach (ver Reichenbach, 1971 [1956]). Mediante la teoría “at-at” de la transmisión causal⁴¹, Salmon postula su criterio de transmisión de marca (TM) basado en las ideas de Reichenbach:

“(TM): Sea P un proceso que, en ausencia de interacciones con otros procesos, permanecería uniforme respecto de una característica Q , la cual se manifestaría consistentemente durante un intervalo que incluye a ambos puntos A y B ($A \neq B$) del espacio-tiempo. Entonces, una marca (que consiste en una modificación de Q en Q'), que ha sido introducida en el proceso P por medio de una única interacción local en el punto A , se transmite al punto B si P manifiesta la modificación Q' en B y en todas las etapas del proceso entre A y B sin intervenciones adicionales” (1984, p. 148).

Para ejemplificar el criterio (TM), pensemos en un haz de luz monocromático proveniente de un láser que apunta en una cierta dirección. Para Salmon este es un proceso

⁴¹ Salmon nota que la noción de transmisión de marca de Reichenbach podría involucrar ciertos compromisos con poderes ocultos o con la noción de necesidad, provocando problemas con algunos supuestos que él mismo asume para su teoría. En efecto, el criterio de Reichenbach afirma que los procesos se distinguen de los pseudo-procesos por su *habilidad* para transmitir una marca. La teoría “at-at” de la transmisión causal, de acuerdo con Salmon, viene a solucionar estos inconvenientes. Ver Salmon (1977) para más detalles.

causal, ya que, si disponemos de un filtro de color (marca) en medio de la trayectoria del haz de luz, a partir de este filtro el haz cambiará de color, es decir, es capaz de transmitir una marca. En cambio, el movimiento de ese mismo punto de luz sobre una pared representa un pseudo-proceso, pues en esa dirección del movimiento el punto de luz no es capaz de transmitir una marca.

Hemos mencionado que la teoría de Salmon se compone de dos elementos fundamentales: la propagación de la influencia causal, que como hemos visto se encuentra caracterizada a través de la noción de proceso causal, y la producción causal. Revisemos, brevemente, este segundo concepto. Salmon afirma que la idea de producción causal puede ser explicada en términos del concepto de “bifurcaciones causales” (*causal forks*)⁴², en particular de los conceptos de “bifurcación conjuntiva” (*conjunctive forks*) y “bifurcación interactiva” (*interactive forks*), ambos conceptos definidos en términos estadísticos. Por un lado, una bifurcación interactiva aparece cuando dos o más procesos físicamente independientes surgen de algún conjunto de condiciones de fondo (causa común) (Salmon, 1984, p. 179). Por otro lado, una bifurcación interactiva se encuentra representada, de acuerdo con Salmon, por la intersección espacio-temporal de procesos. Cuando dos procesos se intersectan y producto de esta intersección los procesos participantes experimentan modificaciones que se mantiene después de interactuar, decimos que estos procesos forman una *interacción causal*.

En resumen, la teoría de Salmon de 1984 pretende explicar algunos aspectos fundamentales respecto de las relaciones causales a través de tres nociones centrales en su teoría (p. 179). En primer lugar, los procesos causales son caracterizados como los medios a través de los cuales la estructura y el orden son transmitidos de una región a otra del espacio-tiempo. En segundo lugar, las interacciones causales son el medio por el cual las modificaciones en la estructura son producidas. Estas interacciones causales son caracterizadas a través de las bifurcaciones interactivas. Por último, las bifurcaciones conjuntivas son las encargadas de la producción de la estructura y el orden.

⁴² Las bifurcaciones causales tienen su origen en el trabajo de Reichenbach (1971[1956]), quien intenta explicar el principio de causa común en términos de este concepto.

Posteriormente, en el año 1992, Dowe publica una serie de artículos (1992a, 1992b, 1992c) donde presenta algunas críticas a la teoría de procesos de Salmon, a la vez que desarrolla, en una primera versión, su propia teoría de la causación. Dowe formula cuatro críticas a la teoría de procesos de Salmon. La primera de ellas tiene relación con un reclamo de circularidad⁴³. En efecto, para Dowe las nociones de “marca” e “interacción causal” utilizadas en la teoría de Salmon son mutuamente dependientes; por lo tanto, la teoría presentaría un problema de circularidad (Dowe, 1992a, p. 201). En segundo lugar, para Dowe la teoría de Salmon presenta un problema de vaguedad en dos sentidos: (i) por un lado, el análisis de las relaciones causales en términos de propagación y producción implica la noción primitiva de “característica” (relacionada con “estructura”), la cual deber ser especificada con mayor claridad, puntualizando al menos su rango de aplicación; (ii) por otro lado, se debe especificar qué tipo de interacciones son válidas en el marco de la teoría de Salmon, pues ningún proceso presenta una ausencia total de interacciones causales locales (1992a, pp. 201–202). Una tercera crítica a la teoría de Salmon tiene relación con el uso de relaciones estadísticas: ni la producción causal ni las interacciones causales, de acuerdo con Dowe, pueden ser analizadas en estos términos (1992a, p. 204). Por último, Dowe nota que la teoría de Salmon no ha evitado la alusión a “poderes ocultos”, hecho que se manifestaría en el uso de contrafácticos y propensiones en las caracterizaciones y definiciones dadas por Salmon (1992a, p. 207)⁴⁴.

Como hemos mencionado, en estos mismos trabajos de 1992 Dowe formula una primera versión de su teoría de procesos, la TCC. De acuerdo con lo que el autor señala, este nuevo enfoque pretende cumplir con los mismos objetivos generales presentados por Salmon, pero introduciendo el concepto de “cantidad conservada”. Dowe presenta su enfoque de procesos de la causación sobre la base de dos definiciones^{45,46} (1992a, p. 210; 1992c, p. 126):

⁴³ Como el mismo Dowe señala, esta crítica fue presentada con anterioridad en diferentes versiones por Kitcher (1985), Sober (1987) y Mellor (1988).

⁴⁴ Otras críticas a la teoría de Salmon, asociadas a problemas de inadecuación epistemológica, son presentadas en Dowe (1992c).

⁴⁵ El hecho de que la teoría de Dowe sea presentada mediante definiciones será un hecho relevante para los argumentos que presentaremos en el Capítulo 7.

Definición 1: una interacción causal es una intersección de líneas de mundo que involucran el intercambio de una cantidad conservada.

Definición 2: un proceso causal es una línea de mundo de un objeto que manifiesta una cantidad conservada.

Dos años más tarde, en 1994, Salmon se hace eco de las objeciones a su teoría de 1984 y presenta una nueva propuesta para analizar la causación basada en el enfoque de Dowe del año 1992 descrito más arriba. Básicamente, la nueva propuesta de Salmon reemplaza la noción de “cantidad conservada” usada por Dowe en su teoría por la noción de “cantidad invariante”. Esta modificación se justifica, de acuerdo con Salmon, por el hecho de que sería deseable que la causación se expresara de una manera que fuera independiente del marco de referencia (1994, pp. 305–306). En este sentido, sería razonable exigir, afirma Salmon, que si dos eventos se encuentran conectados causalmente respecto de un marco de referencia, estos eventos permanezcan conectados en todos los marcos de referencia. Y estas ideas justamente se sustentarían adecuadamente en la noción de “cantidad invariante”, pues estas cantidades, además de poseer la característica de permanecer constantes en el tiempo, se mantienen sin variación en todo marco de referencia. Esto no sucede con las cantidades conservadas, las cuales se mantienen constantes en el tiempo, pero su valor puede diferir de un marco de referencia a otro. De acuerdo con las consideraciones anteriores, Salmon presenta su nueva propuesta del siguiente modo (1994, pp. 303, 308):

Definición 1: una interacción causal es una intersección de líneas de mundo que involucra el intercambio de una cantidad invariante.

Definición 2: un proceso causal es una línea de mundo de un objeto que transmite una cantidad diferente de cero de una cantidad invariante en cada momento de su historia (cada punto del espacio-tiempo de su trayectoria).

Definición 3: un proceso transmite una cantidad invariante (o conservada) de A a B ($A \neq B$) si posee esta cantidad en A y en B y en cada etapa del proceso

⁴⁶ La versión final de la TCC de Dowe (2000) será presentada, con todo detalle, en el siguiente capítulo.

entre A y B sin ninguna interacción en el intervalo semi-abierto $(A, B]$ que involucre un intercambio de esa cantidad invariante (o conservada) particular.

El debate entre Salmon y Dowe respecto de la mejor manera de caracterizar las relaciones causales continúa a través del intercambio de un par de trabajos publicados en los años 1995 y 1997 (ver Dowe, 1995a; Salmon, 1997), donde los autores discuten algunas cuestiones particulares relacionadas principalmente con el problema de la dirección de la causación y las nociones de “transmisión” y “cantidad invariante” involucradas en las propuestas de ambos autores. Finalmente, en el año 2000, como resultado del debate con Salmon, Dowe publica su libro *Physical Causation*, en el cual el autor presenta la versión final de su TCC. Un análisis detallado de esta obra será objeto del siguiente capítulo.

2.5. Comentarios finales

Hemos presentado, a través de un recorrido histórico, las propuestas fisicalistas de la causación. Hemos visto que estos análisis comenzaron a gestarse en la década de 1970 con las propuestas de transferencia de cantidades de Aronson y Fair, las cuales sirvieron como insumo para las posteriores propuestas basadas en procesos de Salmon y Dowe. El sustrato común de estas caracterizaciones de las relaciones causales es que la transferencia/transmisión/intercambio de una cantidad física es lo que captura aquello que es distintivo de este tipo de relación metafísica.

Capítulo 3. *Teoría de las Cantidades Conservadas*

Resumen

La Teoría de las Cantidades Conservadas, a pesar de las diferentes críticas que ha recibido, continúa siendo la teoría de la causación física más robusta y desarrollada con la que contamos. El proyecto de Phil Dowe, su autor, es ofrecer una teoría de la causación que provenga de un análisis empírico y no de uno conceptual. En otras palabras, Dowe pretende establecer qué es la causación en el mundo, dejando de lado el problema acerca del concepto de causa en el sentido en que cotidianamente nos referimos a este término. Asimismo, el análisis causal de Dowe pretende ser científicamente informado, en particular, por la física. De este hecho se deriva, de acuerdo con el autor, el adjetivo de teoría de la causación *física*.

La TCC de Dowe (2000) se articula sobre la base de tres ejes. El primero de ellos busca responder a la pregunta, ¿qué son los procesos e interacciones causales? La respuesta del autor a esta pregunta viene dada a través del establecimiento de dos proposiciones que conforman el núcleo de su teoría:

CC1: un proceso causal (*causal process*) es una línea de mundo de un objeto que posee una cantidad conservada.

CC2: una interacción causal (*causal interaction*) es una intersección de líneas de mundo que involucra el intercambio de una cantidad conservada.

En estas proposiciones, la *línea de mundo* de un objeto es la colección de puntos en un diagrama de Minkowski del espacio-tiempo que representa la historia del objeto; una *cantidad conservada* es cualquier cantidad gobernada por un principio de conservación y una *intersección de procesos* (líneas de mundo) es la superposición en el espacio-tiempo de dos o más procesos.

El segundo eje sobre el cual se articula la TCC se relaciona con el problema de la conexión causal. En particular, aquí Dowe pretende responder a la pregunta acerca de cuál es la conexión entre causas y efectos. Después de establecer que los candidatos más adecuados para ocupar el lugar de los *relata* causales son los hechos o eventos, Dowe

afirma que habrá una conexión causal entre dos *relata* si existe un conjunto de procesos causales e interacciones causales entre ellos, tales que (i) cualquier cambio en los *relata* o cualquier cambio en las cantidades conservadas involucradas ocurre en una interacción causal y (ii) el cambio en las cantidades conservadas se encuentra gobernado por una única ley de la naturaleza.

El tercer y último eje abordado por Dowe en su TCC se relaciona con el problema de la asimetría causal, es decir, el autor aborda la pregunta acerca de qué hace a una causa diferente de su efecto. La respuesta de Dowe a este problema apela a la noción de bifurcación conjuntiva de Reichenbach. De este modo, la dirección de un proceso causal se encuentra determinada, según Dowe, por la dirección de la bifurcación conjuntiva abierta de la cual forma parte este proceso o, de no ser posible lo anterior, por la dirección de la mayoría de las bifurcaciones abiertas que forman parte de la red en la que participa este proceso.

3.1. El programa de Dowe

La TCC de Dowe (2000) es una teoría de la causación física propuesta por su autor como una solución frente a algunas deficiencias que, a su criterio, tendría la teoría previa en esta misma línea brindada por Salmon (1984, 1994). Como fue señalado en el capítulo anterior, en una serie de artículos publicados en el año 1992, Dowe analiza las dificultades de la propuesta de Salmon, a la vez que presenta una primera versión de su TCC. Algunos trabajos posteriores, tanto de Salmon como de Dowe, permitieron agudizar aún más el debate (Dowe 1995a, 1995b; Salmon 1997). Finalmente, en el año 2000, Dowe presenta la versión más acabada de la TCC, presentada en su libro *Physical Causation*. El análisis que se lleva a cabo en este capítulo refiere a esta obra de Dowe.

El objetivo principal de Dowe en su obra del año 2000 es articular y defender una teoría de la causación física. Con este propósito, el autor comienza por considerar que esta tarea puede abordarse desde dos enfoques diferentes: el conceptual y el empírico (Dowe, 2000, p. 1). Según lo entiende Dowe, el análisis conceptual es un análisis del significado, apoyado en el sentido común y en la comprensión cotidiana de los términos. Lo relevante a destacar aquí es que, para el autor, este tipo de análisis es *a priori* y, de ser verdadero, será

necesariamente verdadero (Dowe, 2000, p. 2). Por otra parte, el análisis empírico busca establecer qué tipo de causación se encuentra en el mundo real; el objetivo de este tipo de análisis de la causación es, afirma el autor, mapear el mundo objetivo y no nuestros conceptos. Por lo tanto, solo es posible efectuar este tipo de análisis *a posteriori* (Dowe, 2000, p. 3). Ahora bien, un análisis *a posteriori* de la relación causal deja abierta dos posibilidades en caso de ser verdadero: puede ser una verdad contingente o una verdad necesaria. Dowe elige la primera de estas alternativas. En suma, el autor busca establecer un análisis empírico de la causación física que tenga el estatus de verdad contingente.

En este sentido, Dowe sostiene que es un error común pedir a un análisis de este tipo que se cumpla para todos los mundos lógicamente posibles (Dowe, 2000, p. 6). De este modo, se alinea con otros análisis empíricos de la causación, como los de Aronson, Fair y Salmon, que pretenden cumplirse solo en el mundo actual, ya que no buscan establecer identidades necesarias entre los conceptos involucrados (Dowe, 2000, p. 6). Lo que el autor propone es un análisis empírico basado en la ciencia. Sostiene que, de alguna manera, la ciencia “informa” a la filosofía sobre los elementos que intervienen en una relación causal y, de esta forma, la filosofía puede aprovechar estos resultados usándolos en su favor (Dowe, 2000, p. 8). Es precisamente por este motivo que Dowe sostiene que su análisis es un análisis de la causación *física*.

En líneas generales, el trabajo de Dowe se articula en torno a tres preguntas centrales que guían el desarrollo de su teoría. En primer lugar, se plantea responder a la pregunta: ¿qué son los procesos e interacciones causales? Esta tarea da lugar a una particular visión acerca de los procesos e interacciones causales, donde lo que guía la caracterización propuesta es la distinción entre procesos causales y pseudo-procesos. La idea central de Dowe en este punto es que la posesión de una cantidad conservada, más que la transmisión de una marca (como postula Salmon), es lo que permite distinguir entre procesos causales y pseudo-procesos (Dowe, 2000, p. 89). La segunda pregunta que Dowe se propone responder es: ¿cuál es la conexión entre causa y efecto? Esta pregunta es abordada por su autor en el Capítulo 7 de su libro, ofreciendo un enfoque para la conexión causal basado en las nociones de proceso e interacción causal. Por último, en el Capítulo 8, Dowe enfrenta la pregunta: ¿qué distingue la causa de su efecto? Acerca de este problema, el autor ofrece una

propuesta para dar cuenta de la asimetría causal basada en la noción de bifurcación conjunta de Reichenbach.

En las siguientes secciones seguiremos esta estructura. De este modo, en la Sección 3.2 abordaremos las nociones de proceso causal e interacción causal. Luego, en la Sección 3.3, analizaremos la propuesta de Dowe acerca del problema de la conexión causal. A continuación, en la Sección 3.4, examinaremos la solución de Dowe al problema de la asimetría causal. Por último, en la Sección 3.5 ofreceremos unos breves comentarios finales.

3.2. ¿Qué son los procesos e interacciones causales?

Para responder a la pregunta acerca de qué son los procesos e interacciones causales, Dowe sostiene que su teoría puede ser formulada mediante las siguientes dos proposiciones (2000, p. 90)⁴⁷:

CCI: un proceso causal (*causal process*) es una línea de mundo de un objeto que posee una cantidad conservada.

CC2: una interacción causal (*causal interaction*) es una intersección de líneas de mundo que involucra el intercambio de una cantidad conservada.

Para Dowe, un *proceso* es una línea de mundo de un objeto, donde la *línea de mundo* de un objeto es la colección de puntos en un diagrama de Minkowski del espacio-tiempo que representa la historia del objeto (Dowe, 2000, p. 90). De esta manera, los procesos son “gusanos” en el espacio-tiempo.

En el lenguaje de la teoría especial de la relatividad, las líneas de mundo pueden ser tipo-tiempo (*timelike*), tipo-espacio (*spacelike*) o tipo-luz (*lightlike*). Una línea de mundo es tipo-tiempo si todos sus puntos se encuentran dentro del cono de luz futuro desde el punto de partida considerado. Una línea de mundo es tipo-espacio si algunos de sus puntos se

⁴⁷ En el Capítulo 5 de su libro del año 2000, así como también en Dowe (1992a, 1992c), el autor sostiene explícitamente que tanto *CCI* como *CC2* representan *definiciones* de los conceptos de proceso e interacción causal (Dowe, 2000, p. 93). Sin embargo, en el Capítulo 7 de esta tesis, mostraremos que estas definiciones se convierten en verdades contingentes en la argumentación de Dowe, cuando intenta responder a algunas críticas a su teoría de la causación.

encuentran fuera del cono de luz. Una línea de mundo es tipo-luz si todos sus puntos se encuentran sobre el cono de luz. Según la teoría especial de la relatividad, el tipo de línea de mundo del objeto considerado indica a qué velocidad se mueve dicho objeto: si es tipo-tiempo, su velocidad es menor que la velocidad de la luz, si es tipo-luz, su velocidad es la de la luz, y si es tipo-espacio, al menos en alguna región espacio-temporal supera la velocidad de la luz. Para Dowe, independientemente del tipo de línea de mundo de la que se trate, cualquier línea de mundo es un proceso. Sin embargo, en términos relativistas, un punto dentro del cono de luz y un punto fuera del cono de luz no pueden vincularse causalmente. Siguiendo estas ideas físicas, Dowe dirá que si un proceso es tipo-tiempo o tipo-luz, es un proceso causal; si es tipo-espacio, se trata de un pseudo-proceso.

En otros términos, un pseudo-proceso no es un proceso causal porque ocurre a velocidades superiores a la de la luz. Para ejemplificar esta afirmación, Dowe recurre al caso de un punto de luz moviéndose sobre un muro (Dowe, 2000, p. 90). La trayectoria que describe ese punto de luz es un pseudo-proceso, ya que el movimiento del punto de luz en el muro puede ocurrir a velocidades mayores a la de la luz. Lo mismo ocurre, por ejemplo, si apuntamos con un potente láser a dos puntos distantes en el cielo: el movimiento del punto de luz desde un punto a otro en el cielo podría ocurrir a velocidades mayores a la de la luz, por tanto, este proceso califica como pseudo-proceso.

Ahora bien, para el autor, lo esencial en la distinción entre procesos causales y pseudo-procesos reside en la posesión⁴⁸, por parte del objeto, de una *cantidad conservada* (como se puede ver en *CCI*), esto es, de una propiedad física del objeto cuya magnitud no varía a través del proceso. Así, el movimiento del punto de luz en el muro es un pseudo-proceso no solo porque podría ocurrir a velocidades mayores a la de la luz (lo cual es efectivamente el caso), sino porque en la dirección del movimiento el proceso no posee una cantidad conservada. Lo mismo puede decirse del desplazamiento del punto de luz entre dos puntos distantes en el cielo. Por el contrario, si consideramos la dirección de la trayectoria del haz de luz, este califica como proceso causal, ya que en esa dirección el haz de luz sí posee cantidades conservadas; por ello, es capaz de “causar” en esa dirección. Este fenómeno, de acuerdo con Dowe, califica como proceso causal.

⁴⁸ Aquí “posesión”, de acuerdo con Dowe, debe entenderse como instanciación de una propiedad (Dowe, 2000, p. 92).

Otro ejemplo que permite clarificar aún más la distinción entre proceso causal y pseudo-proceso es el siguiente. Considérense los miles de espectadores de un multitudinario partido de fútbol, quienes, después de celebrar un gol de su equipo favorito, se organizan para hacer “la ola”. Estas ondas, que es posible visualizar cuando las tribunas se observan desde una cierta distancia, califican como un pseudo-proceso, ya que no existen cantidades conservadas en la dirección del movimiento de la ola. Sin embargo, si nos enfocamos en un único espectador, podremos percatarnos que el movimiento que realiza es el de pararse para luego sentarse: considerando la dirección de este movimiento individual, puede decirse que corresponde a un proceso causal, ya que en esa dirección el objeto (el espectador) sí posee cantidades conservadas.

En un intento por precisar aún más su teoría, Dowe introduce una distinción más general respecto de los gusanos espacio-temporales. Sostiene que no todos los gusanos espacio-temporales son líneas de mundo o, lo que es lo mismo, no todos los gusanos espacio-temporales son procesos (Dowe, 2000, p. 91). Para que un gusano espacio-temporal sea una línea de mundo, esto es, un proceso, es necesario que el objeto representado por dicho gusano exhiba identidad a través del tiempo. En otras palabras, la proposición *CCI* supone que las distintas partes del proceso son las mismas en diferentes momentos (Dowe, 2000, p. 91). Esta exigencia de identidad del objeto a través del tiempo permite restringir los gusanos espacio-temporales que califican como procesos, ya que existen objetos que no exhiben identidad a través del tiempo o del espacio. Dowe denomina a estos objetos, los que no exhiben identidad, “basura espacio-temporal”; ejemplos de ellos son los llamados “impostores temporales” (*timewise gerrymanders*) e “impostores espaciales” (*spacewise gerrymanders*).

Los impostores temporales, según Dowe, son objetos putativos que se definen de forma distinta en diferentes momentos (Dowe, 2000, p. 99). Así, por ejemplo, un impostor temporal podría ser “el objeto que poseo en mi bolsillo”, puesto que este objeto será distinto en diferentes momentos. En mi bolsillo puedo tener en un tiempo t_1 una moneda, en un tiempo t_2 un bolígrafo y en un tiempo t_3 un teléfono celular. Por otra parte, los impostores espaciales corresponden a objetos putativos que son la suma mereológica de objetos en un intervalo de tiempo determinado. Como ejemplifica el mismo Dowe, uno de estos objetos puede ser una moneda en mi bolsillo más una lapicera en el escritorio más un

reloj en mi muñeca. Según el autor, es claro que las líneas de mundo de estos impostores no constituyen procesos legítimos.

En definitiva, para la TCC un proceso (causal o no) es una línea de mundo de un objeto, es decir, un gusano espacio-temporal. Este objeto debe exhibir identidad a través del tiempo, lo cual descarta los gusanos espacio-temporales que no cumplen con esta condición y que, por tanto, no son procesos. Quedan excluidos los impostores temporales y espaciales. Ahora bien, entre los gusanos espacio-temporales que son procesos, pueden distinguirse los causales de los no causales, dependiendo de si poseen cantidades conservadas o no, respectivamente. Esta clasificación puede esquematizarse del siguiente modo:

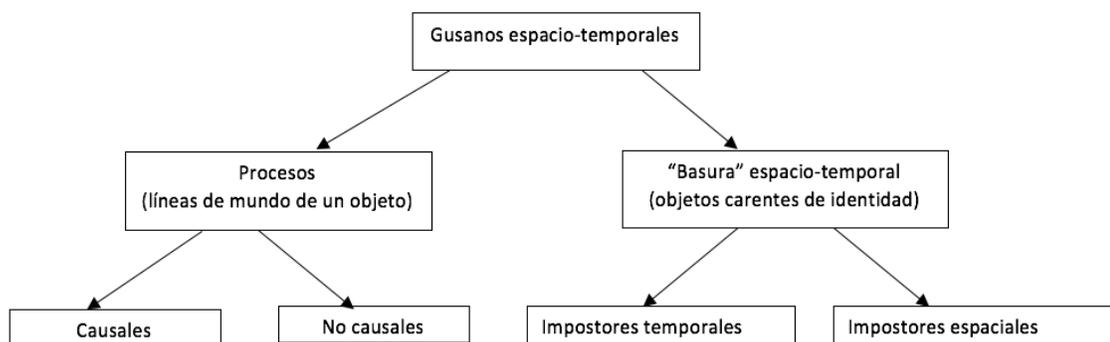


Figura 3.1.: Esquema de clasificación de los gusanos espacio-temporales

Hasta aquí parece quedar claro que un proceso causal es una línea de mundo que posee una cantidad conservada, pero aún no se ha especificado qué se entiende por cantidad conservada. Para Dowe, una *cantidad conservada* es cualquier cantidad gobernada por un principio de conservación. Según el autor, las teorías científicas actuales son nuestra mejor guía para saber cuáles son las cantidades que se conservan: tales teorías deben ser nuestro referente al momento de decidir si una cantidad se clasifica como conservada o no (Dowe, 2000, pp. 91-94). De esta manera, cantidades como energía, momentum y carga eléctrica pueden ser consideradas, de acuerdo con el criterio de Dowe, como cantidades conservadas. Por lo tanto, el papel que juegan los principios de conservación es central a la hora de elucidar la idea de proceso causal, ya que son ellos los que permiten identificar qué

cantidades se conservan (como el momentum o la energía) y cuáles no (como la velocidad) (Dowe, 2000, p. 95).

Un punto relevante para tener en cuenta es que las cantidades que el autor señala como conservadas no pertenecen a una misma y única teoría física. Por un lado, la carga eléctrica es un concepto físico que proviene del electromagnetismo, y es en esta teoría que se cumple el principio de conservación de la carga. Por otra parte, energía y momentum son cantidades que pueden hallarse en el marco de más de una teoría física, por ejemplo, en la mecánica clásica, en la relatividad especial y también en la teoría general de la relatividad, pero los correspondientes principios de conservación no se cumplen del mismo modo en todas las teorías en las que estos conceptos aparecen. Por ejemplo, en mecánica clásica se enuncia un principio de conservación de la energía y, por otra parte, un principio de conservación del momentum. En cambio, en la relatividad especial y en la teoría general de la relatividad no se formula un principio independiente para cada una de estas cantidades, sino que se habla de la conservación de la energía-momentum. Por lo tanto, cuáles son las cantidades conservadas depende de la teoría física que se considere. Es por ello que podemos válidamente preguntarnos a qué cantidades conservadas se refiere exactamente Dowe.

Respecto de *CC2*, una intersección de procesos (líneas de mundo) es la superposición en el espacio-tiempo de dos o más procesos. La intersección ocurre entonces en una *región local* del espacio-tiempo que consta de todos los puntos del espacio-tiempo que son comunes a ambos procesos (Dowe, 2000, p. 92). Cabe destacar que el hecho de que la intersección ocurra en una región local del espacio-tiempo es un claro indicio de que la teoría de Dowe es *local*⁴⁹: la interacción causal es independiente de lo que suceda fuera de la región espacio-temporal en la que ocurren los eventos. En consecuencia, todo proceso causal dependerá de hechos locales sobre el proceso, es decir, de la posesión por parte del objeto de una cantidad física que se conserve, y no dependerá de lo que ocurra en otros lugares del universo fuera de la región local (Dowe, 2000, p. 96).

⁴⁹ Es importante mencionar que algunos críticos de la TCC consideran que la teoría no es local como supone su autor. Como se señalará más adelante, el objetivo central del artículo de Rueger (1998) es mostrar que el criterio de las cantidades conservadas no es suficiente para capturar el supuesto carácter local de la teoría de Dowe. Ver el siguiente capítulo para más detalles acerca de esta crítica.

La TCC de Dowe es, actualmente, la teoría de la causación física más desarrollada y articulada de la que se dispone. Por un lado, en ella confluye toda una tradición de teorías de la causación física, como las de Aronson, Fair y Salmon, llevando a su punto más alto las aspiraciones reduccionistas y fisicalistas para abordar el problema de la causación. Por otro lado, la teoría de Dowe es la mejor teoría disponible para distinguir casos genuinos de causación de casos espurios, fundamentalmente en el terreno de la física.

3.3. ¿Cuál es la conexión entre causas y efectos?

En el Capítulo 7 de su libro Dowe aborda el problema de qué procesos e interacciones causales deben existir en una relación causa-efecto. Para ello, el autor explora dos posibles respuestas, la teoría de proceso *naïve* y las teorías integradoras, para finalmente ofrecer su propia respuesta a esta pregunta basándose en la metafísica de los estados de cosas de David Armstrong. Analizamos a continuación cada una de estas propuestas⁵⁰.

El enfoque de proceso *naïve*, como el autor lo denomina, pretende establecer las condiciones necesarias y suficientes para la conexión entre causa y efecto de la siguiente manera: dos *relata* causales están conectados por una relación causal si y solo si una línea continua de procesos causales e interacciones causales existe entre ellos (Dowe, 2000, p. 147). El problema de establecer las conexiones entre causa y efecto de este modo, según el autor, es que la teoría de proceso *naïve* parece proveer exitosamente una condición necesaria para la causación, pero no tendría éxito para proveer una condición suficiente.

En efecto, Dowe presenta dos casos que sirven como contraejemplo para el enfoque de proceso *naïve* como condición suficiente. El primero de ellos es formulado por Cartwright (1983), quien plantea la situación en que una planta es rociada con un defoliante que mata a nueve de cada diez plantas, pero la planta rociada sobrevive. En esta situación es posible establecer una serie de procesos e interacciones causales que conecten el rociamiento de la planta con su sobrevivencia, sin embargo, no diríamos que la causa de que la planta sobreviva es que fue rociada. El segundo ejemplo que menciona Dowe es presentado por David Papineau (1989): ser un niño o niña con obesidad no es la causa de que en la adultez se sea delgado o delgada, sin embargo, nuevamente, podemos establecer

⁵⁰ El análisis causal de Dowe se basa en la causación singular.

ciertos procesos e interacciones causales que conecten ambas situaciones. Dowe llama a esta clase de casos “conexiones erróneas” (*misconnections*).

El segundo enfoque que explora Dowe para dar respuesta al problema de la conexión entre causas y efectos es el de las soluciones integradoras, como las denomina el autor. Las soluciones integradoras combinan el enfoque probabilista de la relación entre causa y efecto (recuérdese Capítulo 1, Sección 1. 4) con el enfoque de procesos de las relaciones causales, es decir, como hemos visto, la idea de que la causa y el efecto se encuentran conectados por un proceso causal. El propio Dowe propone una de estas soluciones integradoras, donde los procesos causales juegan un papel preponderante en comparación con otras propuestas en este sentido, donde la parte probabilista de la teoría es la que sobresale (por ejemplo, Lewis, 1986a; Eells, 1991). Resumidamente, la propuesta de solución integradora de Dowe considera que *C* es una causa de *E* si y solo si (i) hay un conjunto de procesos e interacciones causales entre *C* y *E* (estos procesos e interacciones se encuentran caracterizados por *CC1* y *CC2*), y (ii) la posibilidad (*chance*) de que ocurra *E* a partir de *C* –dado un proceso causal conectando *C* y *E*– es mayor que la posibilidad de que ocurra *E* si *C* no ocurre (2000, p. 167).

Sin embargo, aunque el autor afirma que esta solución integradora es bastante más sólida que sus rivales, también señala que esta debe ser abandonada pues, por un lado, aún sobreviven algunos casos de conexiones erróneas que no pueden ser explicados por esta nueva propuesta⁵¹ y, por otro lado, porque no se estaría apuntando en la dirección correcta al querer analizar la causación en términos del aumento de posibilidades⁵².

Después de descartar tanto la propuesta de procesos *naïve* como las propuestas integradoras, Dowe ofrece su propio enfoque acerca de la conexión causal. El enfoque de Dowe acerca de la conexión entre causa y efecto da cuenta, por un lado, de los *relata* causales basándose en la metafísica de Armstrong (1997) y, por otro lado, explica cómo estos *relata* se encuentran conectados apelando a las nociones de procesos e interacciones

⁵¹ No se describe aquí el ejemplo presentado por Dowe porque no resulta suficientemente claro.

⁵² El autor no brinda una argumentación en favor de este punto porque considera que esto se encuentra más allá de los alcances de su libro.

causales. De acuerdo con el autor, su enfoque permite sortear satisfactoriamente los casos de conexiones erróneas.

Respecto de los *relata* causales, Dowe acepta en su teoría que estos puedan ser tanto eventos como hechos, pues ambos refieren a objetos y esto se ajusta bien a su TCC. Los eventos, de acuerdo con el autor, pueden definirse como un cambio en una propiedad de un objeto a la vez, mientras que los hechos son objetos que tienen una propiedad a la vez o durante un periodo de tiempo (2000, pp. 169–170). Asimismo, Dowe exige que los *relata* causales, entendidos como hechos o eventos, cumplan ciertas características. En primer lugar, en el marco de una relación causal, estos hechos o eventos deben involucrar cantidades conservadas o supervenir de hechos o eventos que involucran cantidades conservadas (2000, p. 170). En segundo lugar, el autor también distingue entre lo manifiesto y lo físico, afirmando que los hechos causales manifiestos supervienen de hechos causales físicos (2000, p. 170). Por último, Dowe afirma que ningún hecho causal físico es disyuntivo o negativo, lo cual permite una comprensión adecuada, afirma el autor, de los problemas asociados a la prevención y omisión en el marco de su teoría (2000, p. 170).

Por otra parte, respecto del problema de cómo se encuentran conectados las causas con sus efectos, Dowe plantea la siguiente definición de conexión causal para dos interacciones I_1 e I_2 conectadas por un proceso p (2000, p. 171):

Conexión causal: las interacciones I_1 e I_2 están conectadas por una conexión causal en virtud de un proceso causal p solo si alguna cantidad conservada intercambiada en I_2 es también intercambiada en I_1 y poseída por p .

Es importante recordar que, en esta definición de conexión causal, I_1 e I_2 se encuentran definidas en términos del intercambio de una cantidad conservada, mientras que p se encuentra definido sobre la base de la posesión de una cantidad conservada, en correspondencia con las proposiciones *CCI* y *CC2* (Dowe, 2000, p. 171). Ahora bien, Dowe también ofrece una definición más general de conexión causal para objetos que poseen diferentes cantidades conservadas. Supongamos, afirma el autor, que la causa y el efecto se denotan por $q(a)$ y $q'(b)$, respectivamente. En esta notación a y b son objetos, mientras que q y q' son las cantidades conservadas poseídas por estos objetos. De este

modo, el autor presenta la nueva definición de conexión causal del siguiente modo (2000, pp. 171–172):

Conexión causal: hay una conexión causal (o un hilo de conexiones causales) entre un hecho $q(a)$ y un hecho $q'(b)$ si y solo si hay un conjunto de procesos causales e interacciones causales entre $q(a)$ y $q'(b)$ tales que:

- (1) cualquier cambio del objeto de a a b y cualquier cambio de la cantidad conservada de q a q' ocurre en una interacción causal involucrando los siguientes cambios: $\Delta q(a)$, $\Delta q(b)$, $\Delta q'(a)$ y $\Delta q'(b)$; y
- (2) para cualquier cambio en (1) involucrando más de una cantidad conservada, el cambio en las cantidades se encuentra gobernada por una única ley de la naturaleza.

En la colisión de dos bolas de billar a y b (la segunda en reposo), como ejemplifica Dowe, existe un hilo de conexiones causales entre las dos bolas que colisionan. En efecto, en primer término, tenemos la bola a poseyendo una cierta cantidad de momentum, $q(a)$, en un tiempo t_1 ; luego, en la interacción entre a y b ocurren los intercambios de momentum entre ellas, $\Delta q(a)$ y $\Delta q(b)$; por último, en el tiempo t_2 la segunda bola adquiere el momentum $q(b)$ que permite su movimiento.

Resumiendo, Dowe ofrece una explicación para dar cuenta de la conexión entre causa y efecto basada en las nociones de interacción causal y proceso causal que, de acuerdo con su autor, evita satisfactoriamente el problema de las conexiones erróneas (ver Capítulo 4, Sección 4.4, para una crítica respecto de la propuesta de conexión causal de Dowe). Continuemos ahora con la respuesta que ofrece Dowe al problema de la dirección de la causación.

3.4. ¿Qué hace a una causa diferente de su efecto?

En el Capítulo 8 de su libro, Dowe se aboca a la tarea de dar una respuesta a la aparente asimetría de la relación causal, es decir, a la tarea de determinar qué hace diferente una causa de su efecto. Si bien el carácter asimétrico de la relación causal aún es objeto de debate en el ámbito de la filosofía, nuestra experiencia cotidiana parece indicarnos que, en

efecto, la causación presenta esta característica: afirmar, por ejemplo, que la causa de la rotura del vaso de cristal al caer al piso (el efecto) fue el movimiento desprevenido de mi mano no implica afirmar que la rotura del vaso fue la causa del movimiento de mi mano. Dowe, en consecuencia, acepta en su análisis de la dirección de la causación que las relaciones causales son asimétricas y que, por lo tanto, existe una dirección objetiva para la relación causa-efecto de la que debemos dar cuenta⁵³. Para arribar a una respuesta satisfactoria, el autor explora tres tipos de teorías acerca de la dirección de la causación: la teoría temporal, la teoría subjetiva y la teoría física o de bifurcaciones (*forks*).

Ahora bien, el análisis que Dowe lleva a cabo acerca de las tres teorías de la dirección de la causación anteriormente mencionadas se desarrolla sobre la base de una premisa: la veracidad del modelo de retrocausación o causación hacia atrás⁵⁴ (*backwards causation*). En efecto, el autor evalúa las diferentes propuestas que intentan dar respuesta al problema de la dirección causal procurando que estas sean compatibles con los casos de retrocausación. La motivación de Dowe para imponer esta premisa tiene relación con su convencimiento de que la retrocausación es un modelo adecuado para dar cuenta de algunos fenómenos característicos de la mecánica cuántica, como el fenómeno de Bell, por ejemplo (ver Dowe, 2000, pp. 180–187, para más detalles acerca de la retrocausación y la mecánica cuántica).

La primera teoría de la dirección de la causación evaluada por Dowe es la teoría temporal. En términos generales, esta teoría afirma que la dirección causal se reduce a la dirección temporal. Sin embargo, esta alternativa es descartada rápidamente por el autor, pues no es compatible con los modelos retrocausales. En efecto, como señala Dowe, por definición es imposible que una medida posterior tenga alguna influencia causal sobre el estado temprano de una partícula, por ejemplo (2000, p. 188).

La segunda teoría analizada por Dowe que pretende dar cuenta de la dirección de la causación es la teoría subjetiva. Como explica el autor, esta teoría afirma que la dirección causal no es una característica del mundo de los objetos, eventos y procesos, sino de la

⁵³ Dowe menciona que incluso si la relación causal fuera no-simétrica, igualmente tendríamos que dar cuenta de la dirección de la causación (2000, p. 179).

⁵⁴ Muy aproximadamente, la retrocausación es un modelo causal en el cual el efecto precede en el tiempo a la causa.

forma en que percibimos esos objetos, eventos y procesos (2000, p. 188). La teoría subjetiva se presenta bajo dos formas, una forma fuerte y una forma débil. La primera afirma que la causación en sí misma no es un aspecto objetivo del mundo, mientras que la segunda afirma que la causación sí es objetiva, pero la asimetría causal no lo sería⁵⁵ (2000, p. 188). Dowe también descarta esta segunda alternativa debido a que, si bien, en principio, el subjetivismo es compatible con la retrocausación, esta no sería compatible, en particular, con el tipo de retrocausación necesario para explicar el fenómeno de Bell en mecánica cuántica. De acuerdo con Dowe, el problema consiste en que la explicación de este tipo de fenómenos exige que la dirección de los procesos involucrados en él deba ser objetiva y no dependiente de la forma en que experimentamos el mundo (2000, p. 189).

Así, llegamos al último tipo de teoría analizada por Dowe para dar cuenta de la asimetría causal, la teoría física o de bifurcaciones. El primer autor en concebir la dirección de la causación en términos de la noción de bifurcación fue Reichenbach (acerca de la noción de “bifurcación”, recuérdese Capítulo 2, Sección 2.4). Basándose en la propuesta de este autor, Dowe define la noción de bifurcación causal en términos de procesos e interacciones causales, afirmando que una *bifurcación causal* puede ser definida como el caso en el cual dos procesos causales surgen de una interacción causal (2000, p. 192). Esta es, de modo general, la definición de una bifurcación conjuntiva que hemos visto en el Capítulo 2.

Para dar cuenta de la asimetría causal, Reichenbach afirma que la dirección de los procesos causales se encuentra dada por la dirección de las bifurcaciones conjuntivas abiertas⁵⁶. Basándose en la idea de Reichenbach, Dowe elabora su propia propuesta para dar cuenta de la dirección de las relaciones causales. De acuerdo con el autor:

“la dirección de un proceso causal se encuentra dado por la dirección de una bifurcación conjuntiva abierta parcialmente constituida por ese proceso; o, si no existe tal bifurcación conjuntiva, por la dirección de la mayoría de las

⁵⁵ El trabajo de Price (1992) es una de las principales defensas del subjetivismo débil.

⁵⁶ Es importante mencionar que la propuesta de Reichenbach acerca de las bifurcaciones conjuntivas abiertas se basa en la premisa de que estas no existen abiertas al pasado, sino solo al futuro.

bifurcaciones abiertas contenidas en la red en la que se encuentra el proceso” (2000, p. 204).

De este modo, de acuerdo con Dowe, la dirección de un proceso causal se encuentra determinado o bien por su participación en una bifurcación conjuntiva abierta o bien por la asociación de este proceso con la red causal en la que se ubica (2000, p. 205). Asimismo, un proceso causal puede tener dos direcciones, pues la propuesta de Dowe se basa en bifurcaciones conjuntivas las que, a su vez, se definen en términos de las frecuencias de los tipos de eventos (2000, p. 205). Esta propuesta, según el autor, sería compatible con los casos de retrocausación, que es el objetivo que el autor buscaba.

3.5. Comentarios finales

La TCC de Dowe es una teoría de la causación física basada en las nociones de proceso causal e interacción causal. Estas nociones se encuentran definidas por las proposiciones *CCI* y *CC2* presentadas en la Sección 3.2. Sobre la base de los procesos e interacciones causales, Dowe también hace frente al problema de la conexión entre causa y efecto. De modo general, la propuesta del autor es que existe una conexión de tipo causal entre dos hechos o eventos si existe un conjunto de procesos e interacciones causales que conectan ambos *relata* causales. Asimismo, las cantidades conservadas involucradas en los procesos e interacciones deben estar gobernadas por una ley de la naturaleza. Por último, hemos visto que Dowe también elabora una propuesta para dar cuenta de la asimetría causal, basada en la noción de bifurcación conjuntiva de Reichenbach. Su propuesta consiste en que la dirección de un proceso causal se encuentra determinada por la dirección de una bifurcación conjuntiva abierta o por la dirección de la mayoría de las bifurcaciones abiertas que componen la red en la que se encuentra ese proceso.

Capítulo 4. *Críticas a la Teoría de las Cantidades Conservadas*

Resumen

En este capítulo se presentarán las principales críticas que se han formulado a la TCC. No se pretende brindar una lista exhaustiva de críticas, sino solo resaltar aquellas que generaron un cierto impacto en el debate acerca de la teoría de la causación de Dowe.

Nuestro recorrido por las críticas a la TCC comenzará con el debate acerca de la distinción entre procesos causales y pseudo-procesos y su relación con los términos “transmitir” y “poseer” (Sección 4.1). A continuación, presentaremos las críticas relacionadas con la identidad a través del tiempo, tanto de los objetos como de las cantidades físicas conservadas (Secciones 4.2 y 4.3). Luego, expondremos los problemas que presenta la TCC de Dowe para dar cuenta de las interacciones causales en casos de superposición (Sección 4.4). Continuaremos con la propuesta de Dowe acerca de las prevenciones y omisiones y las críticas que se han formulado a esta propuesta (Sección 4.5). A continuación, presentaremos algunas críticas provenientes del ámbito de la filosofía de la física. Revisaremos algunas dificultades asociadas con la noción de interacción causal ofrecida por Dowe y su relación con el cambio en el valor de las cantidades conservadas intercambiadas; también expondremos una crítica que se asocia con la noción de ley de conservación y su relación con la noción de sistema cerrado (Sección 4.6). Finalmente, presentaremos un conjunto de críticas que se asocia con ciertas dificultades que enfrenta la TCC en contextos relativistas generales. En efecto, las leyes de conservación juegan un papel preponderante en la teoría de la causación de Dowe; sin ellas varios de los conceptos fundamentales utilizados por el autor en la construcción de su teoría se ven socavados. El problema reside en que en la relatividad general no es claro que existan leyes de conservación correctamente definidas (Sección 4.7).

Esta última crítica es central para el estudio que se lleva a cabo en esta investigación. Es a partir del examen de esta crítica, junto con las respuestas del autor a ellas, que se abre una vía de análisis para determinar las condiciones de posibilidad de la TCC –o de cualquier teoría de la causación que pretenda formularse en estos términos– en las teorías físicas del espacio-tiempo, desde la perspectiva de tres problemas filosóficos

independientes: *el problema de las leyes de conservación, el problema de la posibilidad física y el problema de la definición*. Este estudio se iniciará en el capítulo siguiente.

4.1. Transmitir en lugar de poseer

La primera crítica a mencionar aquí es la enunciada por Salmon (1994, p. 308), quien sostiene que en la formulación de Dowe, en particular en la proposición *CCI*, es más adecuado hablar de “transmitir” en lugar de “poseer”⁵⁷. Salmon ejemplifica esta propuesta invitándonos a pensar en un clásico ejemplo de pseudo-proceso (este ejemplo aparece por primera vez en Salmon, 1984, pp. 145–146): un haz de luz proveniente de una baliza giratoria concentrada en un punto moviéndose por la pared de un edificio⁵⁸. El argumento de Salmon es que, aun cuando esta situación corresponde a un pseudo-proceso, pues el punto de luz podría llegar a moverse a velocidades mayores a la de la luz, cada punto de la trayectoria del rayo de luz manifiesta (posee) energía, por lo que “poseer” no sería suficiente para caracterizar los procesos causales. Debemos considerar el término “transmitir”, afirma Salmon.

Dowe (1992a, p. 214; 2000, p. 98) aborda el ejemplo de Salmon acerca del punto de luz, manifestando que en esa situación es la pared, y no el punto de luz, el objeto poseedor de la energía. Sin embargo, para Salmon las consideraciones de Dowe son insuficientes. Con el objetivo de mostrar que el término “poseer” no es suficiente para discernir entre procesos y pseudo-procesos, el autor nos propone que consideremos la línea de mundo de la parte de la superficie de la pared que está absorbiendo energía como consecuencia de ser iluminada (1994, p. 308). Esta línea de mundo, de acuerdo con Salmon, manifiesta energía durante todo el periodo en que el punto de luz viaja por la pared, pero al mismo tiempo no sería un proceso causal puesto que la energía no se transmite, sino que es recibida de una fuente exterior (1994, p. 308). El objeto que Salmon pretende que analicemos sería el resultado de la suma de todos los puntos de la pared que están siendo iluminados y considerados solo durante el tiempo que están siendo iluminados. De esta manera, el objeto

⁵⁷ Una crítica interesante proveniente del ámbito de la filosofía de la física respecto de la *posesión* de cantidades físicas es formulada por Luper (2009, Sección 4).

⁵⁸ Otro ejemplo para una objeción similar puede encontrarse en Hitchcock (1995).

a considerar es un pseudo-proceso, pero que sin embargo posee cantidades conservadas (energía) durante el tiempo considerado. La salida de Dowe a este contraejemplo es considerar que el objeto pretendido por Salmon no califica como proceso en su TCC. Para Dowe, los procesos pueden ser causales o pseudo-procesos, pero esto no agota los posibles objetos que pueden ocupar una región del espacio-tiempo en su teoría. Para Dowe, también existen objetos que no califican como procesos, básicamente porque no presentan identidad a través del tiempo (Dowe, 2000, p. 99). Dowe llama a estos objetos “basura espaciotemporal”, y un ejemplo de estos objetos son los “*timewise gerrymanders*” y los “*spacewise gerrymanders*” (acerca de estos conceptos, recuérdese Capítulo 3, Sección 3.2). De este modo, el objeto sobre el que nos invita a pensar Salmon es, de acuerdo con las consideraciones de Dowe, un *timewise gerrymanders* que no tiene cabida en su teoría.

4.2. Identidad de los objetos

La idea de identidad a través del tiempo se encuentra implícita en la TCC. La definición de proceso como línea de mundo que posee una cantidad conservada implica la genidentidad del objeto, es decir, es necesario que el objeto que posee la cantidad conservada sea el mismo a través del tiempo (Dowe, 2000, p. 102)⁵⁹. Dowe asume la identidad del objeto a través del tiempo como no analizable o primitiva a la teoría; sin embargo, Salmon se muestra contrario a esta postura, proponiendo su propia versión de una teoría causal que no contaría con este problema (Salmon, 1997). La objeción de Salmon, más específicamente, tiene que ver con la contradicción de Dowe al invocar elementos no analizables o no accesibles empíricamente, siendo justamente la teoría de Dowe una teoría empírica.

Dowe, asumiendo la crítica de Salmon sobre el problema de la identidad de los objetos a través del tiempo, analiza las posibles salidas a este problema. El autor pone sobre la mesa tres enfoques que intentan dar cuenta de la identidad a través del tiempo, considerando las posibles ventajas y desventajas de cada enfoque en relación con la TCC.

La primera posibilidad es analizar la identidad como identidad estricta. De acuerdo con este enfoque, basado en Armstrong (1980), “un objeto es idéntico a sus otras partes temporales de la misma forma que es idéntico a sí mismo y de la misma manera que

⁵⁹ El autor también aborda esta objeción en Dowe (2000a).

diferentes cosas tienen la misma propiedad” (Dowe, 2000, p. 102). Por lo tanto, un objeto debe estar totalmente presente en un momento para existir en ese momento (Dowe, 2000, p. 102). Dowe sostiene que adoptar el enfoque de la identidad estricta solucionaría la objeción de Salmon, ya que esta visión no requiere conceptos imposibles de analizar empíricamente. Sin embargo, Dowe descarta el enfoque de la identidad estricta, ya que requeriría otro modo de identificar procesos causales como genuinos, porque el filtro de las cantidades conservadas no sería suficiente para descartar los *gerrymanders* temporales (Dowe, 2000, p. 103).

El segundo enfoque que Dowe propone para analizar la identidad del objeto es la similaridad-continuidad a través del tiempo. Esta visión plantea que el mundo está formado por hechos locales particulares (bits) sin conexión lógica entre ellos cuando están espacio-temporalmente separados (Dowe, 2000, p. 103). De esta forma, el punto de vista de similaridad-continuidad plantea que “...los bits en diferentes momentos están conectados como partes temporales del mismo objeto, en el sentido más débil de ‘mismo’, solo si es que las relaciones espacio-temporales y de similitud se mantienen entre los bits” (Dowe, 2000, p. 104). Dowe propone el siguiente ejemplo para aclarar este punto: un objeto redondo, negro y duro en x en t_1 es el mismo objeto en y en t_2 si, aproximadamente, el segundo objeto es también redondo, negro y duro, y en cada punto espacio-temporal entre x e y hay un objeto redondo, negro y duro (2000, p. 104). El autor concluye que nada hay en el enfoque de similaridad-continuidad que contradiga las pretensiones empiristas de no invocar elementos inanalizables o primitivos en una teoría filosófica (2000, p. 104).

Por último, Dowe analiza el enfoque de la teoría causal de la identidad. En palabras del autor: “...para que un objeto muestre identidad a través del tiempo se requiere como condición necesaria que sus partes temporales estén conectadas como causa y efecto” (2000, p. 105). Dowe, descarta rápidamente este enfoque de la identidad, pues si la TCC se basa en ella, entonces tendremos un problema de circularidad. En efecto, bajo este escenario la TCC, que pretende caracterizar la causación, depende de la noción de identidad a través del tiempo y, a su vez, la noción de identidad a través del tiempo depende de la noción de causación (2000, p. 105).

Dowe concluye su análisis acerca de las posibles teorías candidatas para dar cuenta de la noción de identidad de los objetos a través del tiempo, sosteniendo que, si bien la identidad estricta y la similaridad-continuidad a través del tiempo podrían brindar una respuesta a la crítica de Salmon, también presentan sus inconvenientes, y por ello prefiere dejar la noción de identidad del objeto como primitiva en su teoría (Dowe, 2000, p. 107). Sin embargo, Dowe admite que es necesario dar una explicación a este problema, pero también afirma que la filosofía es una empresa permanente en la que los avances siempre plantean nuevas preguntas. Añade, además, como una respuesta provisoria a la crítica de Salmon, que su objeción se encuentra abierta a una crítica: si el concepto de proceso causal no puede ser explicado sin apelar a la noción de identidad a través del tiempo, entonces, apelando a un argumento simétrico, la noción de proceso causal también sería un primitivo inanalizable, lo cual dejaría fuera de juego la propia propuesta de Salmon acerca de la causación.

Como hemos visto, Dowe asume el problema de la identidad como un problema abierto sin solución en su teoría. Así también lo observa Wilfredo Quezada cuando afirma: “Dowe reconoce que él no está en condiciones de reducir o eliminar el estatus primitivo que tiene el concepto de ‘identidad a través del tiempo’ en su teoría y que eso amenaza su carácter estrictamente Humeano, pues introduce algo así como un poder oculto” (2002, p. 94). Un comentario similar es formulado por Hernán Miguel: “por el momento Dowe tomará la noción de identidad como primitiva, aun cuando no es un primitivo inanalizable. Sencillamente en esta etapa del análisis, no disponemos de una forma de dar cuenta de la identidad” (2006, p. 222)⁶⁰.

4.3. Identidad de las cantidades conservadas

En los trabajos de Aronson (1971b) y Fair (1979) ya es posible encontrar algunas declaraciones acerca del problema de la identidad a través del tiempo de las cantidades físicas conservadas. Por ejemplo, Fair afirma que la identificación de la energía y

⁶⁰ El trabajo de tesis de Miguel (2006) es una excelente investigación acerca de las teorías contrafácticas y físicas de la causación. En particular, en el Capítulo 5 puede encontrarse una detallada y minuciosa revisión de las principales dificultades que enfrentan las teorías de la causación física, que se relaciona estrechamente con lo expuesto en este capítulo.

momentum se encuentra asegurada por las leyes de conservación en el marco de un sistema cerrado (1979, p. 234). Del mismo modo, Aronson declara que una de las razones para creer en la identidad de las cantidades físicas tiene relación con la noción de “igualdad numérica” implicada por las leyes de conservación (1971b, p. 144).

Dowe manifiesta no estar de acuerdo con Aronson y Fair respecto de la posibilidad de concebir cantidades como energía o momentum como propiedades que manifiestan identidad a través del tiempo. En efecto, en primera instancia, Dowe se pregunta acerca de qué tipo de identidad a través del tiempo es aquella a la que se refieren Aronson y Fair en sus afirmaciones. Para dar respuesta a esta pregunta, el autor diferencia entre tres tipos de identidad de las cantidades conservadas que podrían sostenerse. Un primer tipo corresponde al sentido usual de identidad que se predica de las propiedades, un segundo tipo está relacionado con el término “la misma cantidad”. A ninguno de los dos tipos anteriores de identidad, afirma Dowe, hacen referencia Aronson y Fair cuando afirman que las cantidades conservadas presentan identidad a través del tiempo. Por último, el tercer tipo de identidad que se podría predicar de las cantidades conservadas se encuentra asociada con la identidad de una sustancia o materia en el tiempo, también llamada genidentidad (Dowe, 2000, p. 56). De acuerdo con Dowe, es en este tercer sentido en el que se piensan las cantidades conservadas exhibiendo identidad a través del tiempo. Sin embargo, el autor concluye que afrontar el problema desde esta perspectiva trae como consecuencia dos puntos de vista disímiles entre sí y sin solución. En efecto, Aronson y Fair asumen que una cantidad conservada, como la energía, presenta identidad a través del tiempo a la vez que no puede no ser poseída. No obstante, afirma Dowe, si la energía es una sustancia, entonces podemos en efecto hablar de identidad a través del tiempo, pero entonces la energía puede existir sin ser poseída, en contra de lo que afirman Aronson y Fair; por otra parte, si es una propiedad cuantitativa que es poseída en el sentido de “instanciada”, entonces efectivamente no puede no existir no poseída, pero ya no podemos hablar de identidad a través del tiempo (Dowe, 2000, p. 57).

La postura de Dowe al respecto es que las leyes de conservación solo requieren una igualdad numérica y no la característica de identidad a través del tiempo (Dowe, 2000, p. 59). Dowe aplica este argumento a la TCC, sosteniendo que hablar de genidentidad de las cantidades conservadas no tiene sentido. Argumenta que su teoría solo requiere que exista

un intercambio de alguna cantidad conservada, este intercambio debe traer como consecuencia el cambio en el valor de dicha cantidad (Dowe, 2000, p. 111). En consecuencia, su teoría no se compromete ni con la identidad a través del tiempo de las cantidades conservadas ni con la ausencia de tal identidad.

En opinión de Quezada, esta dificultad efectivamente se soluciona de acuerdo con los términos expresados por Dowe:

“En otras palabras, no hay espacio en ella [en la teoría de las cantidades conservadas] para genidentidad de las cantidades físicas, puesto que la identidad después de una intersección depende del intercambio de cantidades y ésta a su vez se expresa en términos de la simple noción de cambio en el valor de la cantidad. Esto muestra también que el problema de saber si se ha intercambiado la misma cantidad no es una dificultad para la teoría de las cantidades conservadas: intercambio de cantidades, como lo define Dowe, no requiere (ni descarta) transferencia local de la misma cantidad a través del tiempo.” (Quezada, 2002, pp. 94–95).

Otro autor que concuerda con el punto de vista, según el cual no tiene sentido hablar de genidentidad de las cantidades conservadas, es Dennis Dieks (1986). Para una crítica de las teorías de transferencia y su relación con la noción de identidad a través del tiempo, ver María José García-Encinas (2004).

4.4. Superposición de interacciones causales

En su artículo “Superposición de interacciones causales en la teoría de Phil Dowe”, Hernán Miguel y Jorge Paruelo (2005) sostienen que la TCC no identifica correctamente, en ciertas situaciones, los casos genuinos de causación. Para fundamentar su afirmación los autores proponen un contraejemplo –fundamentalmente a la propuesta inicial de Dowe (1992a, 1992c), pero también aplicable a la versión final de su teoría, Dowe (2000)– y derivan de él toda una familia de ellos. El contraejemplo consiste en suponer dos cuerpos A y B que colisionan, y donde al momento de la colisión el cuerpo B emite espontáneamente una partícula α (Miguel y Paruelo, 2005, p. 128). De acuerdo con la teoría de Dowe, los cuerpos A , B y la partícula α pueden ser identificados como procesos causales, pues son

líneas de mundo que poseen cantidades conservadas, en este caso, momentum lineal y carga eléctrica. Por otra parte, también podemos identificar la colisión de A y B como una interacción causal (donde se intercambia momentum lineal y carga eléctrica), ya que implica la intersección de líneas de mundo de los procesos entrantes (la intersección de A y B) con los procesos salientes (A , B y la partícula α). El problema, de acuerdo con los autores, es que tendríamos que aceptar, según la teoría de Dowe, que la colisión de los cuerpos causó la emisión de la partícula α , situación que no estaríamos dispuestos a aceptar desde la física.

Al analizar la estructura de este contraejemplo, los autores señalan que la situación podría ser analizada por separado, como dos interacciones, una correspondiente a la colisión de los cuerpos A y B y otra correspondiente a la emisión de la partícula α por parte del cuerpo B . No obstante, sostienen los autores, al ocurrir ambas interacciones en el mismo punto del espacio-tiempo, el intercambio de cantidades conservadas (momentum lineal) también ocurre entre el cuerpo A y la partícula α , y esto permitiría, según la teoría de Dowe, identificar (incorrectamente, de acuerdo a lo que dicta la física) la interacción entre los cuerpos A y B como causa de la emisión de la partícula α (Miguel y Paruelo, 2005, p. 129). Los autores argumentan que es posible construir una familia de contraejemplos como el antes descrito; para ello es suficiente pensar en la superposición de dos interacciones independientes que ocurran en el mismo punto del espacio-tiempo y donde, además, al menos una cantidad conservada sea intercambiada en ambas interacciones (2005, p. 130)⁶¹.

Frente a los contraejemplos presentados, Miguel y Paruelo ofrecen una salida para la TCC. Los autores afirman que el uso de condicionales contrafácticos puede ser una herramienta útil para identificar interacciones independientes y, de este modo, distinguir correctamente los casos genuinos de causación de aquellos que no lo son (2005, p. 136). Sin embargo, esta alternativa al mismo tiempo representa una inconsistencia respecto de los objetivos declarados por Dowe, esto es, el hecho de elaborar una teoría de la causación a la luz únicamente de la física (Miguel y Paruelo, 2005, p. 137).

⁶¹ Miguel y Paruelo también dan cuenta de un contraejemplo en el que dos interacciones aparentan encontrarse conectadas causalmente, aun cuando se localizan en diferentes puntos del espacio-tiempo (2005, Sección 5).

4.5. Prevención y omisión

Otro problema al que la TCC debe hacer frente, así como prácticamente cualquier teoría de la causación, es el problema de las prevenciones y omisiones. En general, los casos de causación por prevención tienen como causa un hecho o evento cuya ocurrencia es real, pero donde el efecto es la no ocurrencia de algo; por ejemplo, “la administración adecuada del medicamento previno la muerte del paciente”. Por otra parte, los casos de causación por omisión se caracterizan por tener una omisión como causa (de nuevo, la no ocurrencia de algo) y un efecto cuya ocurrencia es real; por ejemplo, “la no administración de la medicación a tiempo causó la muerte del paciente”. Mas formalmente, Dowe caracteriza estos tipos de causación como causación* (causación estrella); de esta manera, cuando decimos A causa* B existen tres casos: si A es un hecho o evento negativo, entonces tenemos un caso de causación por omisión; por otra parte, si B es un hecho o evento negativo, entonces estamos frente a un caso de causación por prevención; por último, si tanto A como B son hechos o eventos negativos, entonces tenemos un caso de prevención por omisión (Dowe 2000, p. 123). Ahora bien, válidamente Dowe se pregunta cómo eventos o hechos que no ocurren realmente (eventos negativos) pueden ser parte de las relaciones causales.

La solución de Dowe a este problema se basa en considerar que la causación* en realidad no es un caso de causación genuina o real, sino que más bien debe ser comprendida como un caso que involucra ciertos condicionales contrafácticos acerca de casos de causación real (2000, p. 124; ver también Dowe 2001). De acuerdo con el autor, esta solución es general en el sentido de que puede ser aplicada a la mayoría de las teorías de la causación y no solo a la TCC (ver Dowe, 2000, pp. 143–145). En primera instancia, Dowe propone una forma canónica para las prevenciones y omisiones: un caso de prevención puede ser expresado por A causó* no- B y un caso de omisión puede ser representado por no- A causó* B , donde A y B son hechos o eventos genuinos positivos (2000, p. 132).

De esta forma, el esquema que Dowe propone para los casos de prevención es el siguiente (2000, p. 132):

Prevención: A previno B \equiv *A causó* no-B* si

- (1) *A* ocurrió y *B* no, y ocurrió un *x* tal que
- (2) hay una relación causal entre *A* y el proceso debido a *x*, tal que una de las siguientes opciones se cumple:
 - i) *A* es una interacción causal con el proceso *x*, o
 - ii) *A* causa *y*, una interacción causal con el proceso causal *x*, y
- (3) Si *A* no hubiera ocurrido, *x* habría causado *B*.

Donde *A* y *B* son hechos o eventos positivos y *x* e *y* son variables que oscilan entre hechos o eventos.

Por otra parte, el esquema propuesto por Dowe para los casos de omisión es el siguiente (2000, p. 136):

Omisión: no-A causó B* si

- (1) *B* ocurrió y *A* no, y ocurrió un *x* tal que
- (2) *x* causó *B*, y
- (3) si *A* hubiera ocurrido entonces *B* no habría ocurrido, y habría habido una relación causal entre *A* y el proceso debido a *x*, tal que una de las siguientes opciones se cumple:
 - (i) *A* es una interacción causal que involucra el proceso causal *x*, o
 - (ii) *A* causa *y*, una interacción causal que involucra el proceso causal *x*.

donde, de nuevo, *A* y *B* son hechos o eventos positivos y *x* e *y* son variables que oscilan entre hechos o eventos⁶².

Dowe hace frente a algunas objeciones que se han formulado a los esquemas propuestos para los casos de prevención y omisión que acabamos de presentar. Por un lado, se objeta que se podrían usar condicionales contrafácticos más simples para analizar los casos de prevención y omisión. La respuesta de Dowe a esta objeción es que este tipo de propuestas no explicarían claramente la diferencia entre causación* y causación genuina para varias teorías de la causación, como es el objetivo del autor, sino solo para la

⁶² Dowe también ofrece una solución para los casos de prevención por omisión (ver pp. 139–141).

causación expresada en términos contrafácticos (2000, p. 142). Por otro lado, también se objeta que la propuesta de Dowe trae al mundo mucha más causación* de lo que nuestra intuición nos dicta. Para responder a esta objeción el autor apela a que este no es un problema propio de la propuesta que él hace, sino que de cualquier teoría de la causación que acepte la causación indirecta (2000, p. 142).

En un artículo titulado “Causar o dejar que ocurra”, Miguel y Paruelo (2007) hacen notar una crítica respecto de la propuesta de Dowe en relación a la equivalencia entre prevenciones y omisiones. Según Dowe, se cumple una relación de espejo entre los casos de prevención y omisión: “donde un preventor *A* causó* no-*B*, si *A* no hubiera ocurrido, no-*A* hubiera causado *B* por omisión; y donde una omisión no-*A* causó* *B*, si *A* hubiera ocurrido, *A* habría prevenido *B*” (2000, p. 138). De acuerdo con la crítica de Miguel y Paruelo, esta equivalencia no siempre se cumple: esta se cumpliría solo en los casos en que *A* y *B* representan eventos o hechos particulares, pero no sería cierto si *A* y *B* representan estados de cosas (2007, p. 16). Esta misma crítica menciona Miguel (2006), cuando también señala el problema que acarrea el incluir contrafácticos en una teoría de la causación física⁶³.

Respecto del uso de contrafácticos en la TCC por parte de Dowe, igualmente crítico es Quezada (2002), quien en su artículo “Causalidad física: procesos causales y cantidades conservadas” sostiene que, al ser la teoría de Dowe parte de un programa reductivista Humeano a posteriori de la causación, no es adecuado que mantenga en su seno “poderes ocultos” en la forma de contrafácticos o apele a la noción de mundo posible.

4.6. Críticas a la TCC desde el ámbito de la filosofía de la física

Una de las primeras críticas formuladas a la TCC proveniente del ámbito de la filosofía de la física se debe a Hitchcock (1995) y Sungho Choi (2003), quienes sostienen que una teoría de la causación que se basa en cantidades conservadas no estará completa sin una adecuada definición de ley de conservación, con la cual, en general, no se cuenta en las ciencias físicas. El problema que estos autores observan es que, en general, una ley de

⁶³ Otra crítica respecto a la propuesta de Dowe sobre prevenciones y omisiones es de Hunt (2005).

conservación se define mediante la noción de sistema cerrado y este, a su vez, presupone la noción de interacción causal. Esto nos indicaría que la definición de causación en términos de cantidades conservadas es circular.

Por otra parte, Luper (2009) formula una crítica que se relaciona con las intersecciones y los cambios en el valor de las cantidades conservadas. El autor nos invita a pensar en un sistema simple compuesto por dos planetas con la misma masa describiendo una órbita circular alrededor del centro de masa del sistema (p. 69). Para Luper es claro que en el sistema anteriormente descrito los planetas son procesos causales de acuerdo con la TCC, pues estos poseen cantidades conservadas; sin embargo, afirma Luper, no es igualmente claro que la TCC capture adecuadamente el hecho de que estos planetas se encuentran en una interacción causal. Según la TCC, para que exista una interacción causal (i) debe haber una intersección de líneas de mundo y (ii) debe haber un intercambio de cantidades conservadas. De acuerdo a Luper, (i) se cumple solo si incluimos en el análisis a las líneas de mundo del campo gravitacional de los planetas, pero esto obliga a la teoría de Dowe a incluir los objetos físicos como componentes de una relación causal (2009, p. 70). Por otro lado, (ii) se cumple dependiendo de qué entendamos por “cambios en el valor de una cantidad conservada”, pues los planetas conservan constante el módulo de la velocidad, pero cambian en todo momento su dirección. Luper afirma que en la TCC no se especifica si esto cuenta como un cambio en el valor de la cantidad conservada. El mismo autor propone una solución respecto de este último punto. Para Luper, la frase “un cambio en el valor de una cantidad conservada” debe poseer diferentes significados dependiendo de si la cantidad conservada es una cantidad escalar o vectorial (2009, p. 70).

Otra crítica del mismo autor se relaciona con el hecho de determinar qué cantidades contarán como conservadas y cuáles no. Como hace notar Luper, Dowe le deja esta decisión a la ciencia, pues considera que serán nuestras teorías científicas actuales las que nos digan qué cantidades físicas son cantidades conservadas. El problema con esta postura, según Luper, es que la ciencia nos dice que cantidades como el número leptónico, el numérico bariónico y la cantidad TCP (tiempo, carga, paridad) cuentan como cantidades conservadas; sin embargo, este tipo de cantidades físicas no pueden ser transmitidas o intercambiadas, como exige la TCC (2009, Sección 3). Por último, Luper también nota que la TCC de Dowe, al definir las interacciones causales en términos del intercambio de

cantidades conservadas, no permite calificar como causales a las interacciones participantes en sistemas en equilibrio (por ejemplo, un anillo con cierta masa que tiene una masa puntual en su centro) (2009, Sección 5).

4.7. Leyes de conservación en relatividad general

Llegamos así al último conjunto de críticas que revisaremos en este capítulo, las cuales se asocian con la dificultad para enunciar leyes de conservación correctamente definidas en los contextos relativistas generales. Este conjunto de críticas, como profundizaremos al finalizar esta sección, es especialmente relevante para los fines de esta investigación pues, a partir de un examen meticuloso tanto de las propias críticas como de la respuesta de Dowe a ellas, es posible arribar a un análisis más profundo respecto de las condiciones bajo las cuales una teoría de la causación física en términos de cantidades conservadas puede ser formulada en las teorías físicas del espacio-tiempo, como veremos en profundidad en los capítulos siguientes.

De modo general, la crítica respecto de las leyes de conservación en la relatividad general puede ser enunciada del siguiente modo. Como hemos visto, las proposiciones *CCI* y *CC2*, que conforman el núcleo de la TCC, se fundamentan en la noción de cantidad física conservada, las cuales a su vez se encuentran determinadas por las leyes de conservación respectivas⁶⁴. Son las leyes de conservación las que permiten identificar y/o definir aquellas cantidades físicas que contarán como conservadas. Por lo tanto, la TCC se encuentra íntimamente conectada a las leyes de conservación. Es más, la inexistencia de leyes de este tipo en el marco de alguna teoría física podría implicar la inaplicabilidad de la TCC en el marco de esa teoría. El problema para la TCC, argumentan sus críticos, es que en el contexto de la relatividad general existe un amplio debate acerca de la posibilidad de enunciar genuinas leyes de conservación y, por lo tanto, la aplicabilidad de la TCC en el marco de la teoría de Einstein se encontraría en entredicho (ver Rueger 1998, Curiel 2000, Vicente 2002, Hofer 2009, Lam, 2005, 2010, entre otros). Los problemas para enunciar leyes de conservación bien definidas en la relatividad general se deben a la ausencia, en

⁶⁴ Para un estudio de la relación entre causación y leyes de conservación en las teorías de transferencia y proceso en general, ver Kistler (2006, pp. 67–71).

general, de simetrías del espacio-tiempo, las cuales justamente permiten expresar leyes de conservación en forma integral, las necesarias para la TCC⁶⁵.

Es apropiado aprovechar este espacio para dar cuenta de una crítica a la TCC que se encuentra estrechamente relacionada con el problema de las leyes de conservación en la relatividad general. Esta crítica se debe fundamentalmente a Alexander Rueger (1998) y Vicente (2002), quienes afirman, de modo muy general, que las teorías de la causación física de Salmon y Dowe no son buenas candidatas para caracterizar la relación causal como local, como los autores afirman de sus teorías⁶⁶. Respecto de la TCC, según el argumento de Rueger, el hecho de que las cantidades conservadas dependan del espacio-tiempo de fondo es suficiente para sostener que la conexión causal no es una relación intrínseca, ya que dependería de condiciones (como el espacio-tiempo) que están más allá de los eventos relacionados causalmente. Frente a esta situación, Rueger (1998, p. 37) señala que existen dos posibles salidas: o bien (i) se puede insistir en que la noción de causación, en tanto relación local, es correcta y, en consecuencia, concluir que no existen relaciones causales genuinas en el mundo; o bien (ii) se puede considerar la posibilidad de que la intuición local no sea correcta y, por lo tanto, deba ser abandonada. Dowe se hace eco de esta crítica en su trabajo del año 2000, básicamente apelando a que su teoría de la causación tendría el carácter de contingente. La respuesta de Dowe (2000, 2000a) a esta crítica tendrá un papel central para el argumento que desarrollaremos en el Capítulo 7.

Volvamos ahora nuestra atención a las críticas que se relacionan directamente con el problema de las leyes de conservación. Curiel (2000) critica a las teorías de la causación en términos de transferencia por no adecuarse a la relatividad general donde, de acuerdo con su visión, no es posible definir adecuadas leyes de conservación. El objetivo central del autor en su artículo es mostrar que los problemas con la energía en contextos relativistas generales podrían dar lugar a restricciones sobre los tipos de teorías de la causación que pueden ser considerados fundamentales en este contexto físico. En particular, el autor sostiene que, si se considera a la relatividad general, o al menos algunos aspectos de ella,

⁶⁵ En el siguiente capítulo estudiaremos detallada y cuidadosamente el problema de las leyes de conservación en la relatividad general.

⁶⁶ Una respuesta parcial al problema del localismo de la TCC puede ser encontrada en Dowe (2014).

las teorías de la causación en términos de transferencia (dentro de las cuales el autor incluye a la TCC) no pueden ser correctas.

Un punto importante mencionado por Curiel respecto de estas teorías de la causación y su relación con la energía es que una condición necesaria para caracterizar tanto la propagación como el intercambio de energía, y así poder hacer uso de tales teorías, es que los sistemas dinámicos satisfagan el principio de conservación de la energía tanto en su forma diferencial como integral. De este modo, el concepto de conservación de cantidades se vuelve relevante para este tipo de teorías filosóficas: sin una ley de conservación integral no es posible hablar de propagación e intercambio de energía.

Por su parte, Lam (2005, 2010) argumenta que, debido a que en la relatividad general no existe, en general, un genuino principio integral de conservación de la energía-momentum, las nociones de “intercambio” y “posesión” de una cantidad conservada – conceptos centrales en la teoría de Dowe– se ven socavadas (2005, p. 475; 2010, p. 66). En otras palabras, al no haber leyes de conservación, no se cuenta con las herramientas necesarias para regular el intercambio de cantidades físicas en una interacción causal. Por lo tanto, concluye Lam, la caracterización de “causal” en una interacción pierde su significado fundamental (2005, p. 475; 2010, pp. 66–67). El mismo argumento puede encontrarse formulado por Hofer (2009).

En suma, los argumentos antes señalados ponen de manifiesto la existencia de ciertas dificultades para enunciar leyes de conservación legítimas en el ámbito de la relatividad general, lo cual trae consigo consecuencias indeseables para la TCC de Dowe. Si bien este es un hecho que se presenta solamente en contextos relativistas generales, no puede ser ignorado, ya que la TCC necesita de leyes de conservación que la respalden; de lo contrario, no podría siquiera ser formulada. En otras palabras, al no contar con leyes de conservación en el contexto de la relatividad general (o al enfrentar ciertos obstáculos para enunciarlos), la teoría de Dowe tendría dificultades para distinguir procesos causales de no causales en dicho contexto, y esto constituiría un punto débil de su teoría de la causación *física*. Además, la relatividad general es considerada la teoría fundamental de la física para describir el universo a gran escala y, por ello, su marco conceptual no puede ser ignorado por una propuesta que, como la de Dowe, pretende ser empírica y no meramente

conceptual. La importancia de esta crítica es fundamental en el desarrollo de esta investigación, ya que permite poner en evidencia una seria dificultad de la TCC en el ámbito de la relatividad general: una teoría de la causación que pretende ser empírica no debería encontrarse en desacuerdo con nuestra mejor teoría del espacio-tiempo. Es justamente este hecho la motivación inicial de esta investigación.

En efecto, como adelantamos brevemente al inicio de esta sección, un examen detallado de este conjunto de críticas, así como también de la defensa de Dowe frente a ellas, abre una amplia vía de análisis que permite el estudio de las condiciones físicas y metafísicas bajo las cuales una teoría de la causación al estilo de Dowe pueda ser formulada y mostrarse compatible con las teorías físicas del espacio-tiempo, en particular, con la relatividad general. Como veremos en los capítulos siguientes, las condiciones de posibilidad de una teoría de la causación física en contextos relativistas generales pueden ser estudiadas bajo el alero de tres problemas filosóficos independientes: (i) el problema de las leyes de conservación (Capítulo 5), (ii) el problema de la posibilidad física (Capítulo 6) y el problema de la definición (Capítulo 7). Para cada uno de estos problemas se ofrecen argumentos independientes a partir de los cuales formularemos ciertas condiciones que sería razonable imponer a toda teoría de la causación que pretendiera formularse en términos físicos, a la vez que evaluaremos las consecuencias de la aceptación o no de estas condiciones. Iniciamos este estudio en el siguiente capítulo, abordando el problema de las leyes de conservación en la relatividad general.

Capítulo 5. *Física de la causación: el problema de las leyes de conservación en la RG*

Resumen

En el capítulo anterior hemos analizado las principales críticas que se han formulado a la TCC. En particular, hemos examinado en detalle aquella crítica que afirma que la teoría de la causación de Phil Dowe no sería compatible con los escenarios físicos que propone la teoría general de la relatividad (RG), pues no sería del todo claro que en el marco de esta teoría física puedan ser formuladas leyes de conservación de la energía-momentum correctamente definidas. En adición a lo anterior, al finalizar el capítulo precedente hemos afirmado que un examen detallado de esta última crítica, así como también de las respuestas de Dowe a ella, nos permitirá llevar a cabo un estudio pormenorizado de los alcances de la TCC en contextos relativistas generales. Este estudio se realiza desde la mirada de tres debates actuales tanto en fundamentos de la RG como en metafísica. En este capítulo iniciamos este estudio analizando el primero de estos debates: *el problema de las leyes de conservación*.

El problema de las leyes de conservación en la RG ha sido (y sigue siendo) un verdadero desafío tanto para las y los físicos como para las y los filósofos desde los mismos inicios de esta teoría. Prueba de ello son el debate que en la primera mitad de siglo XX mantuvieron Emmy Noether, Félix Klein, Albert Einstein y David Hilbert en relación a la importancia de las ecuaciones de conservación de la energía en la formulación de la RG, así como el actual debate en filosofía de la física o fundamentos de la RG acerca de la interpretación más adecuada para este problema (ver, por ejemplo, Hofer, 2000; Pitts, 2010; Lam, 2011; Dürr, 2019a, 2019b; Read, 2020). En la bibliografía especializada acerca de este tópico se discute tanto la validez de las ecuaciones de conservación en sus formas diferencial e integral (o local y global, respectivamente), por un lado; así como también la validez de las ecuaciones de conservación de la energía-momentum de los campos de materia y de la energía-momentum total (energía-momentum de la materia más energía-momentum del campo gravitacional), por otro lado. Dado el interés de esta investigación,

nos enfocamos en la discusión en torno a la autenticidad de la ecuación integral de conservación de la energía-momentum total.

Esta ecuación integral (EI) presenta la siguiente forma:

$$\int_{S=\partial V} t^{\mu\nu} K_{\mu} n_{\nu} dS = - \int_{S=\partial V} T^{\mu\nu} K_{\mu} n_{\nu} dS = 0 \quad (\text{EI})$$

Para Carl Hofer (2000), esta ecuación de conservación no es una expresión válida, entre otras razones, debido a que no sería aplicable al mundo actual. Detrás de la afirmación de Hofer se encuentra el hecho de que (EI) se deriva a partir de ciertas restricciones relacionadas con la planitud asintótica, característica que nuestro mundo actual no parece presentar. Vincent Lam (2011), por otro lado, sostiene que esta expresión integral de conservación solo se encontraría bien definida cuando ciertas estructuras de fondo –el campo vectorial de Killing– se encuentran presentes. Por su parte, James Read (2020) argumenta que (EI) sí se encuentra bien definida en ciertos modelos dinámicamente posibles (MDP) de la RG (aquellos relacionados con la planitud asintótica). En contraposición a Hofer, Read afirma que, aun cuando el mundo actual no sea asintóticamente Minkowskiano, la expresión anterior se aplica al mundo actual. De hecho, se aplica para el cálculo de energía relacionada con las ondas gravitacionales. En este sentido, Read añade que toda teoría física es, en algún sentido, una idealización y, por lo tanto, estas no se aplican al mundo actual en un sentido fuerte. Por último, Patrick Dürr (2019b) argumenta que los modelos asintóticamente planos son idealizaciones y de ellas, eventualmente, solo podemos predicar un realismo selectivo hacia alguna de sus estructuras, no siendo ninguna de ellas merecedora de un compromiso ontológico serio. Por lo tanto, concluye Dürr, parece más adecuado hablar de un eliminativismo –en lugar de un realismo selectivo– acerca de la existencia de una ley de conservación integral para la energía-momentum gravitacional.

A partir de los argumentos de Dürr, defenderemos la existencia de una ecuación integral de conservación para la energía-momentum total correctamente definida en la RG. Aun cuando los modelos asintóticamente planos sean modelos idealizados, es posible defender que estos tienen cierta conexión con lo real. De hecho, las teorías científicas son, bajo la visión semántica, un conjunto de modelos que utilizan diferentes tipos de

representaciones científicas, como aproximaciones o idealizaciones. Desde este punto de vista, ninguno de los modelos de la teoría tendrá una correspondencia *exacta* con el mundo actual, como pretende Hofer. Algunos de estos modelos representarán con mayor precisión nuestro mundo que otros, pero ninguno de ellos lo hará con total precisión. Por lo tanto, que (EI) solo se encuentre bien definida en algunos MDP's de la teoría que no son demasiado similares al mundo actual no implica que no exista una ley integral de conservación válida en la RG. Nuestros argumentos se basarán, en consecuencia, tanto en una visión semántica de las teorías científicas como en una visión posibilista de las teorías físicas.

Este debate en torno a los fundamentos de la RG, así como la argumentación que ofreceremos, nos llevará a postular la siguiente condición que debe ser evaluada por cualquier teoría de la causación física que pretenda tener cierto alcance en la teoría de la gravitación de Einstein:

Cond1: En las teorías físicas del espacio-tiempo, en particular en la teoría general de la relatividad, las leyes de conservación integrales se cumplen en una cierta clase de sus modelos.

¿Cuáles son las consecuencias de ello para la TCC? Una respuesta concluyente a esta pregunta tendrá que esperar hasta el Capítulo 8. Sin embargo, adelantamos que esta dependerá de si las proposiciones *CC1* y *CC2* de Dowe son vistas como caracterizaciones o como definiciones de los términos involucrados en ella.

5.1. Introducción

Al finalizar el capítulo anterior hemos anunciado que el estudio del alcance y de las limitaciones de la TCC en la RG –recordemos, el problema general que aborda esta investigación– será llevado a cabo siguiendo tres vías de análisis que nos permitirán fijar tres condiciones que usaremos como referencia en este estudio. Estas vías de análisis fueron presentadas en forma de problemas: *el problema de las leyes de conservación*, *el problema de la posibilidad física* y *el problema de la definición*. En este capítulo abordaremos el primero de ellos.

El problema de las leyes de conservación ha sido ampliamente debatido en la bibliografía especializada en filosofía de la física. Algunos autores argumentan que no existen leyes de conservación en la RG en ningún sentido (Hoeyer, 2000; Dürr, 2019a, 2019b), mientras que otros afirman que sí es posible defender la existencia de este tipo de leyes en el marco de la teoría de Einstein (Lam, 2011; Read, 2020)⁶⁷. Este debate tiene implicancias directas para la TCC: la ausencia de ecuaciones de conservación integrales en contextos relativistas generales implica, a su vez, la ausencia de cantidades físicas que merezcan la categoría de “conservada”, siendo este tipo de cantidades fundamentales en la formulación de la teoría de la causación de Dowe. En efecto, recordemos que las proposiciones *CCI* y *CC2* de la TCC se fundamentan en la noción de “cantidad conservada”, a través de expresiones como “poseer una cantidad conservada” o “intercambiar una cantidad conservada”.

Considerando el problema anterior, en este capítulo evaluaremos en qué sentido estamos autorizados a hablar de la existencia o de la inexistencia de leyes de conservación en los contextos relativistas generales. En particular, investigaremos la plausibilidad de las leyes integrales de conservación, las relevantes para la TCC. Esbozaremos algunos argumentos en defensa del punto de vista que sostiene que las leyes integrales de conservación se encuentran correctamente definidas en algunos modelos de la RG; estos argumentos descansan en una visión semántica de las teorías científicas, así como en una visión posibilista de las teorías físicas. Como una consecuencia de esta discusión, fijaremos una primera condición (que llamaremos *Cond1*) que impone el debate actual sobre fundamentos de la RG a cualquier teoría de la causación que pretenda formularse en términos de cantidades físicas conservadas. Usando esta condición como referencia, se evaluará el alcance de la TCC en la RG.

Este capítulo se encuentra estructurado del siguiente modo. Después de un repaso necesario de algunas nociones básicas de RG en la Sección 5.2, en la Sección 5.3 presentaremos el problema de las leyes de conservación desde un punto de vista técnico. A continuación, en la Sección 5.4, examinaremos el debate de las leyes de conservación desde un punto de vista filosófico, presentando los principales puntos de vista al respecto. Por

⁶⁷ Respecto del trabajo de Read (2020), es importante mencionar que existe una versión previa del año 2018.

último, en la Sección 5.5, esbozaremos algunos argumentos para intentar defender la existencia de leyes integrales de conservación en contextos relativistas generales, presentaremos la primera condición y adelantaremos algunas consecuencias para la TCC.

5.2. Nociones básicas de RG

La teoría Newtoniana de la gravitación puede expresarse por medio de dos ecuaciones básicas. Una de ellas describe el campo gravitacional dependiente de la materia y la otra establece qué le sucede a la materia cuando se expone a un campo de este tipo. La primera de estas ecuaciones, la del campo gravitacional, es la conocida ecuación de Poisson, expresada en términos del potencial gravitacional ϕ :

$$\nabla^2\phi = 4\pi G\rho \quad (5.1)$$

donde ρ es la densidad de masa y G la constante de gravitación universal Newtoniana. La segunda de estas ecuaciones relaciona la aceleración con el gradiente del potencial gravitacional:

$$\mathbf{a} = \nabla\phi \quad (5.2)$$

Hasta antes del maravilloso trabajo de James Clerk Maxwell (1865), que unificó las ecuaciones que representaban al campo magnético y al campo eléctrico en un conjunto de cuatro ecuaciones representativas del campo electromagnético, la teoría Newtoniana de la gravitación gozaba de gran prestigio, ya que lograba explicar con gran precisión los fenómenos hasta ese entonces conocidos. Sin embargo, algunas dificultades comenzaron a aparecer para la teoría de Newton. En primer lugar, se descubrieron algunas discrepancias entre las predicciones de la teoría y algunas observaciones, como sucedió, por ejemplo, con la aceleración del movimiento medio de la Luna y el avance del perihelio de Mercurio. Sin embargo, de acuerdo con Roberto Torretti (1983), las discrepancias recién mencionadas no fueron un motivo suficiente para abandonar la teoría, sino que se supuso –en la comunidad de científicos– más bien que estos hechos podrían ser resueltos considerando algunas modificaciones a la teoría de Newton. Torretti sostiene que una dificultad mayor para la gravitación Newtoniana y que obligó a su revisión y posterior reemplazo surgió de la

incomodidad de físicos y filósofos por la acción a distancia inherente a la teoría de Newton (1983). Este problema se hizo evidente con uno de los resultados más sorprendentes de las ecuaciones de Maxwell: la velocidad de la luz en el vacío tiene un valor constante y es igual para todos los observadores. Este hecho, por supuesto, implica una velocidad límite para la transmisión de señales. Con el advenimiento de la relatividad especial –que incluye como uno de sus postulados la constancia de la velocidad de la luz– los problemas para la gravitación Newtoniana se acentuaron y comenzó a tomar fuerza la búsqueda de una nueva teoría de la gravitación.

Antes de revisar cómo se inició esta búsqueda, se hace preciso puntualizar de qué manera la teoría gravitacional de Newton permite acciones a distancia e instantáneas, resultando, de este modo, incompatible con la teoría especial de la relatividad (en adelante, RE). Una primera manera de detectar esta incompatibilidad es percatándose que el potencial gravitacional y la aceleración de un cuerpo en la teoría de Newton, representados por las ecuaciones (5.1) y (5.2), no se transforman adecuadamente bajo las transformaciones de Lorentz. Estas transformaciones de coordenadas son las usadas en la RE, y reflejan la constancia de la velocidad de la luz para todos los observadores inerciales. La segunda forma es analizando la ecuación (5.1) para el potencial gravitatorio. Como se hace evidente, el potencial gravitacional ϕ depende de la densidad de materia del universo ρ , lo cual implica que, si se modifica la distribución de materia en algún punto, el efecto sobre el potencial gravitacional será instantáneo. Esto último, además, implica la idea de simultaneidad absoluta que se cumple en la teoría Newtoniana, pero no en la RE.

Volvamos ahora nuestra atención hacia la búsqueda de Einstein de una teoría de la gravitación⁶⁸. Se podría pensar que el camino que seguiría Einstein sería modificar la teoría de la gravitación existente para hacerla compatible con la teoría relativista que él mismo había desarrollado. Sin embargo, el camino trazado por Einstein fue, en algún sentido, contrario a la propia RE, pues el resultado al que arribaría, la RG, mantiene como válida la primera solo como una aproximación local (se verán detalles sobre este punto más abajo). El puntapié inicial para la RG de Einstein fue la, en ese entonces ya conocida, equivalencia

⁶⁸ Es interesante mencionar aquí que Alfred North Whitehead (1922) postuló una teoría alternativa de la relatividad que mantenía la división entre física y geometría, división que justamente Einstein disuelve con su teoría de la gravitación.

entre masa inercial y masa gravitacional, fundada en la universalidad de la caída libre. Einstein llevó hasta las últimas consecuencias esta equivalencia, lo cual le permitió sentar las bases para una nueva teoría de la gravitación. Examinemos, a continuación, el curso de las ideas seguidas por Einstein en la formulación de su teoría de la gravitación.

La equivalencia entre masa inercial y masa gravitacional fue expresada por Einstein en la siguiente forma⁶⁹. Supónganse dos laboratorios S_1 y S_2 . S_1 es un laboratorio uniformemente acelerado en una determinada dirección; por otra parte, S_2 es un laboratorio en reposo que se encuentra inmerso en un campo gravitacional homogéneo que dota a los objetos de este laboratorio con una aceleración igual en magnitud a la del laboratorio S_1 . Como es bien conocido y probado experimentalmente, las leyes físicas de los laboratorios S_1 y S_2 no difieren en lo absoluto, lo que es debido, como se mencionó más arriba, a la universalidad de la caída libre. Por lo tanto, existe una equivalencia física entre lo ocurrido en un laboratorio uniformemente acelerado y lo ocurrido en un laboratorio en reposo en un campo gravitacional homogéneo. Esto supone, de acuerdo con Einstein, una extensión del Principio de la Relatividad –principio enunciado en su teoría de la RE– a marcos uniformemente acelerados⁷⁰. Esta igualdad física entre marcos acelerados y marcos en reposo en un campo gravitacional homogéneo es conocido como el *Principio de Equivalencia Débil*. El adjetivo “débil” hace referencia a que este principio expresa el comportamiento de los cuerpos que se encuentran sometidos únicamente a la gravedad. Sin embargo, Einstein da un paso más y extrapola este principio a cualquier tipo de fenómeno, no solo los gravitatorios. De tal manera, un marco uniformemente acelerado será completamente equivalente a un marco en reposo sujeto a un campo gravitacional homogéneo. Ningún experimento físico permitirá distinguir un marco del primer tipo de

⁶⁹ La forma en que es expresado aquí el principio de equivalencia corresponde a la versión postulada por Einstein en su artículo de 1907.

⁷⁰ De acuerdo con Torreti (1983), es cuestionable que la igualdad entre un marco uniformemente acelerado y otro en reposo en un campo gravitacional uniforme sea considerado como una extensión del principio de la relatividad. El autor sostiene que:

“...su presentación [del Principio de Equivalencia] como una generalización del principio de la relatividad es completamente engañosa. La equivalencia física entre marcos de referencia en reposo en un campo gravitacional homogéneo y los marcos uniformemente acelerados no elimina la desigualdad física entre estos últimos y los marcos inerciales. Simplemente implica que un marco de referencia en reposo en un campo gravitacional no puede ser inercial” (p. 135).

otro marco del segundo tipo. Este último es conocido como *Principio de Equivalencia Fuerte*⁷¹ (ver Schutz, 2009, p. 115).

Más arriba se ha dicho que, para arribar a su teoría de la gravitación, Einstein eligió un camino que se insinuaba más bien en contra de la RE. La razón es la siguiente. Tal como afirma Torretti (1994, Cap. 2), Einstein buscaba eliminar los marcos de referencia privilegiados, esto es, los marcos de referencia inerciales, dejando de esta manera fuera de acción a la RE, que justamente es válida en marcos de referencia de este tipo. La manera de ver con mayor claridad cómo Einstein elimina los marcos privilegiados es, como sostiene Torretti (1994), invirtiendo la formulación de Einstein del Principio de Equivalencia:

“Si, como él dice [Einstein], un marco en reposo en un campo gravitacional uniforme equivale a un marco exento de gravitación pero uniformemente acelerado respecto de los marcos inerciales, entonces *un marco inercial equivale*, a su vez, a un marco *en caída libre en un campo gravitacional uniforme*” (p. 63, énfasis en el original).

Además, como es experimentado a diario por los cuerpos masivos, la gravedad se encuentra en todas partes y es imposible no verse afectado por ella; luego los marcos inerciales solo pueden existir en la forma de marcos en caída libre en un campo gravitacional. Sin embargo, el Principio de Equivalencia solo habla del *campo gravitacional uniforme*; pero estos solo existen como una aproximación local. En general, el campo gravitacional es variable, ya que depende de la distribución de materia que lo produce. Se concluye, pues, que no existe una idea de marco inercial global (únicamente existirían como una aproximación local) (ver Torretti, 1994, pp. 64–65). Muy probablemente, el afán de Einstein por eliminar los marcos de referencia inerciales se debió a la influencia de Ernst Mach, quien no encontraba un argumento razonable para elegir un marco de referencia privilegiado. De acuerdo con Mach, las fuerzas inerciales⁷² (básicamente, la fuerza de Coriolis y la fuerza centrífuga) se deben a la acción de las

⁷¹ Para una reconstrucción y discusión histórica acerca del principio de equivalencia se recomienda fuertemente ver Torretti (1983, Cap. 5; 1994, Cap. 3).

⁷² Fuerzas que la mecánica clásica trata como aparentes y que se manifiestan al describir un sistema físico cuando se usa un sistema de referencia no inercial. Y que justamente Einstein identifica con la gravedad a través del Principio de Equivalencia.

fuentes materiales (lejanas) del universo. De hecho, Einstein incluye dentro de los principios⁷³ de su teoría de la gravitación clásica las ideas de Mach a través de lo que denomina *Principio de Mach*.

Luego de haber revisado, en los párrafos anteriores, algunos detalles sobre los inicios de la RG de Einstein, en las líneas que siguen se hará referencia a las principales características de la teoría de la gravitación clásica de Einstein, que serán de gran utilidad en las secciones siguientes.

Se inicia el análisis con algunas definiciones matemáticas importantes que serán de utilidad para comprender el significado de la gravitación Einsteniana. La mayoría de estas definiciones pertenece a la rama de las matemáticas conocida como geometría diferencial. En razón de la complejidad inherente a esta disciplina, en la exposición que sigue se sacrificará la rigurosidad matemática para intentar asegurar, por el contrario, una adecuada y precisa exposición física y conceptual de los términos involucrados. Una presentación de estos conceptos con todo su rigor matemático escapa no solo a los objetivos de este trabajo, sino que al espacio con el que se cuenta en este capítulo. No obstante, para el lector que quiera profundizar en las ideas expuestas, se brindarán las referencias pertinentes en todo momento.

Quizás el concepto más fundamental en el contexto de la RG es el de *variedad* (en inglés, *manifold*)⁷⁴. Informalmente se dice que una variedad diferenciable M es un espacio (conjunto de puntos), formado por “parches” que en regiones locales de la misma “se parece” a R^n (el espacio euclideo de la mecánica Newtoniana o el espacio minkowskiano de la RE), y cuyos parches están unidos suavemente. Se definirá con claridad el concepto de “diferenciable” un poco más abajo, pero antes se hace necesario definir algunos conceptos adicionales.

La definición ofrecida muestra una imagen de la variedad como un “conjunto de parches” donde cada uno de ellos es mapeado a R^n . Revisemos, entonces, la definición de

⁷³ Para un interesante análisis acerca de los principios en los que se fundamenta la RG, véase Torretti (1994, Cap. 4).

⁷⁴ La exposición que se ofrece de ahora en adelante, y hasta el final de esta sección, sigue los excelentes trabajos de Wald (1984), Carroll (2004) y Schutz (2009).

mapa. En general, un mapa entre dos conjuntos A y B , $\phi: A \rightarrow B$, es una relación que asigna a cada elemento del conjunto A un elemento del conjunto B ⁷⁵. Es decir, un mapa es una generalización de la idea de función. En particular, para el caso de las variedades, un mapa asigna a cada elemento de un subconjunto de la variedad (un parche) un elemento en R^n . Ahora que ya se cuenta con la noción de mapa, es posible introducir la idea de *continuidad y diferenciabilidad*. La idea es muy simple: ya que un mapa puede ser pensado como una generalización de la idea de función, entonces es razonable exigir que un mapa sea continuo y diferenciable en el mismo sentido que se les exige a las funciones. Así, un mapa de clase C^∞ es continuo e infinitamente diferenciable. En consecuencia, una variedad diferenciable de clase C^∞ es una variedad cuyos mapas son continuos e infinitamente diferenciables. Por otra parte, dos conjuntos A y B serán llamados difeomórficos si existe un mapa $\phi: A \rightarrow B$ de clase C^∞ y su mapa inverso $\phi^{-1}: B \rightarrow A$ entre estos dos conjuntos. En tal caso, el mapa es llamado un *difeomorfismo*. Este es el caso de la RG, donde los mapas son difeomorfismos.

Una *carta o sistema de coordenadas* es un subconjunto U de la variedad M junto con un mapa uno a uno $\phi: U \rightarrow R^n$, tal que la imagen $\phi(U)$ es un conjunto abierto en R^n (ver Carroll, 2004, p. 59) (ver Figura 5.1). Por último, se define el concepto de atlas. Un *atlas* es básicamente un conjunto de cartas que satisfacen dos condiciones: (i) la unión de las cartas coincide con la variedad, esto es, las cartas cubren la variedad M , y (ii) las cartas están unidas suavemente, es decir, si dos cartas que representan dos subconjuntos de la variedad se superponen, en la región de la superposición debe existir un sistema de coordenadas que permita transformar de una carta a otra (para detalles técnicos véase Carroll, 2004, p. 59; Wald, 1984, p. 13). Esta es la noción un poco más precisa de “parches unidos suavemente” mencionada en la definición de variedad más arriba.

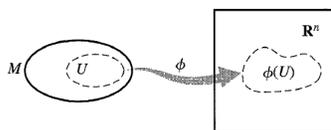


Figura 5.1 (Carroll, 2004, p. 59)

⁷⁵ Por supuesto, los mapas pueden ser clasificados en diferentes categorías o cumplir con ciertas propiedades. Véase Wald (1984, p. 12) y Carroll (2004, pp. 57–58).

Dadas las definiciones anteriores, una definición más precisa de variedad que la ofrecida anteriormente, es la siguiente: una *variedad diferenciable n-dimensional*⁷⁶ de clase C^∞ es un conjunto que puede ser representado mediante un atlas maximal⁷⁷.

Una vez que se cuenta con la definición de variedad, el siguiente paso es equiparla con una serie de estructuras que permitan el objetivo último de esta construcción matemática: realizar operaciones en espacios curvos, como, por ejemplo, operaciones con vectores y tensores, diferenciación o integración. Una de las estructuras más importantes con que se dota a una variedad es la de campo métrico. El *campo métrico* $g_{\mu\nu}$ es un tensor simétrico de tipo $(0, 2)$ ⁷⁸ que provee a la variedad de las herramientas necesarias para establecer medidas de distancias, ángulos y volúmenes en cada espacio tangente a ella (en consecuencia, el campo métrico actúa en cada punto de la variedad). Adicionalmente, el campo métrico permite definir relaciones causales en la variedad, más específicamente, permite discriminar cuáles puntos de la variedad podrían estar conectados mediante una señal. El campo métrico $g_{\mu\nu}$ es análogo a la métrica de Minkowski $\eta_{\mu\nu}$ usada en la RE.

La *signatura* de la métrica se define como el número de autovalores positivos y negativos que esta contiene (por ejemplo, la métrica de Minkowski tiene signatura $(-, +, +, +)$). Por otra parte, si alguno de los autovalores de la métrica es cero, entonces decimos que la métrica es “degenerada”. En cambio, si la métrica es continua y “no-degenerada”, su signatura será la misma en cada punto de la variedad. En la RE y en la RG se utilizan métricas continuas y no-degeneradas. La signatura de la métrica permite clasificar las variedades. Si la métrica es definida positiva (todos sus signos son positivos), entonces la variedad es llamada *euclídea o riemanniana* (ver Carroll, 2004, p. 73; Schutz, 2009, p. 144). Por otro lado, si la signatura posee un signo menos, la métrica se denomina *lorentziana o pseudo-riemanniana*. Tanto en la RE como en la RG se usan métricas pseudo-riemannianas. En particular, una de las premisas principales de la RG es que el

⁷⁶ La dimensionalidad n de la variedad debe coincidir con el conjunto R^n al que está siendo mapeada. Es una de las exigencias para declarar que la variedad se “parece” localmente a R^n .

⁷⁷ Un atlas *maximal* hace referencia a aquel atlas que contiene todas las cartas posibles compatibles con la variedad.

⁷⁸ Un tensor de tipo $(0, N)$ es una función de N vectores a números reales. Esta función es lineal para cada uno de los N argumentos (ver Schutz, 2009, p. 56).

espacio-tiempo puede ser modelado por una variedad pseudo-riemanniana, 4-dimensional, Hausdorff y conectada^{79,80}.

Continuemos revisando los conceptos de espacio tangente, vector tangente y curva. En cada punto de la variedad es posible definir un *espacio tangente* a ese punto que será isomorfo a R^n . Si se denomina p a un punto de la variedad, entonces el espacio tangente a ese punto se denota por $T_p M$. Es importante señalar que los espacios tangentes en distintos puntos de la variedad serán diferentes debido a la curvatura arbitraria de la variedad a nivel global (ver Janssen, 2013, pp. 100–101) (ver Figura 5.2). Adicionalmente, sobre la variedad es posible definir los *vectores tangentes* (o simplemente vectores) como aquellos que se encuentran sobre el espacio tangente en un punto de la variedad. Por último, para algún intervalo $I \subseteq R$, una *curva suave* (y continua) γ sobre una variedad M es un mapa de clase C^∞ de I a M , esto es, $\gamma: I \rightarrow M$. Es importante mencionar que en cada punto de la variedad que pertenezca a la curva podemos asociar un vector tangente que pertenece al espacio tangente⁸¹.

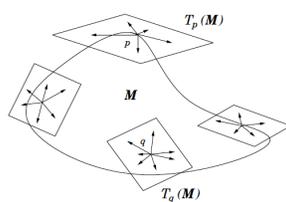


Figura 5.2 (Janssen, 2013, p. 101)

Será instructivo hacer una breve pausa para resumir las herramientas con las que se cuenta hasta ahora. Se ha definido el concepto de variedad diferenciable que, *grosso modo*, es un conjunto compuesto de “parches” o subconjuntos que pueden ser mapeados a R^n .

⁷⁹ Para una definición precisa de estos términos ver Wald (1984), apéndice A.

⁸⁰ En algunos libros de texto se menciona que la variedad usada en RG también es “paracompacta”. Sin embargo, que la variedad sea Hausdorff y conectada implica que es también paracompacta (ver Hawking y Ellis, 1973, p. 57).

⁸¹ Se ha ofrecido una idea muy intuitiva de vector y curva, además se ha pasado por alto la definición de tensor. Sin embargo, las nociones aquí ofrecidas de estos conceptos son suficientes para los objetivos de este trabajo. Si el lector desea profundizar en las definiciones matemáticamente precisas de vector, curva y tensor, se recomienda ver Wald (1984, Cap. 2) y Carroll (2004, Cap. 2).

Cada “parche” de la variedad es mapeado mediante una carta o sistema de coordenadas, y el conjunto de estas cartas define un atlas sobre esta variedad. Adicionalmente, se han definido estructuras adicionales sobre esta variedad, como el campo métrico, el espacio tangente y las nociones de vector tangente y curva, con el propósito de definir ciertas operaciones matemáticas sobre ella. Ya definido todo este aparataje matemático, surge el siguiente problema: como se ha mencionado más arriba, los espacios tangentes, definidos por el conjunto de los vectores tangente, son diferentes a medida que se recorre la variedad debido a la curvatura que esta podría presentar. Luego, en general, no es posible comparar dos vectores situados en puntos diferentes de la variedad, pues pertenecen a diferentes espacios vectoriales (o espacios tangentes). Para poder llevar a cabo esta comparación se hace necesario introducir algunas estructuras adicionales, como la derivada covariante, la conexión afín y el transporte paralelo. Estas son definidos a continuación.

La derivada parcial de un vector o un tensor no transforma adecuadamente bajo un cambio general de coordenadas. Sin embargo, lo que se necesita en la RG es un operador diferencial que transforme tensorialmente bajo variedades arbitrarias (y que, por lo tanto, sea independiente de las coordenadas elegidas) y que se reduzca a una derivada parcial en un espacio-tiempo plano como el de la RE con coordenadas inerciales. El operador que cumple estas características es la *derivada covariante* ∇ ⁸². A modo de ejemplo y con el fin de exponer algunas cuestiones conceptuales importantes, es posible demostrar que la derivada covariante de un vector contravariante V^ν en la dirección ν tiene la siguiente forma⁸³:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma_{\mu\lambda}^\nu V^\lambda \quad (5.3)$$

En la ecuación anterior se puede ver que la derivada covariante actúa sobre un vector contravariante como una derivada parcial del vector que está siendo derivado, pero además se añade un factor de corrección debido al transporte paralelo (el cual se define más

⁸² Para una definición matemáticamente rigurosa de derivada covariante, ver Carroll (2004, Cap. 3), Schutz (2009, Cap. 6) y Wald (1984, Cap. 3).

⁸³ Por supuesto, también es posible deducir la derivada covariante de un vector covariante y de un tensor. Estas ecuaciones no han sido expuestas aquí, ya que el mismo análisis conceptual que hacemos respecto a la derivada covariante de un vector contravariante es aplicable a los vectores covariantes y a los tensores.

abajo). En este factor de corrección (el segundo término del lado derecho de la igualdad) aparece el término $\Gamma_{\mu\lambda}^{\nu}$ llamado *conexión* o *símbolos de Christoffel*⁸⁴, que representa un conjunto de coeficientes que son especificados en algún sistema de referencia. En general, sobre una variedad determinada es posible establecer diferentes tipos de conexiones, cada una de las cuales brindará una definición diferente de derivada covariante. Sin embargo, en la RG, es posible fijar una conexión para una variedad con un campo métrico $g_{\mu\nu}$ exigiendo dos requisitos: (i) la conexión debe estar libre de torsión, y (ii) debe ser compatible con la métrica⁸⁵. La conexión que se deriva de estos dos requisitos es la conocida *conexión de Levi-Civita* y es el tipo de conexión usada en la RG. Como es posible observar en la siguiente ecuación, la conexión de Levi-Civita se encuentra expresada en términos de la métrica:

$$\Gamma_{\mu\nu}^{\sigma} = \frac{1}{2} g^{\sigma\rho} (\partial_{\mu} g_{\nu\rho} + \partial_{\nu} g_{\rho\mu} + \partial_{\rho} g_{\mu\nu}) \quad (5.4)$$

Esta conexión tiene interesantes propiedades que simplifican bastante los cálculos y permiten expresar en forma más sencilla las expresiones para los operadores diferenciales, como el rotacional, la divergencia y el laplaciano.

Más allá de los aspectos matemáticos, intentemos clarificar la interpretación física de la derivada covariante para el caso de la RG. De acuerdo con Sean Carroll (2004), “la derivada covariante cuantifica la razón instantánea de cambio de un campo tensorial en comparación a lo que el tensor sería si fuera ‘transportado paralelamente’” (pp. 102–103). De este modo, la derivada covariante y, por tanto la conexión de Levi-Civita para el caso de la RG, definen una manera específica de mantener un tensor constante a lo largo de un trayectoria y sobre esta base poder comparar tensores cercanos (ver Carroll, 2004, p. 103). Se hace evidente, por la interpretación física de la derivada covariante recién ofrecida, la necesidad de una definición de *transporte paralelo*⁸⁶. Cuando se comparan vectores en el

⁸⁴ En general, los símbolos de Christoffel representan los coeficientes de ponderación de los vectores base que son usados para representar el vector resultante de la derivación de un vector base. Para más detalle acerca de los símbolos de Christoffel, ver Fleisch (2012, Cap. 5)

⁸⁵ Ver Carroll (2004, p. 99) para las formulaciones matemáticas de estos requisitos y las implicaciones de los mismos.

⁸⁶ Para una descripción matemáticamente rigurosa de transporte paralelo, véase Wald (1984, p. 30), Carroll (2004, p. 102) y Schutz (2009, p. 153).

espacio plano muchas veces se necesita mover un vector de un punto a otro manteniéndolo constante todo el tiempo. Por supuesto, en el espacio plano, esto no conlleva ninguna dificultad incluso si el transporte del vector se realiza por trayectorias diferentes. Sin embargo, en un espacio curvo esta operación no es tan sencilla. Transportar un vector en un espacio curvo y que este se mantenga constante dependerá de la trayectoria seguida. Un ejemplo muy recurrente en los libros de texto para poner de manifiesto esta característica de los espacios curvos es el de la esfera. En una esfera, sobre la cual se desea transportar un vector paralelamente a través de dos diferentes trayectorias desde el ecuador hasta uno de sus polos, se arribará a dos diferentes resultados (ver Figura 5.3).

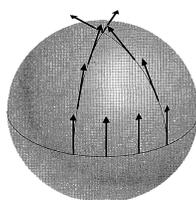


Figura 5.3 (Carroll, 2004, p. 104)

El hecho anterior trae como consecuencia que no sea posible, de una manera natural como se haría en un espacio-tiempo plano, comparar dos vectores (o tensores) que se encuentran en espacios tangentes distintos. Como se ha hecho notar más arriba, la manera adecuada de comparar vectores (o tensores) cercanos es a través de la derivada covariante y su conexión asociada.

Muy asociada a los conceptos de derivada covariante, conexión y transporte paralelo se encuentra la idea de *curvatura*, la cual es crucial en la teoría de la gravitación clásica de Einstein. La noción de curvatura se asocia a la noción de transporte paralelo de la siguiente forma: una variedad presentará curvatura intrínseca si al transportar paralelamente un vector de un punto a otro a través de diferentes trayectorias (formando un circuito cerrado) obtenemos vectores distintos. Así, en un espacio-tiempo plano, el transporte paralelo de un vector llevado a cabo por diferentes trayectorias siempre resultará en el mismo vector original. En la RG, la estructura matemática encargada de medir esta curvatura es el *tensor de Riemann* $R_{\sigma\mu\nu}^{\rho}$, definido por:

$$R_{\sigma\mu\nu}^{\rho} = \partial_{\mu}\Gamma_{\nu\sigma}^{\rho} - \partial_{\nu}\Gamma_{\mu\sigma}^{\rho} + \Gamma_{\mu\lambda}^{\rho}\Gamma_{\nu\sigma}^{\lambda} + \Gamma_{\nu\lambda}^{\rho}\Gamma_{\mu\sigma}^{\lambda} \quad (5.5)$$

Nótese que el tensor de Riemann depende de la conexión Γ y, por lo tanto, este tensor será nulo en un espacio-tiempo plano. Adicionalmente, por su utilidad en las ecuaciones de Einstein, se define el *tensor de Ricci* $R_{\mu\nu}$ como la contracción del tensor de Riemann, y el *escalar de Ricci* R , como la traza del tensor de Ricci. El escalar de Ricci se calcula por:

$$R = g^{\mu\nu}R_{\mu\nu} \quad (5.6)$$

Además, se define el *tensor de Einstein* $G_{\mu\nu}$ como:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \quad (5.7)$$

Por otra parte, es preciso hacer mención del concepto de *geodésica*. Bastará con decir que una geodésica se puede pensar intuitivamente como la distancia más corta entre dos puntos en un espacio curvo. En este sentido, las geodésicas son una generalización del concepto de línea recta en el espacio plano. Sin embargo, en el estudio de la RG se usa una definición equivalente de geodésica⁸⁷, como aquella curva cuyo su vector tangente es transportado paralelamente a sí mismo a lo largo de ella. Es posible demostrar (ver, por ejemplo, Carroll, 2004, p. 106) que la ecuación de las geodésicas⁸⁸ está dada por:

$$\frac{d^2x^{\mu}}{d\lambda^2} + \Gamma_{\rho\sigma}^{\mu} \frac{dx^{\rho}}{d\lambda} \frac{dx^{\sigma}}{d\lambda} = 0 \quad (5.8)$$

⁸⁷ Estas definiciones son equivalentes si la conexión que usamos es la conexión de Levi-Civita, como efectivamente sucede en la RG. Para otro tipo de conexiones, estas dos caracterizaciones de geodésica no son equivalentes.

⁸⁸ Un tema de investigación muy interesante se encuentra asociado a si la ecuación de la geodésica, considerada como la ecuación del movimiento en la RG, se deduce de las ecuaciones de Einstein o si debe ser postulada independientemente. Para una discusión histórica al respecto, véase Torretti (1983, p. 176). Para una discusión más reciente, véase, por ejemplo, Malament (2012) o Weatherall (2018).

Se puede ver fácilmente en qué sentido la ecuación anterior es una generalización de la noción de línea recta en el espacio plano. Si sobre un espacio-tiempo plano se eligen coordenadas cartesianas, donde los símbolos de Christoffel son iguales a cero, el segundo término del lado izquierdo de la ecuación desaparece, quedando únicamente el primer término del lado izquierdo que es justamente la ecuación de la recta, la cual representa la distancia más corta (geodésica) en un espacio-tiempo de este tipo. Una importante propiedad de las geodésicas es que son las trayectorias que siguen las partículas (de prueba⁸⁹) sin aceleración en los espacio-tiempos relativistas generales.

Para finalizar con la noción de geodésica, consideremos brevemente el concepto de *parámetro afín*. Para derivar la ecuación de las geodésicas es necesario parametrizar estas curvas mediante algún parámetro adecuado. El parámetro que cumpla la condición de permitir la derivación de la ecuación de las geodésicas (ecuación (5.8)) es denominado parámetro afín. En la RG, el parámetro afín es cualquier parámetro relacionado con el tiempo propio τ .

Llegados a este punto, con toda la estructura matemática con la que se cuenta, se está en condiciones de presentar las ecuaciones de campo de Einstein. Se ha visto, al inicio de esta sección, que la teoría Newtoniana de la gravitación, expresada a través de las ecuaciones (5.1) y (5.2), puede ser interpretada como una teoría que afirma, por un lado, cómo la materia actuará en presencia de una fuerza gravitacional (la ecuación (5.2)) y, por otro lado, cómo la materia determina un campo gravitacional (la ecuación (5.1)). Podemos pensar el caso relativista general de manera análoga al caso Newtoniano, considerando, por supuesto, la gravedad ya no como una fuerza, sino como una manifestación de la curvatura del espacio-tiempo. De este modo, para el movimiento de la materia en presencia de un campo gravitacional se tiene, por una parte, la ecuación de las geodésicas (5.8) que se ha analizado más arriba y, por otra parte, las ecuaciones de Einstein (las que pueden ser interpretadas como análogas a la ecuación (5.1)) que dictan cómo la materia y la energía influyen sobre el espacio-tiempo para crear curvatura y cómo la curvatura del espacio-

⁸⁹ Una partícula de prueba es un cuerpo que no influye sobre la geometría del espacio-tiempo y, por lo tanto, no influye sobre sí mismo cuando se encuentra en movimiento. En este sentido, es un cuerpo ideal, pero una buena aproximación para el estudio de los fenómenos relativistas (ver Carroll, 2004, p. 108).

tiempo actúa sobre la materia-energía, la que interpretamos como gravedad. Las ecuaciones de campo de Einstein se expresan mediante la ecuación⁹⁰:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi GT_{\mu\nu} \quad (5.9)$$

Como se ve en la ecuación anterior, en el lado izquierdo tenemos la estructura geométrica que se ha analizado en las páginas anteriores. Se tiene el tensor de Ricci $R_{\mu\nu}$ que permite medir la curvatura intrínseca del espacio-tiempo y el campo métrico $g_{\mu\nu}$ que permite establecer “medidas” entre puntos en el espacio-tiempo. En definitiva, en el lado izquierdo de la ecuación (5.9) se encuentra codificada la geometría o curvatura del espacio-tiempo. Por otro lado, en el lado derecho de la ecuación (5.9) se encuentra el tensor de energía-momentum $T_{\mu\nu}$, que representa la distribución de materia-energía en el universo. Dado que una ecuación es una igualdad, el tensor de energía-momentum determina la geometría del espacio-tiempo de la misma forma que la curvatura determina el movimiento de la materia-energía.

Una cuestión que puede parecer evidente, pero que vale la pena resaltar, es que la teoría de la gravitación de Einstein es una teoría de campos. Visualizar la teoría Einsteiniana como una teoría de campos permite desvincular a la gravedad del concepto de acción a distancia, vínculo tan arraigado en la teoría Newtoniana como se ha visto al inicio de esta sección. Las interacciones entre puntos lejanos, en la teoría de Einstein, se llevan a cabo precisamente a través del campo que inunda el espacio y, por lo tanto, las señales necesitan de un tiempo para viajar de un punto a otro.

Las ecuaciones de campo Einstein (5.9) necesitan ser complementadas con una ecuación de conservación, también denominada ecuación de movimiento. Esta ecuación se obtiene a través de la aplicación de un principio denominado *principio de acoplamiento mínimo*, que no es más que una interpretación del Principio de Equivalencia de Einstein. El principio de acoplamiento mínimo sostiene que las leyes de la física expresadas en forma

⁹⁰ Como es sabido, en muchas ocasiones las ecuaciones de Einstein incluyen un término adicional que representa la conocida “constante cosmológica”, añadida por Einstein a sus ecuaciones cuando este se interesó en los problemas cosmológicos. Para profundizar en este asunto, véase Carroll (2004, p. 171).

tensorial en un espacio-tiempo plano (el de la RE) tienen la misma forma en un espacio-tiempo curvo (el de la RG). Es decir, para expresar las leyes físicas válidas en la RE en espacio-tiempos curvos, solo se debe sustituir, en tales leyes, las derivadas parciales por derivadas covariantes. De esta forma, si aplicamos este principio a la ley de conservación de la energía-momentum de la RE:

$$\partial_{\mu} T^{\mu\nu} = 0 \quad (5.10)$$

se obtiene una ley de conservación de la energía-momentum que, bajo cierta interpretación, puede ser considerada como una ecuación de conservación válida en la RG:

$$\nabla_{\mu} T^{\mu\nu} = 0 \quad (5.11)$$

En las siguientes secciones discutiremos largamente sobre esta y otras ecuaciones de conservación.

Antes de finalizar esta sección y para comprender el lugar que ocupa la RE en el contexto de esta discusión, analizamos brevemente a continuación el concepto de *marco inercial local*. Se ha discutido, al inicio de esta sección, el principio de equivalencia y las consecuencias que este principio conlleva. Se ha hecho notar que una de las principales consecuencias de este principio es que es prácticamente imposible construir un marco inercial global (si el campo gravitacional no es homogéneo como generalmente sucede). Sin embargo, sí es posible rescatar la noción de marco inercial local. Aun cuando las inhomogeneidades del campo gravitacional estén presentes, es posible definir, en regiones pequeñas del espacio-tiempo, un marco inercial. En efecto, la forma en que fue definida la variedad, como aquella que localmente se parece al espacio-tiempo de Minkowski, asegura que siempre es posible definir un marco inercial local que permite eliminar el campo gravitatorio y en el que, por lo tanto, serán válidas las leyes de la RE. Ahora bien, es importante destacar que la RE se rescata solo “aproximadamente” en una pequeña región del espacio-tiempo, pues la diferencia entre la métrica $g_{\mu\nu}$ de la variedad y la métrica $\eta_{\mu\nu}$ del espacio tangente, en esta pequeña región, es de segundo orden (solo las primeras derivadas se anulan).

Ahora que ya hemos revisado algunos de los principales conceptos asociados a la teoría de la relatividad de Einstein, estamos en condiciones de avanzar hacia el análisis de uno de los tópicos más importantes en el campo de los fundamentos de la RG: el problema de las leyes de conservación.

5.3. El problema de las leyes de conservación en la RG

El problema de las leyes de conservación es un tema de la teoría de la gravitación de Einstein que ha sido objeto de debate e investigación desde los inicios mismos de esta teoría física. Actualmente no existe un consenso entre físicos y entre filósofos de la física acerca de cómo este problema podría encontrar una solución, por lo que las investigaciones en este campo permanecen activas. El problema, en términos generales, puede ser caracterizado de la siguiente manera: en los espacio-tiempos relativistas generales, a diferencia de lo que sucede, por ejemplo, en la RE y en la mecánica newtoniana, existen ciertas dificultades técnicas para definir genuinas leyes de conservación tanto diferenciales como integrales. En esta sección, presentamos el problema de las leyes de conservación de una manera estándar, es decir, tal como este se enuncia en los libros de texto y artículos académicos acerca del tema, tratando de dejar de lado cualquier interpretación al respecto. Este último asunto será objeto de debate en la siguiente sección.

En esta sección, dividiremos la discusión en dos partes. En la primera de ellas discutimos la conservación de la energía-momentum de los campos de materia, es decir, la conservación del tensor de energía-momentum $T_{\mu\nu}$. En la segunda parte, centramos la discusión en la conservación de la energía-momentum del campo gravitacional $t_{\mu\nu}$.

5.3.1. Conservación de la energía – momentum de los campos de materia

*Leyes de conservación en la RE y en la RG*⁹¹

Iniciemos examinando la formulación de expresiones de conservación desde un punto de vista general, para luego relacionar estos resultados con los escenarios relativistas especial y general. Supongamos que se cuenta con un operador derivada ∇_μ y un campo tensorial

⁹¹ Ver Read (2020) para más detalles y precisiones matemáticas respecto a lo mencionado en esta parte de la exposición.

$T^{\mu_1 \dots \mu_n}$, de modo que al aplicar el operador derivada al campo tensorial antes mencionado se obtiene lo siguiente

$$\nabla_{\mu} T^{\mu_1 \dots \mu_n} = 0 \quad (5.12)$$

Si esta ecuación se expresa en términos de derivadas ordinarias, resulta

$$\partial_{\mu_1} T^{\mu_1 \dots \mu_n} + \Gamma^{\mu_1}_{\sigma \mu_1} T^{\sigma \mu_2 \dots \mu_n} + \dots + \Gamma^{\mu_n}_{\sigma \mu_1} T^{\mu_1 \dots \mu_{n-1} \sigma} = 0 \quad (5.13)$$

o más resumidamente

$$\partial_{\mu_1} T^{\mu_1 \dots \mu_n} + \Delta^{\mu_2 \dots \mu_n} = 0 \quad (5.14)$$

donde los $\Gamma^{\mu}_{\nu\sigma}$ son los símbolos de Christoffel o componentes de conexión asociados al operador derivada ∇_{μ} , y $\Delta^{\mu_2 \dots \mu_n}$ agrupa los términos de conexión de la expresión (5.13).

Ahora bien, con las ecuaciones anteriores en mente podemos evaluar dos posibles escenarios para el operador derivada ∇_{μ} : (i) cuando este operador es considerado plano y (ii) cuando es considerado curvo. Veamos, en primer lugar, el escenario (i). Este escenario implica que los componentes de conexión en la ecuación (5.14) desaparecerán y, en consecuencia, esta ecuación se verá reducida a la siguiente expresión

$$\partial_{\mu_1} T^{\mu_1 \dots \mu_n} = 0 \quad (5.15)$$

Esta última ecuación es válida solo localmente, es decir, en la vecindad de algún punto $p \in M$. Luego, dada la validez local de (5.15), es posible integrar sobre un volumen $V \subset M$ y aplicar el teorema de Gauss para obtener la siguiente expresión⁹²

$$\int_{S=\partial V} n_{\mu_1} T^{\mu_1 \dots \mu_n} dS = 0 \quad (5.16)$$

Es importante precisar que la ecuación (5.15) y, en consecuencia, la ecuación (5.16) son invariantes solo bajo el grupo de Poincaré, es decir, estas expresiones preservan su

⁹² En esta expresión n^a es un vector normal a $S = \partial V$.

forma solo bajo sistemas de coordenadas que se encuentren relacionados por este tipo de transformaciones. Para transformaciones de coordenadas generales, las expresiones anteriores no son válidas (ver Read, 2020, Sección 2.2, para más detalles).

Por otra parte, en el escenario (ii), cuando ∇_a es considerado curvo, los componentes de conexión en la expresión (5.14) no pueden ser eliminados y, por lo tanto, debe mantenerse la expresión general dada por la ecuación antes mencionada. Sin embargo, igualmente es posible integrar la expresión (5.14) y aplicar el teorema de Gauss para obtener la siguiente ecuación

$$\int_S n_{\mu_1} T^{\mu_1 \dots \mu_n} dS = - \int_V \Delta^{\mu_2 \dots \mu_n} dV = 0 \quad (5.17)$$

No obstante, a diferencia de lo que ocurre en el caso (i), esta ecuación integral es dependiente de coordenadas. Esto se puede ver con facilidad al considerar que el lado derecho de esta expresión se encuentra compuesto por objetos geométricos no tensoriales; en consecuencia, el lado izquierdo también será dependiente de coordenadas.

En general, cuando en física se alude a ecuaciones de conservación diferenciales e integrales, se está haciendo referencia a expresiones de la forma (5.15) y (5.16), respectivamente. En efecto, la ecuación (5.15) establece que, en cualquier punto de la variedad, un cambio en la energía-momentum del campo tensorial $T^{\mu_1 \dots \mu_n}$ debe ser balanceado por un cambio igual y opuesto en ese punto. De manera semejante, la ecuación (5.16) puede ser leída de modo que cualquier cambio del campo tensorial $T^{\mu_1 \dots \mu_n}$ a través de la frontera de una región $S \subset M$ debe ser equilibrado por un cambio igual y opuesto en el campo tensorial a través de esa región. En suma, tanto en el caso diferencial como en el caso integral, la energía-momentum del campo tensorial se conserva (en un punto en el caso de la ley de conservación diferencial y en un volumen para el caso de la ley de conservación integral)⁹³.

⁹³ Ver Read (2020, Sección 2.3) para una explicación formal acerca de la legitimidad de las expresiones (5.15) y (5.16) como ecuaciones de conservación diferencial e integral, respectivamente.

Veamos, a continuación, como se aplica la discusión general ofrecida hasta aquí a los marcos relativistas especial y general. Recordemos que tanto en la RE como en la RG se trabaja con tensores de rango 2, por lo tanto, la expresión general (5.13) o (5.12) se reduce a

$$\nabla_{\mu} T^{\mu\nu} = \partial_{\mu} T^{\mu\nu} + \Gamma_{\mu\lambda}^{\mu} T^{\lambda\nu} + \Gamma_{\mu\lambda}^{\nu} T^{\mu\lambda} = 0 \quad (5.18)$$

Ahora bien, en la RE el operador derivada ∇_{μ} es considerado plano como en el escenario (i) analizado más arriba⁹⁴ y, por lo tanto, la expresión (5.18) se reduce aún más⁹⁵, obteniéndose

$$\nabla_{\mu} T^{\mu\nu} = \partial_{\mu} T^{\mu\nu} = 0 \quad (5.19)$$

Esta última ecuación es considerada como una ley de conservación diferencial en la RE, válida para todos los sistemas de coordenadas relacionados por las transformaciones de Poincaré, pues esta expresión conserva la forma de la ecuación (5.15). Como hemos afirmado anteriormente, a partir de esta ley diferencial de conservación podemos construir, sin dificultades, una expresión integral que represente la conservación de la energía-momentum de la materia en una región determinada del espacio-tiempo. Esta ecuación preservará la forma de la expresión (5.16) y, en consecuencia, también será válida en aquellos sistemas de coordenadas vinculados por las transformaciones antes mencionadas. En conclusión, en la RE es posible obtener expresiones de conservación bien definidas, tanto diferenciales (locales) como integrales (globales)⁹⁶, para la energía-momentum de los campos de materia, las cuales serán válidas en una clase de marcos relacionados por el grupo de transformaciones de Poincaré.

⁹⁴ Específicamente, el operador derivada en el caso relativista especial cumple con las condiciones de ser libre de torsión y compatible con la métrica fija de Minkowski $\eta_{\mu\nu}$. Por otro lado, para el caso relativista general, que veremos a continuación, el operador derivada cumple con ser libre de torsión y compatible con la métrica $g_{\mu\nu}$ (ver nota al pie 17 más arriba).

⁹⁵ Esto debido a que, como se mencionó más arriba, para una conexión plana los coeficientes de conexión desaparecen.

⁹⁶ Siguiendo el uso estándar de los términos “local” y “global” en el contexto del debate acerca del problema de las leyes de conservación en la RG, entenderemos que una ecuación de conservación diferencial es una ecuación local, a la vez que una ecuación de conservación integral es una ecuación global. En adelante, y de no especificar algo distinto, el uso que le daremos a estos términos será en este sentido.

Volvamos ahora nuestra atención al caso relativista general. Como es sabido, en la RG el espacio-tiempo es considerado curvo, por lo que ahora estamos situados en el escenario (ii) analizado más arriba. En consecuencia, la ecuación (5.13) ya no puede ser reducida a la forma general de una ley diferencial de conservación, como la ecuación (5.15): solo tendremos una ecuación de conservación diferencial dada por la expresión (5.18).

La ecuación (5.18) es postulada en los libros de texto como una ley de conservación diferencial válida en contextos relativistas generales. Sin embargo, este hecho es objeto de debate pues, como hemos mencionado, la expresión (5.18) no preserva la forma general de una ley diferencial de conservación. Además, al integrar la expresión (5.18) se obtiene una expresión de la forma (5.17), que como ya hemos discutido, presenta la dificultad de ser dependiente de coordenadas. Estas dificultades serán objeto de debate en la siguiente sección.

Resumamos los resultados obtenidos hasta ahora. Hemos afirmado que para el caso relativista especial podemos, en efecto, obtener leyes de conservación bien definidas tanto en su forma diferencial como integral, las cuales son válidas en una clase de marcos relacionados por las transformaciones de Poincaré. En cambio, en el caso relativista general, obtenemos ecuaciones de conservación de la forma (5.18) y (5.17). Estas presentan ciertas dificultades: por un lado, (i) es discutible que la expresión (5.18) sea una ecuación de conservación diferencial válida, pues esta no preserva la forma de la ecuación (5.15). Por otro lado, (ii) la ecuación que obtenemos al integrar (5.18) es dependiente de coordenadas, pues esta tiene la forma de la ecuación (5.17), la cual presenta esta característica.

Isometrías del espacio-tiempo

Ahora bien, en adición al análisis acerca de las leyes de conservación presentado más arriba, debemos incluir en la discusión el papel que juegan las simetrías en la construcción de cantidades invariantes, en particular, las simetrías del espacio-tiempo o también llamadas “isometrías”. La manera usual de comprender el lugar que ocupan las simetrías en la construcción de cantidades conservadas es en el contexto del formalismo Lagrangiano y Hamiltoniano, dentro de los cuales los teoremas de Noether (1918) (en particular, el primer

teorema de Noether) ocupan un lugar central. En términos muy generales, el primer teorema de Noether establece que, a cada transformación continua de coordenadas (o para cada simetría del espacio-tiempo) que deje invariante la acción de un sistema físico, le corresponde una cantidad j^μ , denominada corriente de Noether, que cumple con $\partial_\mu j^\mu = 0$. De este modo, cada una de estas “corrientes de Noether” representa una cantidad conservada que puede ser integrada mediante el teorema de Gauss⁹⁷.

El primer teorema de Noether pone de manifiesto que la conservación de la energía-momentum no solo depende de las características de los campos de materia, sino que también se encuentra estrechamente relacionada con la estructura del espacio-tiempo o, más específicamente, con las isometrías. Estas “isometrías” pueden ser caracterizadas por medio de sus generadores asociados, es decir, por medio de sus campos vectoriales de Killing K^μ , los cuales se definen por la ecuación⁹⁸

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \quad (5.20)$$

La presencia de estos vectores asegura la existencia de un grupo de isometrías que permite construir cantidades invariantes⁹⁹. Básicamente, la idea es que si un espacio-tiempo $(M, g_{\mu\nu})$ posee un campo vectorial tipo-tiempo K^μ , es posible construir la expresión $T^{\mu\nu}K^\mu$ que representa la densidad de energía-momentum relativa a K^μ y cuya derivada covariante es igual a cero (Lam, 2011, p. 1015, para más detalles y precisiones matemáticas):

$$\nabla_\mu (T^{\mu\nu}K_\nu) = (\nabla_\mu T^{\mu\nu})K_\nu + T^{\mu\nu}(\nabla_\mu K_\nu) = \frac{1}{2}T^{\mu\nu}(\nabla_\mu K_\nu + \nabla_\nu K_\mu) = 0 \quad (5.21)$$

Como bien explica Curiel (2000), la construcción de una ley integral de conservación requiere de una familia de curvas temporales (u observadores) tales que se aplica el tensor

⁹⁷ Por ejemplo, en la RE, para una acción que es invariante bajo las transformaciones de Poincaré se tendrán 10 cantidades conservadas, pues este grupo de transformaciones tiene 10 parámetros constantes.

⁹⁸ En particular, un campo vectorial K^μ de un espacio-tiempo $(M, g_{\mu\nu})$ es un campo vectorial de Killing si se cumple que la derivada de Lie de la métrica en la dirección de K^μ es igual a cero, es decir, $\mathcal{L}_K g = 0$, lo cual implica que $\nabla_\mu K_\mu + \nabla_\nu K_\nu = 0$.

⁹⁹ En particular, la presencia de un campo vectorial de Killing K^μ asegura que la métrica $g_{\mu\nu}$ del espacio-tiempo presente una simetría en la dirección de K^μ .

de energía-momentum (como en la ecuación (5.21)) a sus respectivos vectores tangente en cada punto de la región sobre la cual se quiere establecer la ley de conservación. Como resultado de este proceso, se obtiene un campo escalar que representa la densidad de energía-momentum en esa región del espacio-tiempo, y es este campo escalar el que finalmente se utiliza para construir la ley de conservación integral buscada. Un detalle importante: la familia de curvas temporales deben ser las curvas integrales¹⁰⁰ de un campo vectorial de Killing de modo que la ecuación (5.20) se cumpla. En otras palabras, es a partir de la ecuación (5.20) que podemos formular una ley integral de conservación, donde las únicas curvas que satisfacen esta ecuación son las curvas integrales de un campo de Killing.

Ahora bien, si la presencia de los vectores de Killing K^μ se encuentra asegurada y, por lo tanto, la ecuación (5.20) se cumple, entonces a partir de esta expresión, y tal como se ha explicado en el párrafo anterior, es posible construir, usando el teorema de Gauss, la siguiente expresión que representa una ecuación de conservación integral:

$$\int_{S=\partial V} T^{\mu\nu} K_\nu n_\mu dS = \int_V \nabla_\mu (T^{\mu\nu} K_\nu) dV = 0 \quad (5.22)$$

donde n_μ es un vector normal a S , dS es un elemento de superficie y dV es un elemento de volumen. La expresión integral dada por la ecuación (5.22) representa, inequívocamente, una ley de conservación integral con respecto a las curvas integrales de un campo vectorial de Killing K^μ . Esta expresión establece que para cualquier volumen del espacio-tiempo, toda la energía-momentum que entra en ese volumen es equilibrado por una cantidad de energía-momentum que sale de ese mismo volumen. Una formulación equivalente para la expresión (5.22) es la siguiente

$$\int_\Sigma T_{\mu\nu} K^\nu n^\mu = 0 \quad (5.23)$$

la cual representa la conservación de la energía-momentum no gravitacional en un volumen contenido en una hipersuperficie tipo-espacio Σ (ver Lam, 2011, p. 1016; Wald, 1984, p.

¹⁰⁰ Una curva integral es una curva que corresponde a una solución de una ecuación diferencial. En este caso, corresponde a una solución de la ecuación (5.20).

286). En este punto, es importante notar lo siguiente: a diferencia de las expresiones (5.15) y (5.16), las cuales son válidas solo para sistemas de coordenadas relacionados por las transformaciones de Poincaré, la ecuación (5.22) (o la ecuación (5.23)) es válida para cualquier marco de referencia arbitrario, pues esta se encuentra expresada en un lenguaje covariante al contraer el tensor energía-momentum con el campo vectorial de Killing ($T^{\mu\nu}K^\mu$).

En la RE, la presencia de los vectores de Killing se encuentra asegurada, ya que el espacio-tiempo plano de Minkowski presenta las simetrías necesarias. Sin embargo, la situación es bastante más compleja en la RG. En general, no es posible asegurar la existencia de un campo vectorial de Killing; estas estructuras se encuentran presentes solo bajo situaciones particulares o en sistemas físicos idealizados (ver secciones siguientes para más detalles)¹⁰¹. No obstante, como hemos visto, este último hecho no impide que en los espacio-tiempos $(M, g_{\mu\nu})$ que presentan vectores de Killing podamos construir ecuaciones integrales de conservación bien definidas.

5.3.2. Conservación de la energía-momentum total

Debido a que en los espacio-tiempo curvos no siempre se cuenta con simetrías de la métrica, es decir, con los vectores de Killing, no siempre es posible construir una expresión integral de la forma (5.22) o (5.23) para la conservación de la energía-momentum del campo gravitacional siguiendo los pasos indicados en la subsección anterior. En lugar de ello, los físicos han propuesto una expresión diferencial de la forma dada por la ecuación (5.15) para la energía-momentum total, es decir, para los campos de materia más el campo gravitacional, la que sí cumpliría con ser una cantidad conservada. Esta expresión es la siguiente:

¹⁰¹ Una discusión interesante en torno a las leyes de conservación en la RG es que para que estas puedan ser formuladas es necesario agregar algunas estructuras no-dinámicas al espacio-tiempo. El campo vectorial de Killing puede ser comprendido de este modo. Lo anterior, en realidad, sería una consecuencia de la propiedad *background independence* de la RG que impediría establecer expresiones válidas para la conservación de la energía-momentum sin adicionar estructuras adicionales (véase, por ejemplo, Lam, 2011). En la siguiente sección se discutirá con más detalle este problema.

$$\partial_\nu(T^{\mu\nu} + t^{\mu\nu}) = 0 \quad (5.24)$$

donde, como hemos visto, $T^{\mu\nu}$ representa la energía-momentum de los campos de materia, y $t^{\mu\nu}$ representa la energía-momentum del campo gravitacional. La idea detrás de esta ecuación es que a partir de estas dos cantidades en conjunto se pueda construir una ley de conservación de la energía-momentum total, en analogía con el caso de la conservación de la energía-momentum de los campos de materia a partir de la ecuación (5.15). Ahora bien, la cantidad $t^{\mu\nu}$ puede ser definida en términos de un superpotencial $U^{\mu\lambda\nu}$, usando la ecuación (5.24) y las ecuaciones de campo de Einstein (5.9), de la siguiente forma

$$t^{\mu\nu} = \partial_\lambda U^{\mu\lambda\nu} - \frac{1}{8\pi} G^{\mu\nu} \quad (5.25)$$

Sin embargo, debido a que la ecuación (5.24) no define al superpotencial de manera unívoca (puesto que este puede ser elegido arbitrariamente¹⁰²), usualmente se afirma que la cantidad $t^{\mu\nu}$ puede ser expresada a través de diferentes expresiones no equivalentes¹⁰³.

Por otro lado, también es importante notar que la cantidad $t^{\mu\nu}$, que representa la energía del campo gravitacional, es dependiente de coordenadas (este hecho se deduce de la definición de $t^{\mu\nu}$, la cual considera derivadas parciales que, como se sabe, son definidas respecto a un sistema de coordenadas) y, por lo tanto, $t^{\mu\nu}$ no es un tensor, sino un pseudo-tensor, puesto que es una cantidad que no es invariante bajo una transformación de coordenadas arbitraria. En otras palabras, $t^{\mu\nu}$ representa la energía del campo gravitacional como una cantidad que es dependiente del sistema de coordenadas elegido¹⁰⁴. De acuerdo con esta característica del pseudo-tensor podría suceder, por ejemplo, que en dos sistemas de coordenadas distintos obtengamos diferentes cantidades de energía-momentum del campo gravitacional para una misma región del espacio-tiempo¹⁰⁵ o que, incluso, bajo la

¹⁰² Un argumento en contra de esta afirmación se puede encontrar en De Haro (2021).

¹⁰³ Algunas de estas definiciones no equivalentes para $t^{\mu\nu}$ incluyen el pseudo-tensor de Einstein y el pseudo-tensor de Landau-Lifschitz.

¹⁰⁴ Para una definición formal de este pseudo-tensor, véase Misner, Thorne, y Wheeler (1973, pp. 464–466).

¹⁰⁵ Una interesante discusión acerca de esta característica de la energía-momentum del campo gravitacional, expresada a través del pseudo-tensor $t^{\mu\nu}$, es si esta puede ser interpretada como una manifestación de que esta energía-momentum no es “única” o más bien de que no es “genuina”

elección de un sistema de coordenadas adecuadas la energía-momentum del campo gravitacional se anule¹⁰⁶.

A pesar de los inconvenientes con el pseudo-tensor $t^{\mu\nu}$ mencionado anteriormente, es posible construir una expresión integral a partir de la ecuación (5.24). Esta expresión tiene la siguiente forma:

$$\int_{S=\partial V} T^{\mu\nu} n_\nu dS = - \int_{S=\partial V} t^{\mu\nu} n_\nu dS \quad (5.26)$$

Esta ecuación establece que cualquier cambio en la energía-momentum de los campos de materia $T^{\mu\nu}$ en una región determinada será balanceado por un cambio igual y opuesto en la energía-momentum gravitacional $t^{\mu\nu}$ en esa región. Sin embargo, el carácter pseudotensorial de $t^{\mu\nu}$, provoca que la expresión anterior no se encuentre bien definida, pues es dependiente de las coordenadas elegidas (ver Hofer, 2000, p. 194; Read, 2020, Sección 3.3).

No obstante, aún es posible obtener una expresión integral de la conservación de la energía-momentum total bien definida bajo ciertas condiciones que representan circunstancias físicas particulares. Estas condiciones son (Hofer, 2000, p. 194): (i) las integrales deben ser tomadas en el límite $r \rightarrow \infty$, (ii) se asume la planitud asintótica¹⁰⁷ del espacio-tiempo, es decir, $g_{\mu\nu} \rightarrow \eta_{\mu\nu}$ y (iii) el sistema de coordenadas debe ser asintóticamente Lorentziano, pero debe variar arbitrariamente en el interior. Desde un punto de vista general, estas condiciones permiten abordar el problema de la conservación de la energía-momentum total sobre un espacio-tiempo con características Minkowskianas, donde las isometrías son recuperadas (y, por tanto, se reestablece la presencia de los

(precisamente por no ser única) (ver Pitts, 2010; Lam, 2011, pp. 1019–1020). Ver siguiente sección para más detalles.

¹⁰⁶ En la bibliografía técnica, el problema de encontrar un objeto adecuado para la energía-momentum gravitacional se conoce como el “problema de la no localización de la energía gravitacional”. Véase, por ejemplo, Misner, Thorne y Wheeler (1973, pp. 466–468).

¹⁰⁷ En términos generales, la planitud asintótica es una característica del espacio-tiempo por la que se supone la materia ubicada en una región compacta del mismo, de modo que a grandes distancias de ella el espacio-tiempo puede considerarse plano. Un espacio-tiempo asintóticamente plano puede pensarse como un sistema físico aislado (Wald, 1984, p. 269). Ver siguiente sección para más detalles.

vectores de Killing) y, por lo tanto, puede obtenerse una expresión de la forma (5.22). En efecto, contrayendo la ecuación (5.26) con las componentes del campo vectorial de Killing, se obtiene

$$\int_{S=\partial V} t^{\mu\nu} K_{\mu} n_{\nu} dS = - \int_{S=\partial V} T^{\mu\nu} K_{\mu} n_{\nu} dS = 0 \quad (5.27)$$

que representa una ley de conservación integral bien definida, pues es una expresión válida para cualquier marco arbitrario de referencia (tal como la ecuación (5.22)). Si bien algunos autores argumentan que las condiciones (i) a (iii), que permiten arribar a la ley de conservación integral anterior, representan condiciones físicas poco realistas (por ejemplo, Hofer, 2000; Curiel, 2000), lo cierto es que la expresión (5.27) es usada, por ejemplo, para calcular la energía perdida por un sistema debido al transporte de ondas gravitacionales (ver, por ejemplo, Hofer, 2000, p. 194). En la próxima sección se discutirán en detalle este y otros problemas relacionados con las ecuaciones de conservación en contextos relativistas generales.

5.4. Diferentes interpretaciones respecto del problema de las leyes de conservación

En esta sección examinaremos las diferentes interpretaciones que se han ofrecido en la bibliografía filosófica reciente al problema de las leyes de conservación presentado en la sección anterior. En general, estas interpretaciones tienen relación con la legitimidad de las expresiones postuladas en la RG para representar genuinos principios de conservación diferenciales o integrales, así como también con la existencia y fundamentalidad¹⁰⁸ de la energía-momentum gravitacional, nociones que se suponen definidas a partir de las expresiones de conservación¹⁰⁹.

¹⁰⁸ En términos generales, en el debate en torno a las leyes de conservación en la RG, una cantidad es considerada “fundamental” si existen ciertas estructuras matemáticas en el seno de la teoría que den cuenta apropiadamente de esa cantidad. Estas estructuras, además, deben guardar cierta coherencia con el resto del aparato formal de la teoría.

¹⁰⁹ Como veremos en las subsecciones siguientes, el debate filosófico en torno a las ecuaciones de conservación se relaciona no solo con la validez de estas expresiones, en el sentido de encontrarse

Debido a que ciertas expresiones de conservación serán centrales para la discusión que desarrollaremos a continuación, recordemos nuevamente aquí los resultados más importantes obtenidos en la sección precedente. Veamos, en primer lugar, los resultados a los que hemos arribado para la conservación de la energía-momentum de los campos de materia. Por un lado, hemos dicho que en la física de campos clásica (pre RG), la ecuación (5.19)

$$\partial_{\mu} T^{\mu\nu} = 0 \quad (5.19)$$

es presentada en los libros de texto como la ecuación de conservación diferencial de la energía-momentum de la materia, a partir de la cual podemos derivar, mediante la aplicación del teorema de Gauss, una versión integral para la conservación de la energía-momentum de la materia ordinaria. Por otro lado, cuando nos centramos en el espacio-tiempo curvo de la RG, la bibliografía técnica presenta la ecuación (5.18) como la versión diferencial de la ley de conservación de la energía-momentum de los campos de materia

$$\nabla_{\mu} T^{\mu\nu} = \partial_{\mu} T^{\mu\nu} + \Gamma_{\mu\lambda}^{\mu} T^{\lambda\nu} + \Gamma_{\mu\lambda}^{\nu} T^{\mu\lambda} = 0 \quad (5.18)$$

Esta última expresión, mediante el uso de ciertas isometrías del espacio-tiempo (y en la medida en que estas se encuentren presentes), puede ser integrada, obteniendo así la versión integral para la conservación de la energía-momentum de los campos de materia, dada por la ecuación (5.22)

$$\int_{S=\partial V} T^{\mu\nu} K_{\nu} n_{\mu} dS = \int_V \nabla_{\mu} (T^{\mu\nu} K_{\nu}) dV = 0 \quad (5.22)$$

adecuadamente definidas, sino también con la existencia de la energía-momentum gravitacional y su fundamentalidad. En general, se supone que la existencia y fundamentalidad de la energía-momentum gravitacional dependen de la formulación de ecuaciones de conservación bien definidas. Respecto a la fundamentalidad de ciertas cantidades físicas, como la energía o la masa y su relación con los principios de conservación, ver Norton (2015, Sección 3.5). En la discusión siguiente nos enfocaremos principalmente en el debate acerca de la legitimidad de las ecuaciones diferenciales e integrales de conservación, así como también en la existencia de la energía-momentum gravitacional. Dejamos en un segundo plano el problema de la fundamentalidad. En la Sección 5.5 justificamos esta decisión.

Volvamos ahora nuestra atención hacia los resultados que se obtuvieron para la conservación de la energía-momentum de la energía total (energía-momentum de los campos de materia más energía-momentum gravitacional). En los libros de texto de RG se presenta a la ecuación (5.24) como la expresión que representa la ley de conservación diferencial de esta de la energía-momentum total

$$\partial_\nu(T^{\mu\nu} + t^{\mu\nu}) = 0 \quad (5.24)$$

Y la expresión (5.27) se presenta como la versión integral de esta misma cantidad:

$$\int_{S=\partial V} t^{\mu\nu} K_\mu n_\nu dS = - \int_{S=\partial V} T^{\mu\nu} K_\mu n_\nu dS = 0 \quad (5.27)$$

Recuérdese que la ecuación (5.27) se obtiene a partir de la ecuación (5.24) solo en el caso de que ciertas condiciones, principalmente asociadas a la noción de planitud asintótica, se cumplan.

Cuando examinamos la discusión de la sección anterior desde un punto de vista filosófico, las expresiones que de acuerdo a la teorización física aparecen como representantes de las leyes de conservación tanto para la energía-momentum de la materia ordinaria como para la energía-momentum total, principalmente las expresiones (5.18), (5.19), (5.22), (5.24) y (5.27) mencionadas más arriba, reciben diferentes interpretaciones respecto de su validez o plausibilidad. Algunos autores argumentan en favor de un punto de vista eliminativista acerca de la existencia de ecuaciones de conservación bien definidas en la RG (Hofer, 2000; Dürr, 2019a, 2019b). Otros aceptan la validez de ciertas expresiones que representan leyes de conservación en la RG, dependiendo de si existen o no ciertas estructuras de fondo sobre las cuales definir las (Lam, 2011). Por último, se han formulado también argumentos en favor de un punto de vista realista acerca de la existencia de ecuaciones de conservación apropiadas en la RG (Read, 2020). Comencemos examinando los argumentos en favor del punto de vista eliminativista.

5.4.1. El eliminativismo de Hoefer

Uno de los trabajos canónicos en relación al debate de las leyes de conservación en la RG se debe a Hoefer (2000). En su artículo titulado “*Energy Conservation in GTR*” el autor defiende un punto de vista eliminativista respecto de la existencia de leyes de conservación en esta teoría, afirmando que no existe algo así como un principio o ley de conservación de la energía-momentum ni para la materia ordinaria (campos de materia) ni para el campo gravitacional. Sus argumentos se basan en el análisis de lo que el autor denomina “la herencia recibida” respecto de las leyes de conservación en la RG.

En primer lugar, el autor argumenta en contra de la legitimidad de la ecuación (5.18) como representante de una ley de conservación diferencial para la energía-momentum de la materia ordinaria. El autor sostiene que, de acuerdo con la “herencia recibida”, esto es, la corriente de pensamiento que se ha asumido en los libros de texto de la RG para presentar el tópico de las leyes de conservación, la ecuación (5.18) se justifica como una ecuación de conservación válida haciendo uso del principio de equivalencia fuerte. Este principio, recordemos, afirma que todas las leyes físicas, en el origen de un sistema de referencia inercial local, son las mismas que su contraparte clásica de la relatividad especial (Hoefer, 2000, p. 191) (recuérdese la Sección 5.1 donde se ha discutido con más detalles acerca del principio de equivalencia). De este modo, la ecuación (5.18) representaría un principio de conservación válido para la energía-momentum de los campos de materia en la RG debido a que la ecuación (5.19) es la ley de conservación diferencial correcta en un espacio-tiempo plano, como el de la RE. Sin embargo, afirma el autor, la “herencia recibida” también ha promovido un punto de vista contrario al anterior. De acuerdo con este segundo punto de vista, la ecuación (5.18) no corresponde a una ley de conservación legítima en la RG, sino más bien esta nos indicaría el intercambio de energía-momentum entre el campo gravitacional y la energía ordinaria. Es decir, esta última ecuación correspondería a una expresión que describe cómo la energía-momentum gravitacional cambia el balance de energía de un sistema. Esta interacción estaría representada por los términos adicionales que aparecen en la ecuación (5.18) respecto de su homóloga para un espacio-tiempo plano, la ecuación (5.19). Además de lo anterior, Hoefer afirma que es conflictivo aceptar la expresión (5.18) como una ecuación de conservación válida puesto que, a diferencia de lo que sucede con la expresión (5.19), no es posible arribar a una ecuación integral de

conservación a partir de ella, puesto que no es posible aplicar el teorema de Gauss¹¹⁰. Hoefler se inclina por este último punto de vista, afirmando que en la RG no existe un principio de conservación diferencial de la energía-momentum para los campos de materia. El autor afirma que detrás de esta dificultad se encuentra el hecho de que el principio de equivalencia fuerte no estaría suficientemente justificado para aceptar la inferencia que nos lleva de la ecuación (5.19) para un espacio-tiempo plano a la ecuación (5.18) para un espacio-tiempo curvo.

En segundo lugar, Hoefler también asume una posición eliminativista respecto de la existencia de una expresión diferencial que represente la conservación de la energía-momentum de los campos de materia más la energía-momentum gravitacional (energía-momentum total), como la ecuación (5.24) que se presenta en los libros de texto para desempeñar este papel. Hoefler presenta dos argumentos en contra de la validez de esta ecuación, ambos relacionados con la naturaleza del pseudo-tensor $t^{\mu\nu}$. Por un lado, (i) el autor señala que el pseudo-tensor no se encuentra unívocamente definido (Hoefler, 2000, p. 193) (recuérdese la sección anterior respecto a cómo se define el pseudo-tensor $t^{\mu\nu}$); por otro lado, (ii) el autor argumenta que debido a la naturaleza no tensorial de $t^{\mu\nu}$, esta cantidad puede tomar diferentes valores dependiendo del marco de referencia elegido, incluso pudiendo ser nulo si se escoge un marco de referencia adecuado¹¹¹ (Hoefler, 2000, p. 193). Como ha sido mencionado en la sección anterior, este último hecho es el que se conoce en los libros de texto como el carácter no local de la energía-momentum gravitacional, pues el pseudo-tensor $t^{\mu\nu}$ no es suficiente para capturar el intercambio local¹¹² de energía-momentum entre la gravedad y la materia. El autor concluye, en consecuencia, que ni la ecuación (5.18) para la energía-momentum de los campos de materia ni la ecuación (5.24) para la energía-momentum total pueden representar cantidades

¹¹⁰ Esto se debe a que el teorema de Gauss solo es aplicable a expresiones con derivadas ordinarias (parciales) y no a expresiones que contienen derivadas covariantes.

¹¹¹ En un trabajo reciente, De Haro (2021) afirma que el pseudo-tensor $t^{\mu\nu}$ no solo es dependiente de las coordenadas escogidas, sino que además este tampoco presenta propiedades de transformación entre estos sistemas de coordenadas.

¹¹² En este caso, el término “local” hace referencia a una región definida del espacio-tiempo y no se relaciona necesariamente con los términos “local” y “global” asociados a las cantidades diferenciales y locales, respectivamente.

bien definidas para la conservación diferencial de la energía-momentum en la teoría de la gravitación de Einstein.

Por último, Hofer también rechaza la expresión (5.27) como una ley de conservación integral para la energía-momentum total. En efecto, el autor afirma que la expresión (5.27) solo puede ser obtenida bajo ciertas condiciones –las condiciones de planitud asintótica– que poco tienen que ver con las características de nuestro mundo actual. Por una parte, (i) Hofer argumenta que debido a que nuestro mundo actual no es asintóticamente Minkowskiano, la ecuación (5.27) no se aplicaría a la gravedad en nuestro mundo actual (2000, p. 194) y, por otra parte, (ii) el autor afirma que aceptar la ecuación (5.27), debido a la presencia del pseudo-tensor $t^{\mu\nu}$ en ella, implica posicionarse en contra de un principio filosófico que ha tenido un avance progresivo en la física teórica en los últimos años, esto es, el de dar preferencia como cantidades físicamente significativas a aquellas que son independientes de la elección de un sistema de coordenadas (Hofer, 2000, pp. 194-195).

El eliminativismo de Hofer se manifiesta claramente, entonces, en el hecho de que el autor rechaza tanto las versiones diferenciales como integrales de las leyes de conservación –tal como son presentadas en la bibliografía técnica– como expresiones válidas que puedan representar el balance de energía y momentum en el marco de la RG. Como una consecuencia del hecho anterior, el autor también impugna la afirmación que sostiene que en la RG existe una noción clara y definida de “energía-momentum gravitacional”, pues esta noción se define a partir de las expresiones diferenciales e integrales de conservación que el autor rechaza. Hofer concluye, de este modo, que la energía-momentum gravitacional no califica como una propiedad fundamental en contextos relativistas generales¹¹³.

¹¹³ Nótese que se ha concluido que la energía-momentum gravitacional no sería una cantidad fundamental en la RG debido a que las ecuaciones o leyes de conservación asociadas a ella no se encontrarían correctamente definidas. Esta conclusión se funda en cierta concepción estándar que comprende que la fundamentalidad de cantidades físicas, como la masa o la energía, se decide en virtud de su conservación. En relación al comentario anterior, es importante destacar que, en términos generales, el debate en torno a las leyes de conservación se da bajo el supuesto antes mencionado. En general, los autores realizan un análisis de las conclusiones que es posible extraer en relación a la energía-momentum gravitacional teniendo en consideración únicamente la estructura formal y principios de la RG. Recuérdese nota al pie 41.

En un trabajo reciente, Dürr (2019a) llega a una conclusión similar a la de Hoefel respecto a la existencia de la noción de energía-momentum gravitacional en la RG, aun cuando acepta la validez de las expresiones (5.18) y (5.22) para la conservación de la energía-momentum de los campos de materia. Después de analizar varias propuestas que intentan capturar la noción de energía gravitacional local en la RG, como el enfoque pseudo-tensorial –el que mayoritariamente se asume en los libros de texto y sobre el cual se basan las discusiones de esta sección y de la precedente–, la propuesta tensorial de Lorentz y Levi-Civita y las propuestas de Brian Pitts (2010) y David Baker (2005), Dürr concluye que la posición más adecuada respecto de la noción de energía-momentum gravitacional en la RG es la que se asemeja al antirrealismo o eliminativismo que se adopta frente a las fuerzas inerciales en la mecánica clásica. Esto se debe a que ninguna de las propuestas que pretenden caracterizar la noción de energía-momentum gravitacional local sería completamente satisfactoria.

5.4.2. Lam y las estructuras de fondo no-dinámicas

Volvamos ahora nuestra atención hacia la propuesta de Lam (2011) respecto del problema de las leyes de conservación en la RG. Lam afirma que en la RG efectivamente se cuenta con ciertas leyes de conservación de la energía-momentum, pero solo cuando ciertas estructuras matemáticas se encuentran disponibles, las cuales pueden ser interpretadas como una estructura de fondo sobre la cual se definen las correspondientes cantidades. Veamos más en detalle los argumentos del autor.

Para comenzar, el autor rechaza el hecho de que la ecuación (5.18) represente una genuina¹¹⁴ ley de conservación. Esto se debe a que, a diferencia de lo que sucede en un espacio-tiempo plano donde es posible obtener una expresión integral que se derive de la ecuación (5.19), en un espacio-tiempo curvo como el de la RG esta opción no se encuentra generalmente disponible a partir de la ecuación (5.18) (Lam, 2011, p. 1014). Al igual que

¹¹⁴ Es importante destacar que Lam asume que una ley de conservación es “genuina” cuando es expresada en su forma integral. Por lo tanto, de acuerdo con el autor, la ecuación (5.18) no sería una “genuina” ley de conservación, debido a que, en principio, no puede expresarse en una versión integral. En el debate acerca del problema de las leyes de conservación en la RG, esto no es un hecho universalmente aceptado. Otros autores, por ejemplo, Read (2020), afirman que tanto las leyes en su forma diferencial como en su forma integral pueden ser postuladas como genuinas leyes de conservación.

hace notar Hofer, Lam argumenta que en la expresión (5.18) aparecen términos adicionales que representan la interacción del campo gravitacional con los campos de materia, pero al mismo tiempo esta expresión no considera la energía-momentum del campo gravitacional, la que es necesaria para el balance de energía (Lam, 2011, p. 1014).

En segundo lugar, Lam también analiza la expresión (5.22), que expresa la conservación de la energía-momentum integral para los campos de materia. El autor afirma que esta ecuación también presenta dificultades, pues para arribar a ella es necesario que el espacio-tiempo presente algunas simetrías, que son caracterizadas por medio del denominado campo vectorial de Killing (recuérdese la discusión presentada en la sección anterior acerca de estas estructuras matemáticas). El autor reconoce que esta expresión puede ser interpretada como una genuina ley de conservación. No obstante, también puntualiza que las simetrías del espacio-tiempo, necesarias para formular la ecuación (5.22) no se encuentran en general disponibles para el universo en evolución, sino solo en casos particulares que, sin embargo, son muy útiles para diversas situaciones físicas como, por ejemplo, para estudiar los agujeros negros de Kerr-Newman (Lam, 2011, p. 1016).

Para Lam, las dificultades con la expresión (5.22) se deben a la naturaleza dual del campo métrico, esto es, como un objeto que codifica tanto los efectos inercio-gravitacionales como la estructura crono-geométrica del espacio-tiempo, hecho que puede ser caracterizado por la propiedad de independencia de fondo o *background independence*, distintiva de la RG¹¹⁵.

En efecto, para el autor, las peculiaridades de las expresiones de conservación para los campos de materia se deben a dos hechos que se encuentran conectados. Por un lado, serían ciertas características del espacio-tiempo mismo las que impedirían formular estas leyes de conservación. Es decir, para formular la ecuación integral de conservación (5.22) es necesario, como ya se ha mencionado más arriba, contar con la presencia de ciertas estructuras matemáticas representadas por los vectores de Killing. Estos objetos matemáticos pueden interpretarse, afirma Lam, como estructuras de fondo no dinámicas respecto de las cuales puede definirse adecuadamente una ley de conservación integral

¹¹⁵ En Dürr (2019a) puede encontrarse una crítica a la propuesta de Lam. Dürr sostiene que la noción de *background independence* es demasiado vaga y discutida para ser asociada, sin ambigüedades, al problema de las leyes de conservación en la RG.

(como la ecuación (5.22)). De hecho, los campos vectoriales de Killing puede ser vistos como un marco de referencia inercial global en relación al cual se definen estas leyes integrales (2011, p. 1017). Para el caso de la RG, es necesario contar con esta estructura no dinámica, pues como sabemos, la teoría de Einstein se caracteriza por postular un espacio-tiempo dinámico, y sería este hecho el que impediría formular leyes de conservación integrales. Por otro lado, los problemas con las expresiones (5.18) y (5.22) también pueden ser interpretados, sostiene Lam, como una consecuencia de no haber considerado el carácter universal de la energía-momentum gravitacional (2011, p. 1017). Todo sistema físico interactúa con el campo gravitatorio y, por lo tanto, su energía-momentum puede ser transformada en energía-momentum gravitacional y la energía-momentum gravitacional en energía-momentum ordinaria. Ahora bien, estas dos interpretaciones acerca de las expresiones de conservación para la materia ordinaria en realidad serían la misma, debido, como se dijo antes, al carácter dual del campo métrico que se expresa a través de la propiedad de independencia de fondo.

Respecto de la ecuación (5.24), que se postula como una ley de conservación diferencial para la energía-momentum de la materia ordinaria más la energía-momentum del campo gravitacional, Lam evalúa los dos argumentos que Hofer ofrece para cuestionar la legitimidad de esta expresión. Por una parte, (i) Lam señala que en la física actual, en efecto, existe la tendencia a aceptar como cantidades físicas relevantes a aquellas que se muestran independientes de la elección de un sistema de coordenadas, por lo que desde este punto de vista el pseudo-tensor $t^{\mu\nu}$ no calificaría como una cantidad física en toda regla. Por otra parte, (ii) el autor afirma que este pseudo-tensor tampoco se encuentra definido de forma unívoca, pues se puede definir de varias maneras equivalentes (2011, p. 1018) (recuérdese discusión de la sección anterior). De todos modos, parece que el autor no cree que estos argumentos sean concluyentes, pues discute brevemente una alternativa ofrecida por Pitts (2010), quien señala que algunas de las dificultades con las expresiones de conservación para la energía-momentum gravitacional se disolverían si se abandona el supuesto de unicidad asumido en (ii). Además, Lam también reconoce que no habría

inconsistencias con el hecho de aceptar la característica de dependencia de coordenadas del pseudo-tensor $t^{\mu\nu}$ que representa a la energía-momentum gravitacional¹¹⁶.

Por último, respecto de la ecuación (5.27), si bien Lam afirma que la conservación integral de la energía-momentum total no puede ser definida para el caso general (es decir, para un espacio-tiempo genérico que no cuente con las simetrías necesarias), el autor acepta que (5.27) es una expresión bien definida cuando se asumen las condiciones relacionadas a un espacio-tiempo plano, como se mencionó en la sección anterior. En consecuencia, sostiene Lam, estas expresiones solo tienen sentido si se introducen las estructuras de fondo apropiadas, donde una de estas estructuras posibles es la noción de planitud asintótica (2011, p. 1022). El autor también puntualiza que en estos casos las nociones de energía y/o momentum total se encuentran bien definidas y son, en general, nociones globales¹¹⁷, es decir, considerando al espacio-tiempo (asintóticamente plano) como un todo¹¹⁸.

Para finalizar la revisión de la posición defendida por Lam, es importante destacar que, si bien difiere del punto de vista de Hofer en el sentido que Lam sí cree que existen expresiones bien definidas para la conservación de la energía-momentum de la materia ordinaria y del campo gravitacional, el autor concuerda con Hofer en que las nociones de energía-momentum (y masa) no serían propiedades fundamentales en la RG. En efecto, dado que no es posible definir ecuaciones de conservación de la energía-momentum para el caso general (sino solo en casos particulares) y asumiendo que es en virtud de estas expresiones de conservación que las nociones de energía y momentum (y masa) adquieren el estatus de propiedades fundamentales, entonces estas nociones no tendrían esta característica en la RG. Que contemos con expresiones de conservación bien definidas

¹¹⁶ Un argumento en defensa de esta posición puede encontrarse en Dürr (2019a). Mientras la estructura que sea dependiente de coordenadas tenga el mismo grupo de simetrías que el espacio-tiempo dado, sostiene Dürr, la dependencia de coordenadas no tiene por qué ser problemática (2019a, p. 3).

¹¹⁷ Se debe notar que el término “global” usado aquí no tiene el mismo sentido que los términos “local” y “global” utilizados para diferenciar las expresiones de conservación en sus formas diferencial e integral.

¹¹⁸ Existen programas de investigación actuales como, por ejemplo, los programas cuasi-locales, que intentan definir la energía-momentum para regiones extendidas pero finitas del espacio-tiempo. Estos intentos, de todos modos, también requieren de ciertas estructuras de fondo para definir tales cantidades. Ver Lam (2011, Sección 7); Szabados (2009) y De Haro (2021, Sección 4), para más detalles.

cuando ciertas estructuras de fondo se encuentran disponibles, que estas se muestren útiles para ciertas situaciones físicas relevantes y que a partir de ellas podamos calcular y predecir exitosamente, no es suficiente para considerar a las nociones de energía-momentum (y masa) como fundamentales.

5.4.3. El realismo de Read

Revisemos, en último lugar, el punto de vista realista acerca de la existencia de ecuaciones de conservación adecuadamente definidas en la RG, cuyo principal defensor es Read (2020). Dicho muy aproximadamente, Read afirma que, desde una posición filosófica funcionalista acerca de las cantidades físicas, es posible defender la existencia de expresiones matemáticas que representen adecuadamente tanto la conservación integral como diferencial de la energía-momentum, así como también la existencia de la energía-momentum gravitacional en contextos relativistas generales¹¹⁹. Antes de avanzar hacia la presentación de los argumentos del autor, vamos a asumir, siguiendo justamente a Read (2020), la siguiente nomenclatura. En primer lugar, llamaremos modelos cinemáticamente posibles (MCP) de la RG a aquellos dados por la tripleta $(M, g_{\mu\nu}, T_{\mu\nu})$, donde M es la variedad diferenciable cuadri-dimensional, $g_{\mu\nu}$ es el campo métrico y $T_{\mu\nu}$ es el tensor energía-momentum de los campos de materia (recuérdese Sección 5.1). En segundo lugar, denominaremos modelos dinámicamente posibles de la RG (MDP) a aquellos MCP cuyos objetos geométricos satisfacen las ecuaciones de campo de Einstein (para más detalles acerca de esta distinción, ver Read 2020, Sección 2.1). Esta diferenciación entre los MCP y los MDP será relevante no solo para la presentación de los argumentos de Read en esta sección, sino que también para la discusión que desarrollaremos en la sección siguiente.

Veamos ahora en detalle los argumentos del autor. Comencemos con la ecuación (5.18). Read reconoce que esta ecuación, en efecto, no presenta la forma estándar de una expresión de conservación (como su homóloga (5.19) para la RE) pero, a diferencia de los dos puntos de vista anteriores, no es claro para el autor que la ecuación (5.18) no represente una legítima ley de conservación diferencial para los campos de materia debido a que, como afirman Hofer y Lam, los términos adicionales que aparecen en ella, respecto de la

¹¹⁹ Una revisión crítica de la posición funcionalista de Read puede encontrarse en Dürr (2019b).

ecuación (5.19), representan la interacción entre los campos de materia y el campo gravitacional sin, a su vez, tener en cuenta la energía-momentum de este último. Para Read, este argumento puede ser matizado teniendo en cuenta los siguientes dos razonamientos: por una parte, (i) asumir que la expresión (5.18) representa la interacción entre los campos de materia y el campo gravitacional omite toda una discusión perteneciente a los fundamentos de la RG sobre qué objeto matemático representa adecuadamente al campo gravitacional. Mientras algunos autores defienden que este estaría representado por el tensor métrico $g_{\mu\nu}$, otros abogarían por los coeficientes de conexión $\Gamma_{\mu\lambda}^{\nu}$ y otros por el tensor de Riemann $R_{\sigma\mu\nu}^{\rho}$ (Read, 2020, Sección 2.3). Por otra parte, (ii) que (5.18) represente la interacción entre los campos de materia y el campo gravitacional implica que sea cual sea el objeto que tome el lugar del tensor de energía-momentum en los MCP, representará tal interacción. Como esto no sería siempre cierto, se requiere una explicación adicional de por qué los términos extra en (5.18) representarían la interacción entre estos dos campos (Read, 2020, Sección 2.3). De este modo, Read no descarta que la expresión (5.18) pueda representar una legítima ley de conservación diferencial para la materia ordinaria.

Ahora bien, ¿cuál es la evaluación de Read respecto a la validez de la expresión (5.24) como una ley de conservación diferencial para la energía-momentum total? El autor argumenta que la legitimidad de la ecuación (5.24) dependerá de la posición que se asuma respecto a los objetos matemáticos de carácter no tensorial como $t^{\mu\nu}$ (Read, 2020, Sección 3.1). En efecto, si asumimos un punto de vista antirrealista¹²⁰ acerca de las cantidades no tensoriales, entonces el pseudo-tensor $t^{\mu\nu}$, como consecuencia de su característica de ser dependiente de coordenadas, no puede representar cantidades físicas relevantes, como la pretendida energía-momentum gravitacional. En consecuencia, desde esta posición, podemos concluir que la ecuación (5.24) no representa una legítima ecuación de conservación diferencial para la materia ordinaria, pues esta se encuentra compuesta de una cantidad no tensorial, como $t^{\mu\nu}$, que no puede representar entidades físicas. Por otro lado, si se asume un punto de vista realista acerca de los objetos matemáticos no tensoriales, es decir, si se asume que los objetos dependientes de coordenadas pueden representar también cantidades físicas relevantes, entonces la ecuación (5.24) puede ser postulada como una

¹²⁰ No confundir con los puntos de vista realista y antirrealista (eliminativista) acerca de la existencia de expresiones de conservación en la RG que estamos examinando en esta sección.

expresión diferencial válida para la conservación de la energía-momentum total. Bajo este último enfoque acerca del carácter ontológico de los objetos matemáticos de una teoría, que es defendido por Read, la energía-momentum gravitacional, representada por el pseudo-tensor $t^{\mu\nu}$, se definiría adecuadamente de acuerdo con un marco de referencia dado, de manera análoga a lo que sucede con el concepto de velocidad relativa en la mecánica Newtoniana (Read, 2020, Sección 3.1)¹²¹.

En tercer lugar, respecto a la expresión (5.27), Read señala que, si bien es una expresión que se encuentra definida bajo ciertas condiciones particulares pero muy útiles, representa una legítima ecuación de conservación integral para la energía-momentum total, pues es una expresión que no depende de la elección de un sistema de coordenadas. Por lo tanto, con independencia de la discusión anterior respecto a la interpretación que demos a las cantidades que son dependientes de un sistema de coordenadas, como $t^{\mu\nu}$, la ecuación (5.27) puede postularse como una expresión válida para la conservación de la energía-momentum en la RG (Read, 2020, Sección 3.2).

Ahora bien, equipado con la distinción anterior entre posiciones realistas y antirrealistas acerca de las cantidades dependientes de coordenadas, Read se propone refutar la posición eliminativista de Hofer. Revisaremos los argumentos del autor brevemente aquí, pues estos serán referidos en la siguiente sección. De acuerdo con Read, Hofer se niega a aceptar la validez de las ecuaciones (5.24) y (5.27) –aun cuando esta última expresión es independiente de la elección de un sistema de coordenadas–, pues este asume una posición antirrealista acerca del pseudo-tensor $t^{\mu\nu}$. En particular, respecto de la ecuación (5.27), Read se muestra en desacuerdo con los argumentos esgrimidos por Hofer

¹²¹ Dos comentarios respecto de los argumentos que esgrime Read para aceptar la validez de la ecuación (5.24). En primer lugar, debemos mencionar que Dürr (2019b) esgrime una posición diferente respecto de la dependencia de coordenadas de $t^{\mu\nu}$. El autor afirma que, en efecto, ni la relatividad de coordenadas ni el carácter no tensorial de este objeto matemático impide atribuirle un significado físico. Sin embargo, sostiene Dürr, la dependencia de coordenadas de los pseudo-tensores es una dependencia viciosa, pues las coordenadas preferidas no se encuentran en sintonía con las simetrías del espacio-tiempo. En segundo lugar, de acuerdo con Dürr (2019b), aceptar la validez de la expresión (5.24) implica también aceptar que la no unicidad del pseudo-tensor $t^{\mu\nu}$ (recuérdese la discusión acerca de este problema en la subsección anterior) no es problemática. Dürr no se encuentra de acuerdo con esta conclusión y afirma que esta representa más bien una esperanza y no un argumento, pues las investigaciones acerca de los eventuales beneficios (o no) de la no unicidad del pseudo-tensor $t^{\mu\nu}$ están en curso. Recordar Notas al Pie 35 y 36.

para negar la legitimidad de esta ecuación. Recordemos, brevemente, los argumentos de Hofer: (i) debido a que nuestro mundo actual no es asintóticamente Minkowskiano, la ecuación (5.27) no se aplica a la gravedad en nuestro mundo actual (2000, p. 194); (ii) aceptar la ecuación (5.27), debido a la presencia de $t^{\mu\nu}$, va en contra de la tendencia actual en física teórica de dar preferencia a las cantidades que son independientes de coordenadas por sobre las que no lo son (Hofer, 2000, pp. 194-195).

Read rechaza las afirmaciones (i) y (ii). Por un lado, respecto a la afirmación (i), el autor ofrece los siguientes argumentos. En primer lugar, (a) Read afirma que si bien el mundo no es asintóticamente Minkowskiano, esto no impide que la expresión (5.27) se aplique al mundo actual (recordemos a favor de Read que, de hecho, esta se aplica para el cálculo de la pérdida de energía en un sistema físico debido a las ondas gravitacionales). Por lo tanto, afirma Read, las dos partes componentes de la afirmación (i) parecen no estar relacionadas (2020, Sección 3.3). En segundo lugar, (b) el autor afirma que toda teoría física es una idealización, por lo que estas no se aplican al mundo actual tal cual es. En consecuencia, sostiene Read, la objeción (i) de Hofer se aplicaría a una cantidad intolerablemente alta de teorías físicas (2020, Sección 3.3). Por último, (c) Read señala que no debemos solo enfocarnos en el MDP cosmológico que modela nuestro mundo actual, sino más bien en la totalidad de los MDP's de la RG. Es suficiente, afirma el autor, con que en una subclase de modelos de la teoría se pueda definir una expresión independiente del marco de referencia (como la ecuación (5.27)) para concluir que existe una noción bien definida de energía-momentum gravitacional independiente del marco de referencia en la RG (2020, Sección 3.3). Por otro lado, respecto de la afirmación (ii) de Hofer, Read brinda el siguiente contraargumento: (d) la ecuación integral de conservación (5.27) implica una noción de energía-momentum gravitacional que es independiente de coordenadas, pues la expresión antes mencionada lo es. Por lo tanto, afirma Raed, la objeción (ii) de Hofer parece no tener cabida en este tipo de expresiones (2020, Sección 3.3).

Como una conclusión a la discusión anterior, Read afirma que tanto la ecuación (5.24) como la ecuación (5.27) son expresiones legítimas para la conservación de la energía-momentum total en la RG. En particular, si asumimos una posición realista acerca de las cantidades dependientes de un sistema de coordenadas, entonces la ecuación diferencial (5.24) es una expresión válida en todos los MDP's de la RG. Por otro lado,

independientemente del punto de vista realista o antirrealista asumido acerca de las cantidades dependiente de coordenadas, Read sostiene que la expresión integral (5.27) se encuentra bien definida cuando las estructuras de la planitud asintótica (campo vectorial de Killing) se encuentran presentes.

Un contraargumento a las afirmaciones (a), (b) y (c) de Read es el que presenta Dürr (2019b), quien señala que, para distinguir con claridad entre las posiciones de Hofer y Read respecto de los modelos asintóticamente planos de la RG, es necesario diferenciar entre dos tipos de representaciones científicas: aproximaciones e idealizaciones. Para hacer esta distinción, Dürr hace referencia a Norton (2012), quien sostiene que un modelo aproximado corresponde a una descripción inexacta de un sistema objetivo, mientras que en el caso de un modelo idealizado, este se interpreta como una descripción inexacta de algún modelo que intenta reproducir al sistema objetivo considerando algunas de sus características más relevantes. Luego, Dürr afirma que un modelo exitoso implica diferentes compromisos realistas dependiendo de si este se basa en una aproximación o en una idealización. En efecto, el autor señala que, mientras en general se asume que una aproximación otorga realismo a la totalidad del modelo, en una idealización, en cambio, solo podemos hablar de un realismo selectivo hacia los *working posits*¹²² de ese modelo (2019b, Sección 3.3).

El punto de Dürr es que Hofer abogaría por el hecho de que la planitud asintótica sea vista como una idealización, mientras que Read parece estar más inclinado a considerar la planitud asintótica como una aproximación o como una posición funcional de una idealización. La posición de Dürr al respecto es en favor de Hofer. Para el autor, la planitud asintótica corresponde a una idealización y no a una aproximación, pues aunque lejos de una fuente (una estrella, por ejemplo) el espacio-tiempo es aproximadamente plano y, por lo tanto, los cálculos y predicciones que se hagan en base a la expresión (5.27) son perfectamente válidos; el modelo cosmológico actual (el modelo Friedman-Lemaitre-Robertson-Walker, FLRW) nos indica que nuestro universo se encuentra en expansión, por

¹²² Por ahora, solo diremos que una forma de definir los *working posits* es afirmando que corresponden a aquellos enunciados de un modelo idealizado que contribuyen al éxito predictivo y explicativo de ese modelo. En contraposición a los *working posits*, tenemos a los *idle posits*, es decir, aquellos postulados ociosos de un modelo. En la sección siguiente discutiremos en detalle acerca de estos términos.

lo que difícilmente un modelo asintótico pueda ser considerado una aproximación (en el sentido anterior) de nuestro universo. Más bien, según Dürr, esta sería una extrapolación idealizada del espacio-tiempo que rodea a una estrella en un momento particular. Tal y como lo presenta el autor, el último recurso para ser realistas respecto de la planitud asintótica es que ciertas estructuras de dicha planitud exhiban ciertas características que se puedan asociar a un *working posit* del modelo. Sin embargo, Dürr afirma que la planitud asintótica no presenta ninguna característica que permita asociarla a un *working posit* (2019b, Sección 4.2.3). Por lo tanto, concluye el autor, sería más adecuado hablar de un eliminativismo, en vez de un realismo como argumenta Read, respecto de la energía-momentum gravitacional. Volveremos a esta discusión en la sección siguiente.

Ahora bien, volviendo a los argumentos presentados por Read en defensa de la expresión (5.27), podemos preguntarnos: ¿existe la energía-momentum gravitacional en la RG? El autor afirma que es posible responder positivamente a esta pregunta si asumimos un enfoque funcional acerca de las cantidades físicas, en particular, de la energía-momentum gravitacional. En efecto, si bien el autor concuerda con Hofer (2000) y Lam (2011) en que las cantidades físicas como la energía y la masa podrían no ser propiedades fundamentales en contextos relativistas generales –pues estas cantidades solo se encontrarían bien definidas en ciertos modelos de la RG en los que se cumple la ecuación (5.27)–, es posible defender la existencia de estas cantidades desde un punto de vista práctico considerando la utilidad que tienen las ecuaciones del tipo (5.27) en el cálculo y predicción de una diversidad de fenómenos físicos. De este modo, afirma Read, “...es plausible sostener que en situaciones como aquellas en las cuales la ecuación [(5.27)] se cumple, existe una cantidad en la RG que cumple el rol funcional de energía-momentum gravitacional” (2020, Sección 3.3.3). Habría dos razones, sostiene el autor, que justificarían la afirmación anterior. Para que una estructura matemática cumpla el rol funcional de la energía-momentum gravitacional, esta debe: (i) cumplir un rol análogo al que cumple la energía gravitacional en el sentido tradicional, esto es, como una cantidad que permite balancear la energía de la materia de un sistema, y (ii) tener alguna relación con los grados de libertad gravitacionales en la teoría. La ecuación (5.27) cumple, de acuerdo a Read, tanto con (i) como con (ii). Cumple con (i), pues podemos ver por comparación con la mecánica Newtoniana o clásica que la energía-momentum gravitacional en la RG tiene un papel

análogo a la energía gravitacional en la teoría de Newton: equilibrar la energía de un sistema de modo que su suma sea constante. Y cumple con (ii) debido a que esta expresión relaciona el pseudo-tensor $t^{\mu\nu}$ con los coeficientes de conexión, estos últimos, de acuerdo con cierta interpretación, relacionados con los grados de libertad gravitacional (2020, p. 20).

Veamos, en la siguiente sección, de qué manera las conclusiones a las que podamos arribar en el debate en torno a las leyes de conservación en la RG podrían impactar en la formulación de una teoría de la causación física en términos de posesión/transferencia/intercambio de energía, como la TCC de Dowe.

5.5. El problema de las leyes de conservación y la TCC

Recordemos que las proposiciones *CCI* y *CC2* se encuentran fundamentadas en expresiones como “poseer una cantidad conservada” o “intercambiar una cantidad conservada”. Este hecho implica que, para evaluar estas expresiones, es necesario contar con ciertas herramientas: las leyes de conservación integrales. En esta sección analizaremos el estado de estas leyes en la RG y estableceremos en qué sentido podemos hablar de su existencia; esto nos conducirá a postular la condición *Cond1*. De este modo, esta condición viene a imponerse desde los propios fundamentos de la RG al debate acerca del alcance y las limitaciones de la teoría de Dowe en la teoría de la gravitación de Einstein. Al finalizar, adelantamos brevemente algunas consecuencias ulteriores para la TCC.

Antes de presentar nuestros argumentos en relación al problema de las leyes de conservación, y dada la envergadura de este problema, evidente en el gran número de discusiones y variantes que presenta, se hace necesario acotar la discusión, al menos, en los aspectos que se mencionan a continuación. En primer lugar, debemos notar que el debate en torno a las leyes de conservación se da en dos niveles diferentes, pero relacionados. En un primer nivel, se discute si existen o no expresiones matemáticas que genuinamente puedan llamarse de “conservación” en el sentido de que estas representen adecuadamente el balance de energía-momentum entre los campos de materia y el campo gravitacional. Como una implicancia de esta discusión, también se debate acerca de la existencia de algún objeto matemático en el aparataje formal de la RG que capture adecuadamente la noción de

energía-momentum gravitacional (por ejemplo, el pseudo-tensor $t^{\mu\nu}$). En un segundo nivel, se discute si es posible otorgar a esta cantidad el estatus de “fundamental”¹²³ en el marco de esta teoría. Como se mencionó en las secciones anteriores, en el debate en torno a las leyes de conservación, estos dos niveles se encuentran conectados por el hecho de la fundamentalidad de una cantidad como la energía-momentum estaría dada en virtud de su conservación (respecto a la relación entre la conservación y la fundamentalidad de ciertas cantidades como la masa y la energía, ver Norton, 2015, Sección 3.5). En la discusión que sigue nos enfocaremos en el primero de estos dos niveles, dejando de lado el debate acerca de la fundamentalidad. Esta decisión se encuentra justificada por el hecho de que la plausibilidad de las proposiciones *CCI* y *CC2* de la teoría de Dowe depende de las conclusiones a las que arribemos respecto a la existencia (o no) de leyes o ecuaciones de conservación que permitan definir cantidades físicas conservadas, pero es independiente de si consideramos o no a la energía-momentum gravitacional como una cantidad fundamental.

Una segunda cuestión que es necesario delimitar respecto al problema de las leyes de conservación en la RG es la siguiente. En la Sección 5.3 hemos dividido este problema en la discusión acerca de la conservación de la energía-momentum de los campos de materia, por un lado, y en la discusión sobre la conservación de la energía-momentum total, que incluye la del campo gravitacional, por otro lado. Ahora bien, dado que pretendemos evaluar el alcance de la TCC en contexto relativistas generales, donde uno de los conceptos centrales de esta teoría es justamente el de “campo gravitacional”, nos centraremos únicamente en el examen de las expresiones de conservación que incluyan la energía de este campo, como las ecuaciones (5.24) y (5.27).

En adición a lo anterior, se hace necesario restringir aún más nuestro espacio de discusión en relación a este último aspecto. En lo sucesivo, nos enfocaremos en el análisis de las expresiones integrales de conservación, omitiendo sus homólogas diferenciales. Esta restricción se justifica por el hecho de que son las ecuaciones integrales de conservación las que permiten calcular el balance de energía-momentum en una región determinada del espacio-tiempo (y no solo en un punto infinitesimal, como en las diferenciales), requisito

¹²³ Respecto al término “fundamental”, recuérdese Nota al Pie 41.

necesario para la TCC (ver Curiel, 2000). En consecuencia, en lo que resta de esta sección analizaremos los argumentos que rodean a la expresión (5.27) únicamente.

Conviene mencionar aquí un asunto importante respecto de la relación entre las expresiones diferenciales e integrales de conservación que se tornará relevante en la discusión que ofreceremos en los capítulos siguientes. En general, se considera que las ecuaciones diferenciales de conservación tienen un carácter más fundamental que las leyes integrales de conservación, pues es posible que existan las primeras sin las segundas, pero no a la inversa (Read, 2020). En esta investigación adherimos a este punto de vista; por lo tanto, en la RG existirán modelos en los que solo se cumplen las leyes de conservación diferenciales y otros modelos (subconjunto del anterior) en los que serán válidas tanto las leyes diferenciales como integrales de conservación. En efecto, recordemos que la ecuación integral (5.27) se deriva de la ecuación diferencial (5.24) bajo restricciones particulares (asociadas a la planitud asintótica), lo cual implica que la expresión (5.24) se cumple en un conjunto de modelos de una cardinalidad mayor al conjunto de modelos en que se cumplen ambos tipos de ecuaciones de conservación. Retomaremos esta discusión en el siguiente capítulo.

El último sentido en el que se hace necesario acotar la discusión, se encuentra relacionado al enfoque técnico desde el cual se analiza el problema de las leyes de conservación en esta investigación. El enfoque que históricamente se ha asumido para analizar este problema es el enfoque pseudo-tensorial. Sin embargo, en las últimas décadas ha surgido una forma alternativa de encarar las dificultades que nos presenta la conservación de la energía-momentum. Este es el enfoque “cuasi-local”, que asocia la energía-momentum a dominios extendidos del espacio-tiempo, pero finitos (ver Szabados, 2009; De Haro, 2021, para una detallada revisión del programa cuasi-local)¹²⁴. Como ya el lector o la lectora habrá notado en las secciones anteriores, dado el objetivo central de esta tesis, aquí nos acercamos al debate en torno a las leyes de conservación en la RG desde el enfoque pseudo-tensorial, dejando de lado el programa cuasi-local.

¹²⁴ Estos dos enfoques no son completamente independientes. Existen expresiones de conservación cuasi-locales derivadas de expresiones pseudo-tensoriales (ver De Haro, 2021, Sección 5).

Una vez que nuestro espacio de análisis ha sido razonablemente acotado, volvamos nuestra atención hacia los argumentos que podemos brindar en el marco del debate presentado en la sección anterior. Empecemos por resumir las diversas afirmaciones que se han hecho respecto de la ley integral de conservación (5.27):

- (a) Aceptar (5.27), debido a la presencia de $t^{\mu\nu}$, implica alinearse con la posición que admite que los objetos dependientes de coordenadas pueden representar cantidades físicas significativas (Hofer, 2000).
- (b) Contra (a). (5.27) es una expresión válida, pues es independiente de coordenadas. Si $t^{\mu\nu}$ lo es o no, es otro asunto (Read, 2020).
- (c) (5.27) es una expresión válida, pero depende de la incorporación de ciertas estructuras de fondo (Lam, 2011).
- (d) El mundo no es asintóticamente Minkowskiano, por lo tanto, (5.27) no se aplica en el mundo actual (Hofer, 2000).
- (e) Contra (d). Que el mundo sea asintóticamente Minkowskiano no impide que (5.27) se aplique al mundo actual. De hecho, lo hace (Read, 2020).
- (f) Contra (d). Toda teoría física es una idealización, por lo que estas no se aplican al mundo actual en un sentido fuerte (Read, 2020).
- (g) Contra (d). Para decidir acerca de la existencia de una ecuación integral de conservación, no solo debemos enfocarnos en el MDP cosmológico actual, sino que en la totalidad de los MDP's de la teoría (Read, 2020).
- (h) Contra (e), (f) y (g), a favor de (d). Los modelos asintóticamente planos son idealizaciones, y de una idealización solo es posible hablar de un realismo selectivo hacia los *working posits* de ese modelo. Sin embargo, las estructuras de la planitud asintótica no son *working posits*, sino más bien *idle posits*. Por lo tanto, parece más adecuado hablar de un eliminativismo –en lugar de un realismo selectivo– acerca de la existencia de una ecuación de conservación integral para la energía-momentum gravitacional (Dür, 2019).

Podemos dividir estos argumentos en dos grupos. Por un lado, tenemos las afirmaciones (a), (b) y (c), que se relacionan al problema de si debemos aceptar o no como cantidades físicamente relevantes aquellas que sean dependientes de coordenadas, como $t^{\mu\nu}$. Por otro lado, tenemos las afirmaciones (d) a (h), que se encuentran asociadas al debate

en torno a la representación científica. A continuación, llevaremos a cabo una breve evaluación del segundo conjunto de afirmaciones con el objetivo de mostrar que el problema acerca de la existencia de leyes de conservación adecuadamente definidas en la RG y, en consecuencia, el problema de la existencia de la energía-momentum gravitacional, aún se encuentra abierto. En particular, nos interesa mostrar que el argumento (h) de Dürr (2019b) no es en ningún caso definitivo ni concluyente en favor de un antirrealismo o eliminativismo acerca de la energía-momentum gravitacional. Por el contrario, este argumento no solo deja completamente abierto el debate en torno a la existencia de esta cantidad física, sino que también muestra algunas luces para una posible vía de solución en relación a este problema. Queda pendiente, por supuesto, un abordaje del primer conjunto de afirmaciones respecto a la dependencia de coordenadas.

Con el objetivo de brindar una evaluación del argumento (h) lo más clara posible, retomemos el argumento de Dürr destacando las premisas que lo componen:

(h) (P1) Los modelos asintóticamente planos son idealizaciones y (P2) de una idealización solo es posible hablar de un realismo selectivo hacia los *working posits* de ese modelo. (P3) Sin embargo, las estructuras de la planitud asintótica no son *working posits*, sino más bien *idle posits*. (C) Por lo tanto, parece más adecuado hablar de un eliminativismo –en lugar de un realismo selectivo– acerca de la existencia de una ley de conservación integral para la energía-momentum gravitacional (Dürr, 2019b).

Comencemos evaluando la premisa (P1) del argumento de Dürr, que afirma que los modelos asintóticamente planos son idealizaciones (y no aproximaciones) del sistema objetivo que se pretende explicar. Como hemos mencionado en la sección anterior, el autor basa su argumento en la distinción que introduce John Norton (2012) entre aproximaciones e idealizaciones: mientras el referente de una aproximación coincide con el sistema objetivo, el referente de una idealización no coincide con él, sino más bien con otro sistema que se acerca en algunos aspectos al sistema objetivo. Esta caracterización de Norton en términos referenciales intenta capturar la diferencia que comúnmente es atribuida a estos modelos representacionales. Regularmente se afirma que una aproximación tiene cierta “cercanía” con el sistema objetivo en el sentido de que se incorporan en el modelo

aproximado ciertos parámetros de la realidad de una manera simplificada; en cambio, de una idealización se afirma que en el modelo se emplean además ciertas suposiciones que no se ajustan a la realidad.

Para poner la premisa (P1) en contexto, se debe mencionar, en primer lugar, que si bien el enfoque de Norton acerca de estos modelos representacionales captura bastante bien la diferencia entre ambos, este no es ni único ni exhaustivo (ver por ejemplo, McMullin, 1985; Suppe, 1989; Chakravartty, 2001; Weisberg, 2007; Frigg y Hartmann, 2020). En segundo lugar, en general se entiende en el debate acerca de la representación científica que en gran parte de los modelos científicos se emplean tanto aproximaciones como idealizaciones (entre otras formas de representación), por lo que hablar de un modelo idealizado o aproximado en un sentido puro, tal como lo hace Dürr, no es por completo preciso.

Ahora bien, hechas las aclaraciones anteriores, parece ser bastante claro, en acuerdo con Dürr, que la planitud asintótica es en efecto una idealización. Escapa a los alcances de esta investigación discutir las características técnicas de la planitud asintótica (ver, por ejemplo, Wald, 1984, Capítulo 11, para estos fines); en vez de ello, bastará con precisar que los espacio-tiempos asintóticamente planos representan sistemas aislados en la RG (ver Wald, 1984, p. 269). Y justamente los sistemas aislados, además de los planos sin fricción y las masas puntuales, son uno de los ejemplos paradigmáticos de las idealizaciones en física (Saatsi, 2016; Frigg y Hartmann, 2020). Por lo tanto, no parecen haber muchas dudas acerca de la veracidad de la premisa (P1) del argumento de Dürr.

Continuemos examinando la premisa (P2), que afirma que de un modelo idealizado solo podemos comprometernos con un realismo selectivo hacia los *working posits* de ese modelo. Si bien esta afirmación no es conflictiva en sí misma, podemos adelantar que, en el contexto del problema de las leyes de conservación en la RG, no goza de la misma plausibilidad que la premisa (P1) antes examinada, y esto por dos razones interrelacionadas.

En primer lugar, y quizás la razón más importante para poner en cuestionamiento esta premisa, es que el punto de vista realista selectivo es solo una de las maneras de abordar el problema del realismo (o antirrealismo) y su conexión con las representaciones científicas. Existen otros enfoques acerca de qué compromisos metafísicos podemos adquirir o no a

partir de los modelos científicos que se muestran exitosos en términos explicativos y predictivos. En efecto, la bibliografía en torno a la concepción semántica de las teorías científicas y su relación con el debate realismo/antirrealismo es bastante amplia. En el marco de este debate, la supuesta relación entre antirrealismo (o eliminativismo) e idealizaciones en ningún caso se encuentra zanjada. Por un lado, las posturas realistas afirman que, si bien los modelos no son literalmente verdaderos, estos lo serían al menos en términos aproximados. Algunas de estas posturas realistas basan su defensa, en particular respecto de las idealizaciones, en el hecho de que el éxito predictivo y explicativo de este tipo de representación se debería a que estas presentan algún tipo de conexión con la realidad (Saatsi, 2016). Algunos defensores realistas son Frederick Suppe (1989) con su visión cuasi-realista o modal acerca de las teorías científicas, que sostiene que los modelos nos muestran cómo sería el mundo si las suposiciones que contiene ese modelo fueran ciertas; y Newton da Costa y Steven French (2003), con el enfoque que afirma que las teorías científicas y sus modelos deben considerarse como parcialmente (y no totalmente) verdaderos (para otras defensas realistas acerca del enfoque de modelos de las teorías científicas, ver McMullin, 1985; Brzezinski y Nowak, 1992; Niiniluoto, 2007). En este contexto, el realismo selectivo abordado por Dürr en su argumento es un punto de vista más, entre otros, que pretende responder a los ataques antirrealistas del enfoque de modelos de las teorías científicas.

La segunda razón para poner en cuestionamiento la premisa (P2) tiene relación con la noción de *working posit*. En términos generales puede decirse que los *working posits* de un modelo son aquellas estructuras que tienen algún tipo de correlato con la realidad. Por otro lado, los *idle posits* son considerados como aquellas estructuras que no presentan ningún tipo de conexión con la realidad, pues solo existen para dar sentido al modelo. Sin embargo, no existe un consenso entre filósofos y filósofas de la ciencia acerca de qué debemos entender *exactamente* por cada una de estas nociones (ver por ejemplo, Kitcher, 1993, p. 149; Psillos, 1996, p. S311), de modo que podamos en efecto identificar en un modelo científico qué estructuras corresponden a cada categoría (por ejemplo, Stanford 2006; Vickers, 2017).

Por último, evaluemos la premisa (P3) de Dürr, que afirma que la planitud asintótica no es un *working posit* de un modelo, sino más bien un *idle posit*, es decir, un enunciado

ocioso de la teoría que no tiene ninguna correlato en la realidad. Un primer punto a destacar acerca de la premisa (P3) es que su verdad o falsedad se encuentra condicionada a la definición de los términos *working/idle posits*, que como vimos más arriba, es un tema de debate en la actualidad. En segundo lugar, y más importante aún, creemos que Dürr no ha argumentado adecuadamente en favor de la premisa (P3). Revisemos su argumento brevemente. Por un lado, el autor afirma que el comportamiento en el infinito es la característica sobresaliente de la planitud asintótica. Dürr no aclara qué quiere decir con esta afirmación, pero esta parece estar fundamentada en el hecho de que las condiciones que impone la planitud asintótica dependen de ciertas suposiciones acerca del comportamiento del espacio-tiempo en el infinito. Por otro lado, el autor afirma que los *working posits* de los modelos asintóticamente planos exitosos son el comportamiento de caída correcto hasta escalas cosmológicas y no su comportamiento de caída en el infinito. Esta afirmación tiene relación con el hecho de que, en los modelos asintóticamente planos, lejos de la fuente pero más cerca que las escalas de orden cosmológicas, el espacio-tiempo puede ser considerado plano para todo efecto práctico. Sobre la base de estas dos afirmaciones, Dürr concluye que la planitud asintótica es un *idle posit*.

Sin embargo, para demostrar que la planitud asintótica es un postulado ocioso de algunos modelos de la RG, es necesario argumentar que ninguno de sus enunciados tiene un referente en la realidad. Y no queda claro en qué sentido el argumento antes presentado muestra algo como esto. De hecho, tal como Dürr presenta su argumento, el problema de si la planitud asintótica es un *working* o un *idle posit* pareciera tener que ver con cómo definimos este concepto, lo cual no parece ser el camino correcto.

Creemos que es posible formular algunos argumentos en defensa de la planitud asintótica como un *working posit*. Algunas luces acerca de estos argumentos podrían ser obtenidas al considerar la caracterización de Stathis Psillos (1996, 2018) acerca de este término, quien sostiene que los *working posits* son aquellas afirmaciones teóricas que contribuyen sustancialmente al éxito explicativo y predictivo de la teoría. Pareciera ser que, en algún sentido, la planitud asintótica cumple con esta característica, pues gracias a las suposiciones que esta impone es posible, por ejemplo, obtener mediciones correctas de la energía perdida a través de ondas gravitacionales en un sistema físico, como fue mencionado más arriba en este capítulo.

Considerando el análisis de las premisas (P1) a (P3) efectuado en los párrafos anteriores, debemos concluir que no es claro que no exista una ley de conservación integral para la energía-momentum gravitacional en la RG debido a que esta solo puede ser correctamente definida en los modelos asintóticamente planos de la teoría. De hecho, como intentamos mostrar, existen varios caminos disponibles para argumentar en favor de la legitimidad de la ecuación (5.27).

Ahora bien, alguien podría preguntarse qué sentido tiene defender la existencia de una ecuación integral de conservación (y consecuentemente la existencia de la energía-momentum gravitacional) en modelos de la RG que no se corresponden con el mundo actual. O, dicho de otro modo, los modelos en los que existen leyes integrales de conservación bien definidas no son aquellos que manifiestan las características que parece tener nuestro mundo actual, ¿qué importancia o validez tiene entonces la existencia de leyes integrales de conservación en estos modelos? Esta es un cuestionamiento directamente relacionado con el argumento de Hofer presentado más arriba, que, recordemos, sostiene que la ecuación (5.27) no se aplica a nuestro mundo actual puesto que este no es asintóticamente Minkowskiano como exige esta ecuación para encontrarse bien definida.

Sin embargo, creemos que es posible responder a esta objeción haciendo uso de algunas ideas provenientes del enfoque semántico de las teorías científicas, y considerando un punto de vista posibilista de las teorías físicas. A diferencia de la visión sintáctica, que caracteriza las teorías científicas como reconstrucciones axiomáticas a partir de un conjunto de oraciones en un determinado lenguaje lógico, la visión semántica sostiene que las teorías científicas quedan mejor caracterizadas por el conjunto de sus modelos. Bajo este último enfoque, una teoría (en este caso física) no es más que el conjunto de los modelos que de ella se puedan desprender. Estos modelos representarán en menor o mayor medida ciertos rasgos de nuestro mundo actual dependiendo de las características de las representaciones científicas (aproximaciones e idealizaciones, por ejemplo) utilizadas en los modelos. De este modo, algunos modelos serán “más parecidos” al mundo actual y otros menos, pero ninguno de ellos representará *exactamente* a nuestro mundo. Este es un hecho más o menos acordado: las teorías científicas son solo aproximadamente verdaderas. Además, para explicar y predecir los fenómenos naturales se usan uno o varios modelos de una teoría. Por ejemplo, se usa el exitoso modelo FLRW para modelar nuestro universo a gran escala, por

un lado, y los modelos asintóticamente planos para calcular la energía perdida como ondas gravitacionales en un sistema binario, por otro lado. Considerando las aclaraciones anteriores, no parece adecuado afirmar, por lo tanto, que únicamente aquel modelo que más se acerque a la realidad será el que nos informe respecto a los compromisos ontológicos implicados por la teoría.

Adicionalmente, debemos notar lo siguiente. Detrás de la afirmación de que la energía-momentum gravitacional no existe en la RG debido a que la ecuación (5.27) solo se encuentra bien definida en aquellos modelos que no exhiben las características de nuestro mundo actual, se esconde una visión actualista de las teorías físicas, la cual afirma que solo existe aquello que es actual. Sin embargo, bajo una visión posibilista de las teorías físicas, podemos arribar a una conclusión bastante diferente respecto a la existencia de la energía-momentum gravitacional. En efecto, haciendo uso del lenguaje de mundos posibles podemos afirmar que cada modelo de la RG puede ser visto como un mundo físicamente posible¹²⁵. Algunos de estos mundos podrán ser más similares al nuestro que otros, pero todos son posibilidades físicas dadas por las ecuaciones de campo Einstein de la RG. En este sentido, podemos afirmar que las teorías físicas nos informan de aquello que es posible si nuestro mundo exhibiera las características dadas por el modelo en cuestión. Las teorías son, en consecuencia, el conjunto de todas estas posibilidades. Bajo esta mirada, ¿existe la energía-momentum gravitacional? Deberíamos responder afirmativamente a esta pregunta, pues la energía-momentum gravitacional se encuentra correctamente definida en un modelo posible de la RG. Estaríamos hablando de una existencia posible, pero existencia, a fin de cuentas.

Los argumentos esbozados anteriormente no pretenden constituirse como una defensa concluyente en favor de la existencia de la energía-momentum gravitacional. Nos sentiríamos conformes, en vez de ello, si esta discusión es suficiente para convencer al lector o a la lectora de que el debate acerca de los compromisos ontológicos que se puedan atender o no a partir de las representaciones utilizadas en los modelos científicos, como las idealizaciones y las aproximaciones, se encuentra completamente abierto. Como consecuencia, incluso aceptando que los modelos asintóticamente planos son modelos

¹²⁵ En el próximo capítulo abordaremos en detalle la noción de posibilidad física.

idealizados, tenemos una base conceptual y argumentativa suficientemente razonable para elaborar algún tipo de defensa en favor de la existencia de la energía-momentum gravitacional en la RG.

Sobre la base de los argumentos presentados en este capítulo, desde el debate acerca de los fundamentos de la RG hasta el problema de las leyes de conservación, es posible presentar la siguiente condición que debería ser considerada por cualquier teoría de la causación que pretenda formularse en términos de cantidades físicas conservadas:

Condl: En las teorías físicas del espacio-tiempo, en particular en la teoría general de la relatividad, las leyes de conservación integrales se cumplen en una cierta clase de sus modelos.

La *Condl* asegura que, al menos en una cierta clase de modelos de la RG, contamos con genuinas leyes integrales de conservación y, por lo tanto, en esos modelos existen cantidades conservadas bien definidas.

Ahora bien, ¿qué sucede con la TCC si resultara que la *Condl* no es una condición plausible porque se demuestra con argumentos concluyentes que no existen genuinas leyes de conservación integrales en la RG? O en particular, ¿qué podemos decir de la TCC en aquellos modelos de la RG en que no contamos con leyes integrales de conservación? ¿Diríamos en estos casos que la TCC pierde todo sentido, es decir, diríamos que se convierte en una teoría falsa, o más bien deberíamos concluir que en estos casos no hay causación? Una respuesta a esta pregunta tendrá que esperar hasta el Capítulo 7. Sin embargo, podemos adelantar que las consecuencias para la TCC de no contar con leyes de conservación integrales bien definidas (ya sea en algunos modelos de la RG o en todos) dependerá de si las proposiciones *CCI* y *CC2* son consideradas como caracterizaciones o como definiciones de los términos involucrados en ellas.

Capítulo 6. *Metafísica de la causación I:* *el problema de la posibilidad física*

Resumen

En el capítulo anterior hemos abordado *el problema de las leyes de conservación* y hemos postulado la condición *Condl*. Esta establece que, en la RG, las leyes de conservación en su forma integral se cumplen en una cierta clase de sus modelos. Este es el primero de los tres problemas que analizamos en esta investigación con el objetivo de intentar determinar los alcances y las limitaciones de la TCC en los escenarios relativistas generales. En este capítulo se aborda el primero de los dos problemas restantes, *el problema de la posibilidad física*.

El problema de la posibilidad física, en términos generales, consiste en intentar dilucidar, tanto en los debates acerca de los fundamentos de la física como en los debates en filosofía de la física, la noción frecuentemente utilizada de “mundo físicamente posible”. El Enfoque Estándar acerca de la posibilidad física sostiene, en palabras sencillas, que lo físicamente posible está determinado por aquello que dictan las ecuaciones dinámicas de una teoría. De este modo, cualquier modelo de la RG se debería considerar como una solución físicamente posible. Sin embargo, esto no se ajusta a lo que efectivamente ocurre en la práctica y la teorización científica. En este sentido, argumentaremos que, en muchas ocasiones fácilmente rastreables, entran en juego otras variables, más allá de las ecuaciones dinámicas, para delimitar la posibilidad física. Entre ellas podemos contar ciertos principios metafísicos, matemáticos y físicos.

La inclusión de las leyes de conservación para caracterizar lo físicamente posible es especialmente relevante para evaluar las limitaciones de la teoría de la causación de Dowe en los escenarios relativistas generales. En efecto, sobre la base de una adecuada argumentación, podremos postular la siguiente condición que sería deseable imponer a una teoría de la causación en términos de cantidades físicas conservadas que pretenda ser aplicable en la RG:

Cond2: La noción de posibilidad física debe incluir ciertos elementos, adicionales a las ecuaciones dinámicas, que se encuentren en concordancia con la práctica y la teorización física. En particular, debe incluir, al menos, las leyes de conservación (en su forma diferencial o en su forma diferencial e integral).

Las implicancias de esta condición para la noción de posibilidad física en el marco de la RG son claras. La imposición de la *Cond2* trae como consecuencia que los mundos físicamente posibles de esta teoría no sean todos los modelos que se derivan de las ecuaciones de Einstein, como postula el Enfoque Estándar. Para que los modelos de la RG merezcan la etiqueta de “mundo físicamente posible”, estos deben presentar adecuadas leyes de conservación, ya sea únicamente en su forma diferencial o en su forma diferencial e integral.

Tomando como referencia para el análisis la condición anterior, evaluaremos las consecuencias para la TCC. Mostraremos que la imposición de las condiciones *Cond1* y *Cond2* simultáneamente trae consigo algunas consecuencias importantes para la TCC. Una primera consecuencia es que las proposiciones *CCI* y *CC2* de la teoría de Dowe pueden ser aplicadas en un conjunto de mundos físicamente posibles de la RG, esto es, en aquellos que cumplen con las leyes de conservación tanto en su versión diferencial como en su versión integral. Sin embargo, la aplicación conjunta de ambas condiciones también trae como consecuencia que en ciertos modelos o mundos físicamente posibles de la RG no existan ecuaciones de conservación integrales, las cuales son necesarias para definir cantidades conservadas. ¿Qué sucede con la TCC en estos mundos posibles? Una respuesta a esta pregunta será ofrecida en el capítulo siguiente¹²⁶.

6.1. El problema de la posibilidad física

Las teorías físicas no son meramente un repertorio de descripciones de lo que sucede en el mundo actual, sino que pretenden tener un ámbito de aplicación mucho más amplio. Tales

¹²⁶ La discusión llevada a cabo en este capítulo se encuentra basada en el siguiente trabajo: Herrera, M. y López, C (2020). “¿Qué hace físicamente posible a un mundo posible?”. *Principia: An International Journal of Epistemology*, 24(1), 65–88.

teorías no se limitan, simplemente, a describir cómo la evolución de un sistema físico concreto tuvo lugar en un laboratorio en particular, en un momento del día específico y bajo cierta configuración de parámetros experimentales, sino que pretende describir de igual manera una plétora de situaciones *posibles*, con ligeras o más sustanciales variaciones¹²⁷. De hecho, se supone que nuestras teorías físicas son capaces de describir y funcionar igualmente bien en escenarios muy diferentes a los del mundo actual. Y en la misma medida en la que ellas describen (o prohíben) ciertas situaciones posibles, también se pronuncian respecto de otras. Para expresar este rasgo tan relevante de la física, se suelen utilizar nociones modales tales como “es *físicamente imposible* para un sistema físico alcanzar una velocidad superior a la velocidad de la luz”, “es *físicamente posible* un universo que no hubiese dado lugar a la vida en la Tierra”, “es *físicamente necesario* que la carga del electrón sea negativa”. Tales proposiciones dicen algo fundamental acerca de nuestro mundo y otros posibles. Pero, ¿qué significa que algo sea *físicamente* posible, imposible o necesario?

En la bibliografía filosófica existe un enfoque sumamente extendido para dar una respuesta a estas preguntas. Este enfoque acerca de la modalidad, al que se denomina “Enfoque Estándar”, sostiene que la noción de posibilidad *física* viene dada por la estructura nomológica de las teorías físicas: ser *físicamente* posible es estar de acuerdo con las leyes de la física, ser *físicamente* imposible es estar prohibido por las leyes de la física y ser *físicamente* necesario es ser exigido por las leyes de la física. La manera usual de introducir este enfoque es en términos del poderoso y ya familiar lenguaje de “mundos posibles”. Aunque introducido en el vocabulario filosófico durante la Modernidad gracias a Leibniz, la noción de mundo posible adquirió un extraordinario y exitoso desarrollo en la metafísica analítica tradicional, así como en lógica, principalmente a través de la propuesta de semánticas para las lógicas modales. Durante la segunda mitad del siglo XX, la noción de mundo posible se adentró también en el ámbito de la filosofía de la física, y colaboró a apuntalar conceptualmente muchas nociones, entre ellas, la de posibilidad física tal como es abordada por el Enfoque Estándar.

¹²⁷ Por ejemplo, para un estudio del uso de razonamientos contrapuestos en física, ver Wilson (2021).

Sin embargo, podríamos preguntarnos, ¿resulta adecuado este enfoque a la hora de caracterizar la noción de posibilidad física tal como se suele utilizar y aplicar en los múltiples contextos científicos de la física? En este capítulo analizamos el Enfoque Estándar de la modalidad física para abordar la noción de posibilidad física a la luz de la práctica científica. Argumentaremos que, en muchas ocasiones, este enfoque resulta demasiado permisivo a la hora de recoger y sistematizar muchos de los supuestos modales presentes en la comunidad científica. En particular, argumentaremos –sobre la base de algunos ejemplos– que el Enfoque Estándar permite ingresar al “reino de la posibilidad física” ciertas situaciones que usualmente los físicos dejan fuera. Esto conduce a considerar dos aspectos centrales. En primer lugar, la noción de posibilidad física es más compleja que lo que el Enfoque Estándar dictamina. Y, en segundo lugar, resulta necesario *algo más*, además de las leyes físicas, no solo para caracterizar adecuadamente la posibilidad física de manera conceptual, sino también para poder expresar con mayor precisión y fuerza el contenido modal de las teorías físicas.

La elucidación del *problema de la posibilidad física*, caracterizado en los párrafos anteriores, es de particular relevancia en el análisis del problema general que abordamos en esta investigación respecto del alcance que podría tener una teoría de la causación física en contextos relativistas generales. En efecto, si bien muchos de los argumentos formulados por los críticos de la TCC respecto de la compatibilidad entre esta teoría y la RG descansan en una formulación basada en mundos posibles, gran parte de sus autores (sino todos) asumen implícitamente, y sin ningún tipo de cuestionamiento, el Enfoque Estándar acerca de la posibilidad física. Por otra parte, Dowe tampoco ofrece, ni en la formulación de su teoría ni en las respuestas a sus críticos, un análisis acerca de este problema. La dificultad que presenta esta actitud reside en que el Enfoque Estándar, que considera únicamente las leyes dinámicas de una teoría como generadoras de posibilidades físicas, puede ser cuestionado, como hemos adelantado más arriba.

Sería deseable, entonces, considerar la discusión respecto de la posibilidad física en el análisis acerca de los límites de una teoría de la causación como la TCC en la RG. Esto lo haremos formulando una condición que se desprende de la propia argumentación que ofrecemos en este capítulo en relación al problema de la posibilidad física. Esta condición,

que denominaremos *Cond2*, nos proveerá un marco de referencia desde el cual evaluar estos límites.

Sobre esta base, el resto del presente capítulo se estructura del siguiente modo. En primer lugar, en la Sección 6.2 presentaremos una breve introducción acerca del problema de la modalidad. Luego, en la Sección 6.3 analizaremos el Enfoque Estándar acerca de la posibilidad física, exponiendo sus principales características y también sus limitaciones. A continuación, en la Sección 6.4 argumentaremos que, a diferencia de lo que plantea el Enfoque Estándar, en la práctica y teorización física se incluyen ciertos factores extra-nómicos al momento de determinar qué debemos entender por “físicamente posible”. Posteriormente, en la Sección 6.5 presentaremos dos casos de estudio para intentar ilustrar nuestros argumentos. Finalmente, en la Sección 6.6 relacionaremos la discusión que hemos presentado en las secciones anteriores con las dificultades que presenta la TCC en la RG, y postularemos la *Cond2*. Evaluaremos las consecuencias que impone esta restricción a la teoría de la causación de Dowe en el marco de la teoría de la gravitación de Einstein.

6.2. El reino de la modalidad

Expresiones modales como “*x* es necesario” o “es posible que *y*” son ubicuas tanto en el lenguaje filosófico como en el científico. Mientras que existe una enorme variedad de hechos que consideramos *actuales*, es decir, hechos que ocurren efectivamente de una manera determinada y no de otra, también se hace referencia cotidianamente a situaciones que, por ejemplo, o bien podrían haber ocurrido de otra manera (“es *posible* que el ejército soviético no hubiese vencido al ejército alemán en la Segunda Guerra Mundial”), o bien que *nunca podrían* ocurrir (“es *imposible* recuperar información de un agujero negro”). Sin embargo, en la misma medida en la que recurrentemente se recae en nociones modales para expresar diversas proposiciones, todas ellas no son juzgadas de la misma manera: algunas afirmaciones tendrían una “fuerza modal” mayor que otras, como si existiera cierta gradualidad en relación a su aceptación o rechazo. No es lo mismo decir “es imposible que 1+1 sea 3” que decir “es imposible que la velocidad de la luz sea un setenta por ciento menor”. O, “es imposible que cuerpo y alma sean dos sustancias distintas”. Algunas de estas proposiciones parecen ser modalmente más o menos exigentes, en tanto su rango de

aplicación resulta más o menos amplio. Estas intuiciones acerca del discurso modal exigen una serie de distinciones que permitan esclarecer y precisar grados y alcances de los términos modales, abriendo las puertas a un verdadero “*reino de la modalidad*”.

6.2.1. La naturaleza de las cosas y la evidencia relativa a un agente: entre lo epistémico y lo metafísico

A la hora de comenzar a trazar la geografía de este reino de la modalidad, primero es necesario disponer de las herramientas adecuadas. Si bien existen divergencias acerca de cuáles sería tales herramientas, muchos análisis de nociones modales en metafísica, filosofía del lenguaje e incluso filosofía de la física (el campo de interés en este trabajo) recaen fuertemente en la noción de “mundo posible”. Probablemente sean los éxitos alcanzados en el ámbito de la lógica modal y la filosofía del lenguaje (en gran medida debidos a Kripke, 1959, 1963) los que catapultaron el vocabulario de mundos posibles a la metafísica analítica y a buena parte de la filosofía de la física (además de Kripke, ver Lewis, 1973a; Earman, 1986; para enfoques filosóficos sobre nociones modales que prescindan de la noción de “mundo posible”, ver Fine, 1994; Jubien, 1996, 2009). De esta manera, decir “es *posible* que *p*” o “es *necesario* que *p*” debe ser entendido en términos de “existe un mundo posible donde *p* es verdadero” y “en todo mundo posible *p* es verdadero”, respectivamente. No se profundizará en las particularidades lógicas y semánticas que atraviesan la noción de mundo posible (para una introducción, ver Humberstone, 2015) ni tampoco en las discusiones metafísicas a las que ha dado lugar como, por ejemplo, si tales mundos posibles deben ser entendidos como existentes en igual grado que el mundo actual (posición comúnmente llamada “concretismo”, ver Lewis, 1986b), o meramente como ficciones útiles para esclarecer intuiciones semánticas y lingüísticas (“abstraccionismo”, ver Plantinga, 1974; Kripke, 1980)¹²⁸. Puesto que la noción de “*físicamente* posible” de acuerdo con el Enfoque Estándar, promocionada principalmente por John Earman (1986) en la filosofía de la física, recae en buena medida en una estructura de mundos posibles, esta estructura será asumida en lo que resta de la presente investigación.

¹²⁸ Ver el artículo “Two concepts of possible worlds” de Van Inwagen (1986) para mayores detalles sobre esta distinción.

En buena medida, trazar la geografía del reino de la modalidad consiste, por un lado, en distinguir *el alcance* de las afirmaciones modales (alcance expresado en términos de los constituyentes fundamentales de este reino que son los mundos posibles), y, por el otro lado, en *cualificar* tal alcance. Se refinará, en primer lugar, el último punto antes de continuar con el primero.

En la bibliografía existen al menos dos maneras de *cualificar* el alcance de las intuiciones modales respecto de diferentes tipos de afirmaciones. Intuitivamente, la noción de mundo posible permite hablar de manera teórica y sistemática de situaciones o contextos diferentes al actual. Pero este “hablar de” se puede introducir de dos maneras diferentes: en términos *epistémicos* o en términos *metafísicos* (Kment, 2017; Soames, 2011). De un modo general, los mundos *epistémicamente* posibles permiten establecer la verdad o falsedad de ciertos enunciados modales sobre la base de la evidencia disponible (en el mundo actual) para un agente determinado. Por lo tanto, las nociones de posibilidad y de necesidad están estrechamente delimitadas por el tipo y el grado de evidencia al alcance, abriendo la puerta a que las divergencias en el tipo y el grado de evidencia disponible conduzcan a diferencias en lo que es considerado posible o necesario. En particular, se dice que una proposición p es *epistémicamente* necesaria para un agente determinado A , si (i) la evidencia empírica con la que cuenta A y (ii) el razonamiento puro es suficientes para descartar $\neg p$ (Kment, 2017; Soames, 2005). Lo relevante de esta manera de entender la posibilidad y los mundos posibles es que no se compromete con algún tipo de modalidad intrínseca en las cosas o en las situaciones del mundo, sino que la modalidad es siempre relativizada al tipo de información que un agente posee.

Sin embargo, algunos autores se comprometen con una visión de la modalidad que no está relativizada a la evidencia que un agente posee, sino que se encuentra determinada por la naturaleza misma de las cosas (ver, por ejemplo, Ellis, 2001). Cuando se dice que es posible que el mundo tuviese cierta propiedad que no tiene el mundo actual, o que es posible que cierto arreglo de cosas estuviese dispuesto de otra manera, no se está haciendo una afirmación respecto del conocimiento de ciertos agentes sobre la base de cierta información, sino que se está haciendo una afirmación respecto de cómo el mundo, *en sí mismo*, podría haber sido. Desde esta perspectiva, entonces, las nociones modales permiten echar luz sobre la naturaleza misma de las cosas en el sentido de permitir la identificación,

por ejemplo, de aquellas propiedades que permanecerían fijas ante diferentes variaciones. En este sentido, la noción de necesidad se articula con la noción de esencia: si decimos que la esencia de x es “ser F ”, entonces decimos que en todo mundo posible en el cual x exista, x instancia la propiedad F . Por lo tanto, se puede afirmar que una proposición p es *metafísicamente* necesaria, si p es verdadera en todo mundo posible; y el contenido modal de p dice algo intrínseco respecto de las naturalezas de las cosas involucradas en p (por ejemplo, que la esencia de x es instanciar la propiedad P donde sea que x exista).

Existe cierta discusión en la bibliografía respecto de si la distinción entre mundos *epistémicamente* posibles y mundos *metafísicamente* posibles es legítima, o si, finalmente, confluyen en uno de los dos puntos (ver, por ejemplo, Chalmers y Jackson, 2001). No se ahondará en esta discusión aquí, y simplemente asumiremos que cuando se habla de si x es posible o necesario (en algún sentido a especificar ulteriormente) no se está relativizando la afirmación a la información o evidencia que un agente determinado pueda tener respecto de x , si no que se está haciendo referencia a su acepción metafísica.

6.2.2. Lo X-mente posible

Ahora es momento de abordar el primer punto mencionado más arriba: las distinciones respecto del *alcance modal* de las afirmaciones. En la bibliografía acerca de mundos posibles, es usual considerar que el espacio maximal de mundos posibles está dado por los mundos *lógicamente* posibles. Relativo a este espacio, es posible construir una infinidad de proposiciones que pueden ser *lógicamente necesarias*, *lógicamente imposibles* o *lógicamente posibles*. Una proposición lógicamente necesaria es comúnmente considerada una *verdad necesaria* o una *verdad lógica*, en el sentido de que es verdadera en todo mundo lógicamente posible: no existe un mundo posible en el cual la proposición sea falsa. Una proposición que, en su lugar, es *lógicamente imposible*, como por ejemplo una contradicción lógica, es falsa en todo mundo lógicamente posible: no existe un mundo posible en el cual tal proposición sea verdadera. Una proposición *lógicamente posible* refiere a una proposición que es verdadera al menos en un mundo posible¹²⁹.

¹²⁹ Sin embargo, sería un error considerar que tal espacio maximal de mundos lógicamente posibles permanece fijo e inalterable. Como es bien sabido, existen varios sistemas lógicos que no solo extienden la lógica clásica (como la lógica modal), sino que además no incluyen entre sus principios

De acuerdo con el párrafo anterior, el alcance de las afirmaciones modales está dado en términos de cuantificadores universales y existenciales que se aplican al espacio maximal de mundos posibles. En particular, la posibilidad está dada por la cuantificación existencial aplicada sobre el espacio maximal de mundos posibles. La idea central de trazar el mapa del reino de la modalidad depende, en buena medida, de si se considera que tal cuantificación sobre mundos es acotada o no.

Para algunas afirmaciones filosóficas y científicas (las más interesantes para los fines de esta investigación), la noción de posibilidad lógica puede resultar poco esclarecedora por ser demasiado permisiva. Muchas de las afirmaciones modales son mucho más específicas y acotadas: cuando los físicos dicen que es *imposible* extraer información de un agujero negro, probablemente no quieren decir que la proposición “se extrae información de un agujero negro” constituya una contradicción lógica. Es muy simple imaginar un mundo *lógicamente* posible, pero con otras leyes físicas, donde se pueda extraer información de un agujero negro. Cuando se dice “la esencia del oro es ser el elemento número 78 en la tabla periódica” y que por lo tanto es *necesario* para ser oro tener la propiedad “ser el elemento 78 de la tabla periódica”, no se está afirmando una verdad lógica, válida en todos los mundos *lógicamente* posibles. Por lo tanto, para poder especificar el contenido semántico modal de tal clase de afirmaciones, resulta conveniente *restringir* el dominio del espacio maximal de mundos posibles al cual nos referiremos con términos como necesidad, imposibilidad o posibilidad. En otras palabras, al afirmar “es necesario que *p*” se puede querer restringir la cuantificación sobre mundos lógicamente posibles a un sub-conjunto acotado. Con esta estrategia, es posible generar una geografía del reino modal mucho más detallada: lo *metafísicamente* posible, lo *físicamente* posible, lo *biológicamente* posible, lo *X-mente* posible, según el criterio que sea seleccionado.

Bajo este enfoque, “*físicamente* necesario” es solo una manera de expresar muchas afirmaciones modales más débiles que, por ejemplo, lo *lógicamente* necesario, pero más fuertes que una simple contingencia. Entre cualquier principio o verdad de la lógica clásica

lógicos principios básicos de la lógica clásica (como el Principio de Explosión en algunas lógicas para-consistentes, ver Priest, 2008). Lo relevante aquí es que lo *lógicamente posible* está relativizado al sistema de lógica que utilicemos para evaluar si la verdad de una proposición es necesaria, imposible o posible.

y la proposición “el libro que está sobre mi mesa es de color azul” parece haber una plétora de proposiciones con diversa fuerza modal. Las proposiciones *físicamente* necesarias, por ejemplo, parecen hallarse en este espacio intermedio y resulta filosófica y científicamente relevante circunscribir con precisión cuál es su dominio de aplicación y cuál es su naturaleza. En otras palabras, ¿qué es lo que hace *físicamente* posible a un mundo posible?

6.3. La posición tradicional: el Enfoque Estándar de la modalidad física

La introducción del lenguaje modal en las ciencias físicas ha sido de vital importancia, y la adopción de lo *físicamente* posible, lo *físicamente* necesario y lo *físicamente* imposible ha ganado su merecido espacio en este reino de la modalidad. Desde posiciones *esencialistas científicas* (como la sostenida por Ellis, 2001 o Putnam, 1975), lo *físicamente* (o *científicamente*) necesario colapsaría con lo *metafísicamente* necesario, ya que es tarea de la actividad científica el descubrimiento de las propiedades esenciales del mundo natural. Haciendo a un lado esta visión particular, lo relevante a discutir en esta sección es que la noción de posibilidad física, expresada en términos de mundos posibles, es de uso corriente no solo en el ámbito de la filosofía de la física, sino también en el ámbito científico general.

Si bien la noción de posibilidad física puede ser entendida de diferentes maneras, tal como mencionamos en la introducción, existe un enfoque ampliamente extendido en la bibliografía filosófica que entiende lo físicamente posible exclusivamente en términos de la estructura nomológica de las teorías físicas. Este enfoque, al que en esta tesis denominamos Enfoque Estándar, asume que una proposición del tipo ‘*x* es físicamente posible’ significa que existe al menos un mundo posible en el cual, de acuerdo con las leyes de nuestras mejores teorías físicas vigentes, *x* es el caso. En otras palabras, la ocurrencia de *x* no está prohibida por ninguna ley física actualmente vigente. Probablemente sea Earman (1986) quien más ha promocionado el Enfoque Estándar.

Para empezar, resulta necesario caracterizar a qué se hace referencia con la noción de “mundo”. De acuerdo con el vocabulario introducido por Earman (1986), “mundo” es “un mundo espacio-temporal de cuatro dimensiones” (p.13), y un “mundo *posible*” es una colección de eventos posibles que representan historias alternativas a las del mundo *actual*. Por “mundo *actual*”, Earman entiende el conjunto de eventos que han sucedido, suceden y

sucedarán. Sin embargo, ¿cuál es el criterio por el cual seleccionamos cierto subconjunto de mundos posibles como mundos *físicamente* posibles específicamente? Earman (1986) lo determina en términos de aquellos mundos “que satisfacen las leyes naturales que se obtienen en el mundo actual” (p. 13). Por lo tanto, de acuerdo con el Enfoque Estándar presentado por Earman, un mundo *físicamente* posible es un mundo espacio-temporal de cuatro dimensiones que representa una historia alternativa a la efectivamente ocurrida en el mundo actual, y en el cual se satisfacen las mismas leyes físicas que se satisfacen en el mundo actual.

Siendo más sistemáticos, podemos definir el Enfoque Estándar sobre la base de dos compromisos explícitos y uno implícito. Los compromisos explícitos son:

- (a) Compromiso con un enfoque modal expresado en términos de “mundos posibles”.
- (b) Compromiso con el hecho de que son las leyes físicas las que determinan la posibilidad, la necesidad y la imposibilidad *física*.

El primer compromiso solo expresa la preferencia por adoptar un enfoque que entiende las condiciones de verdad de afirmaciones modales invocando una estructura de mundos posibles¹³⁰. Como ya fue señalado, la noción de mundo posible fue introducida por Leibniz para hacer referencia a las distintas maneras en que Dios podría haber creado el mundo. Siglos después, la noción fue retomada por la filosofía analítica –sobre todo por los avances en lógica y en matemática debidos a Gottlob Frege a fines del siglo XIX y por los trabajos de Ludwig Wittgenstein a inicios del siglo XX. Sin embargo, es a mediados del siglo XX que la noción de mundo posible se tornó particularmente central para la filosofía analítica: filósofos como David Lewis, Robert Adams, Alvin Plantinga y Saul Kripke no solo discutieron qué debe entenderse por mundo posible, sino que también buscaron elucidar la naturaleza metafísica de estos mundos, es decir, si se trataba de mundos posibles reales (tal como nuestro mundo actual) o, en cambio, eran meras formas del lenguaje o construcciones mentales.

¹³⁰ Hacia fines del siglo XIX y principios del siglo XX, la noción de mundo posible no era, en absoluto, popular en el ámbito filosófico y científico de la física. En su lugar, la noción de posibilidad física era analizada en términos puramente actualistas, y se encontraba mucho más ligada a lo que *en realidad* podría ser construido dado el mobiliario y las relaciones presentes en el mundo actual (para detalles históricos, ver Uffink, 2001). Para enfoques sobre posibilidad física que prescinden de la noción de mundo posible, ver Fine (1994), Jubien (1996, 2009).

Se dejarán de lado estas discusiones aquí para dar preferencia a una breve discusión acerca de la jerarquización de los mundos posibles, pues esta discusión jugará un papel central en la siguiente sección. Como se señaló en la Sección 6.2., es usual encontrar en la bibliografía filosófica un amplio abanico de criterios para delimitar la posibilidad física o, en este caso, los mundos posibles; sin embargo, en este trabajo se discutirá acerca de la relación jerárquica entre tres clases de mundos posibles: mundos físicamente posibles, mundos metafísicamente posibles y mundos lógicamente posibles. A su vez, se asumirá en esta investigación el trato usual que se le da en la bibliografía filosófica a la relación entre estas categorías de mundos posibles, la cual sostiene que (ver, por ejemplo, Vaidya, 2007) los mundos físicamente posibles (MFP_1) son un subconjunto de los mundos metafísicamente posibles (MMP) y éste, a su vez, es un subconjunto de los mundos lógicamente posibles (MLP). Por lo tanto, si un mundo es físicamente posible, es tanto un mundo metafísicamente posible como también un mundo lógicamente posible, mientras que las relaciones inversas no se cumplen (ver Figura 6.1).

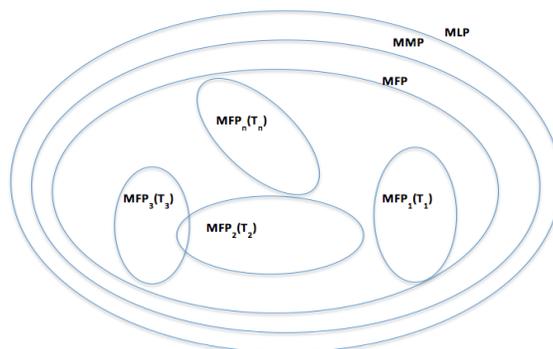


Figura 6.1. Jerarquización de los diferentes niveles de mundos posibles

También es importante resaltar que, en el nivel de la posibilidad física, se pueden definir múltiples conjuntos de mundos físicamente posibles dependiendo de qué teoría física sea considerada. Así, de acuerdo con la Figura 6.1, $MFP(T_n)$ indica el conjunto de mundos físicamente posibles en el contexto de la teoría física T_n . Los conjuntos así definidos pueden ser disjuntos o pueden intersectarse, es decir, pueden existir historias del mundo que son posibles de acuerdo con más de una teoría física. De igual manera, estos conjuntos de mundos físicamente posibles pueden ser mutuamente excluyentes, es decir,

pueden existir historias del mundo que son posibles en relación con una teoría física, pero imposibles en relación con otra.

Por otro lado, de acuerdo con el compromiso (b), resulta natural entender la posibilidad física en términos del siguiente vínculo. En general, se considera que la estructura matemática de las teorías físicas que describen el mundo actual, y que representa su estructura nomológica, se basa en ecuaciones diferenciales (que representan las leyes físicas) cada una de cuyas soluciones representa una posibilidad física diferente (o un diferente mundo posible). En general, cuando se pide un repertorio de *las* leyes de la física (independientemente de *qué* sean esas leyes), se provee una lista de ecuaciones diferenciales como la segunda ley de Newton, la ecuación de Schrödinger, las ecuaciones de campo de Einstein, etc. De acuerdo con esta perspectiva, las ecuaciones diferenciales de la física proveen el abanico de la posibilidad física expresado en sus soluciones. Naturalmente, entre esas soluciones posibles, una es la que representa la evolución en el mundo actual.

Claramente, negar el compromiso (b), ya sea para restringir o relajar el criterio acerca de que son las leyes físicas las que determinan el ámbito de la posibilidad física, no solo equivale a apartarse del Enfoque Estándar, sino que también introduce una potencial disputa acerca del alcance efectivo del dominio de la posibilidad física: al cambiar el criterio que recorta el espacio maximal de los mundos posibles a un sub-conjunto menor, ese recorte podría claramente ser diferente y comprender conjuntos no co-extensivos de mundos posibles, conduciendo por consiguiente a una comprensión distinta de la posibilidad física.

Pero antes de pasar a este punto, será útil enunciar un compromiso implícito en el Enfoque Estándar:

(c) Compromiso con una visión “robusta” de ley física

Este compromiso no ha sido suficientemente desarrollado en la bibliografía relevante para abordar la noción de posibilidad física. No solo existe una amplia variedad de visiones respecto de qué son las leyes físicas, sino diferentes maneras de agrupar tales visiones (para algunas sistematizaciones posibles, ver Carroll, 2020; Belot, 2011; Cohen y Callender, 2009). Algunos consideran que las leyes físicas son meras regularidades que supervienen al

mosaico Humeano de particulares y relaciones (ver Lewis, 1973a; Loewer, 1996); otros postulan que son relaciones entre universales primitivos (Armstrong, 1978, 1983) o relaciones entre particulares que conllevan una necesidad intrínseca o primitiva (Maudlin, 2007); mientras que algunos afirman incluso que no existen tales leyes físicas en absoluto, sino meros principios generales que describen comportamientos en modelos altamente idealizados (Mumford, 2004; Cartwright, 1983). Queda abierto a discusión si una visión regularista-Humeana de las leyes físicas es compatible con el Enfoque Estándar: si las leyes físicas son meras regularidades que supervienen a la configuración *actual* del mosaico Humeano, cualquier historia alternativa de tal configuración podría conducir a distintas leyes. Pero, sin lugar a dudas, quien sostenga que *no existen* leyes físicas como tales, no puede sostener al mismo tiempo el Enfoque Estándar, ya que éste recae, precisamente, en la existencia de tales leyes para limitar la posibilidad física. Por lo tanto, el Enfoque Estándar se compromete, al menos, con la existencia de genuinas leyes físicas, aunque, discutiblemente, podría no implicar un compromiso con algún tipo de *necesitarismo fuerte* respecto de las leyes.

Valga la siguiente aclaración para señalar cuál es exactamente el foco del argumento de esta sección contra el Enfoque Estándar de la modalidad física y el papel que la distinción nómico-extranómico juega en él. Si bien el Enfoque Estándar, como se ha afirmado anteriormente, se compromete con una visión al menos robusta de ley, esto no es central para la argumentación que se ofrece aquí: un regularista o un esencialista tendría los mismos problemas a la hora de delimitar el rango de lo físicamente posible, problemas anclados en qué tipo de elementos teóricos se consideran como nómicos. Qué es *metafísicamente* una ley no juega un papel central a la hora de delimitar lo nómico de lo extra-nómico. Tampoco es relevante qué se incluye dentro de la noción de ley: si relajamos el concepto de ley lo suficiente para permitir que cierto tipo de enunciados o elementos teóricos se consideren leyes, claramente se estaría, con el mismo movimiento, ampliando el rango de lo que es físicamente posible; pero no se habrá roto el lazo que une lo físicamente posible con el concepto de ley. Y esta ruptura es, precisamente, el objetivo que se persigue en esta sección. Por lo tanto, qué sea o no una ley, qué compromisos metafísicos respectos de qué es una ley subyace al Enfoque Estándar o qué cosas consideramos como leyes no son cuestiones centrales en la discusión que se pretende dar aquí. Lo central es, para ser

enfáticos, que cualquiera sea la noción de ley que se sostenga (siempre y cuando no lleve a una trivialización), cabe interrogarse si otros elementos no juegan acaso también un papel activo en delimitar la posibilidad física. La intención es restringir el campo de lo físicamente posible, pero no relajando la concepción acerca de qué es una ley científica, sino rompiendo el vínculo ley física-físicamente posible.

Antes de finalizar esta sección cabe abordar, brevemente, la distinción entre leyes naturales y leyes científicas (físicas en este caso). Por una parte, las *leyes físicas* pueden ser consideradas como aquellas ecuaciones que aparecen en el marco de una teoría científica y que describirían, con información adicional y de manera aproximada, el comportamiento de un cierto rango de fenómenos. Si la teoría está axiomatizada (lo cual no siempre sucede), entonces tales axiomas también contarían como leyes físicas. Como se ve, “ser una ley física” es una propiedad que elementos lingüísticos satisfacen o no. Por otra parte, las *leyes de la naturaleza* pueden ser entendidas como el sustento modal de ciertas regularidades o patrones presentes en la naturaleza; en este sentido, son de una naturaleza distinta a la de las leyes científicas. Si bien es cierto que en general los científicos y algunos filósofos de las ciencias pretenden que las leyes científicas representen también leyes de la naturaleza (en el sentido que sus “*truthmakers*” sean algún elemento de la naturaleza misma; por ejemplo, porque las leyes físicas en realidad expresan en lenguaje matemático necesidades reales o poderes causales), esto no tiene que ser necesariamente asumido. En sentido estricto, las leyes científicas no tienen *truthmakers* (por eso el problema de la inducción, por ejemplo), sino que existe cierta evidencia experimental que les brinda cierto apoyo inductivo. Es decir, entre las leyes naturales y las leyes científicas no necesariamente existe un mapeo. En el sentido de la distinción establecida aquí, el Enfoque Estándar implica un compromiso con leyes científicas y no con leyes de la naturaleza (si bien Earman expresa el espíritu del enfoque en términos de leyes de la naturaleza). Es decir, la posibilidad física, en el sentido expuesto más arriba, está dada por las leyes físicas de una teoría particular y esta concepción de posibilidad no implica necesariamente un compromiso con una visión particular acerca de las leyes de la naturaleza.

6.4. Factores extra-nómicos en la posibilidad física

Si bien el Enfoque Estándar acerca de la modalidad física ofrece una manera sistemática y efectiva de entender la noción de posibilidad física, presenta algunas limitaciones o ambigüedades. Una de sus ambigüedades radica en cómo entender el criterio mismo de lo *físicamente* posible en términos de leyes. Entre sus limitaciones, por ejemplo, se cuenta el hecho de que el enfoque no permite hablar de lo *físicamente* posible en general, sino que requiere una relativización a una teoría física en particular. Finalmente, una cuestión central reside en si el Enfoque Estándar actúa como un buen criterio a la hora de aceptar o desechar instancias de *posibilidad* física que están en consonancia con la práctica científica y con el poder explicativo esperado de una teoría en particular.

Earman, por ejemplo, define la posibilidad física estrictamente en términos de las leyes que se satisfacen en el *mundo actual*. Tim Maudlin entiende la posibilidad física en los mismos términos cuando afirma que una situación es físicamente posible cuando su ocurrencia es consistente con las leyes que “nos rodean” (Maudlin, 2007, p. 18). Sin embargo, conviven otras lecturas de este criterio nomológico para la posibilidad física. Richard Bradley y Norman Swartz, por ejemplo, afirma que “un mundo físicamente posible es cualquier mundo posible que tiene *las mismas* leyes naturales que tiene el mundo actual” (1979, p. 6). A primera vista, las dos definiciones parecen equivalentes, pero al analizarlas con detenimiento, no lo son: mientras que la primera definición solo pone como criterio para la posibilidad física la satisfacción de, *al menos*, las leyes del mundo actual, la segunda exige que *solo* sean las leyes del mundo actual las que determinen lo físicamente posible. Por ejemplo, imaginemos un mundo posible exactamente igual al nuestro, con exactamente las mismas leyes físicas, condiciones iniciales y de contorno que nuestro universo. La única excepción es que ese mundo tiene una simetría adicional respecto del mundo actual. Si consideramos a las simetrías como formando parte de la estructura nomológica de una teoría física, o como meta-leyes con un estatus superior al de las leyes dinámicas o los principios de conservación (ver, por ejemplo, Lange, 2007), tal mundo posible sería *físicamente* posible según la definición de Earman y Maudlin, pero no bajo la definición de Bradley y Swartz. Claramente, la definición de Bradley y Swartz es más estricta que la de Earman y Maudlin. Ambas definiciones, por lo tanto, caracterizan la posibilidad *física* de manera extensional e intensionalmente distinta.

La ambigüedad podría ser subsanada simplemente con acordar qué definición de posibilidad física adoptar. Sin embargo, no existe una manera clara y unívoca de llegar a un acuerdo tal. Gran parte del problema depende de poder establecer cuál de estas definiciones recoge el uso de nociones modales respecto de la posibilidad física de una mejor manera, las cuales podrían variar de contexto en contexto. Probablemente, a nivel de la ciencia física en general, una definición como la de Earman o Maudlin sea la más adecuada ya que resulta más laxa y abierta a variaciones o inclusiones; pero a nivel intra-teórico, quizás sea preferible la definición de Bradley y Swartz. De cualquier manera, es relevante tener en mente que existe una doble lectura de la posibilidad física en la bibliografía; doble lectura que podría no conducir a conjuntos de mundos físicamente posibles co-extensivos.

Una segunda limitación del Enfoque Estándar de la modalidad física es que resulta incapaz de definir la posibilidad física sin la relativización a una teoría en particular. Hasta el momento se ha procedido como si existiera una noción de posibilidad física aplicable en general, pero estrictamente sería un error suponerlo. En primer lugar, muchas de las teorías físicas son incompatibles entre sí; incluso muchas de las *mejores* teorías físicas actuales son incompatibles entre sí¹³¹. Esto conduce a que no sea posible responder de manera unívoca a la pregunta de cuáles son las leyes físicas del mundo actual, sino que siempre tengamos que relativizar la respuesta a una teoría específica. Naturalmente, esto conduce a que, bajo el Enfoque Estándar, la noción misma de posibilidad física esté necesariamente relativizada y no tenga una aplicación general a toda la física. Dicho en otras palabras, la noción de lo físicamente posible *en general* está tan vacía de contenido filosófico y contenido científico relevante como las afirmaciones de que “existe un mundo físicamente posible donde las partículas pueden tener velocidad infinita” y “no existe un mundo físicamente posible donde las partículas puedan tener velocidad infinita”. Mientras que la primera afirmación es verdadera para la mecánica clásica, no lo es para la relatividad especial.

A la luz de lo discutido en el párrafo anterior, una definición apropiada de posibilidad física en términos del Enfoque Estándar exige una relativización a una teoría física en particular. Por lo tanto: un mundo posible es *físicamente* posible de acuerdo a una teoría *T* si y solo si tal mundo satisface las leyes físicas de *T*. Bajo esta relativización,

¹³¹ Respecto al problema de la incompatibilidad entre teorías, ver la edición especial Bueno, O. y Vickers, P. (2014). “Is science inconsistent?”. *Synthese*, 191(13), 2887–2889.

probablemente la definición de Bradley y Swartz sea la más adecuada, ya que es sumamente razonable demandar que la posibilidad física bajo una teoría en particular involucre solo las leyes de esa teoría en particular, ya que, de otra manera, se podría estar lidiando con una teoría diferente y, por lo tanto, con un subconjunto diferente de mundos posibles.

Finalmente, cabe hacer una pregunta más general respecto del Enfoque Estándar: ¿es un buen enfoque para capturar la noción de posibilidad física? Las leyes físicas, ¿son el único criterio adecuado para circunscribir el mundo de lo físicamente posible, en el sentido de ser lo suficientemente permisivo para permitir dentro del ámbito de la posibilidad física ciertos mundos posibles y lo suficientemente exclusivo para desechar a otros no deseados? La tesis que se defiende en esta sección afirma que la apelación pura y exclusiva a las leyes físicas para caracterizar la posibilidad física podría ser un criterio demasiado amplio y poco específico en varias situaciones científicamente relevantes. En otras palabras, es un criterio demasiado permisivo que permite aceptar como *físicamente* posibles mundos posibles que podrían ser dejados afuera según los criterios de la práctica científica. Si bien las leyes físicas de una teoría son fundamentales en determinar cuál es la estructura de lo físicamente posible, no es el único elemento que juega un papel central. En general, el punto que se defiende es que el poder expresivo y explicativo de una teoría física, incluso en la construcción de qué modelos son válidos (posibles) para representar procesos físicos, no se agota en cuáles son sus ecuaciones fundamentales y el conjunto de sus soluciones. Por lo tanto, la noción de posibilidad física tampoco debería agotarse en ella. Antes de analizar algunos casos concretos, caben algunos comentarios generales al respecto.

En primer lugar, es claro que no existe una manera más natural que otra de circunscribir los mundos físicamente posibles. Desde un punto de vista puramente lógico, el calificativo de “físicamente” añadido a la noción de mundo posible solamente indica que el cuantificador que cuantifica sobre todos los mundos posibles es un cuantificador acotado a un subconjunto delimitado de mundos posibles; delimitación establecida sobre la base de un criterio que es introducido de manera externa respondiendo al tipo particular de expresiones que queremos admitir. En principio, el cuantificador podría acotarse de cualquier manera para expresar la noción de posibilidad física, ya que no hay elemento que fije de manera unívoca, ni de manera intrínseca, el criterio. El punto filosófico y

científicamente relevante es si tal criterio resulta efectivo y exitoso a la hora de recoger el uso de nociones modales en física, y potenciar el poder explicativo de nuestras teorías científicas. Un criterio demasiado amplio de la noción de posibilidad física podría debilitar el poder explicativo de una teoría científica al incluir más casos de los que realmente se pretende incluir, perdiendo especificidad en su aplicación y explicación. Si bien las teorías científicas nos dicen *qué* es lo posible de acuerdo con esa teoría y cuáles son las situaciones alternativas en las cuales la teoría sigue teniendo validez, también la teoría necesariamente prohíbe ciertas situaciones o deja deliberadamente fuera de su alcance otras. Una noción refinada y sofisticada de posibilidad física debiera ser capaz de lograr este doble objetivo.

La pregunta que motiva esta propuesta es: ¿se necesita *algo más* además de las leyes físicas para caracterizar la posibilidad física? Laura Ruetsche (2013) ha argumentado que, a la hora de expresar más adecuadamente el poder explicativo de una teoría mediante una interpretación, es necesario incluir diversos *factores extra-nómicos*. Su enfoque, que ella califica de *pragmático y particularista*, entiende que diferentes factores extra-nómicos (como condiciones iniciales o de contorno, e incluso arreglos experimentales particulares) imponen ciertos detalles en la explicación que “importan en la misma medida en la que se supone que las leyes importan en general: importan para establecer generalizaciones (y tal vez, también, para establecer qué propiedades son adecuadas en tales generalizaciones) y, por lo tanto, para establecer qué explicaciones son adecuadas para tales generalizaciones” (Ruetsche, 2013, p. 4). En esta investigación, se asume un punto de vista similar en relación con la noción de posibilidad física. En lo crucial, se quiere hacer hincapié en que el criterio más adecuado para calificar lo *físicamente* posible es aquél que resulta más relevante y significativo para la física en la medida en que recoge y sistematiza el uso de nociones modales en la práctica científica, y no aquél que resulta más sistemático o más simple desde un punto de vista puramente formal o conceptual.

6.5. Dos casos de estudio

Como se ha advertido en las secciones anteriores, la noción de posibilidad física podría ser más compleja de lo que el Enfoque Estándar de la modalidad física generalmente asume. En esta sección se busca motivar este diagnóstico recurriendo al análisis de algunos

ejemplos concretos de la física donde esta complejidad se torna evidente. En particular, se pretende mostrar que el Enfoque Estándar resulta demasiado permisivo y que, usualmente, los físicos recurren a criterios extra-nómicos al emplear la noción de posibilidad física. En otros términos, se afirma que la comunidad física no utiliza *solamente* las leyes de la física como criterio para decidir si ciertas situaciones son físicamente posibles o no –situaciones físicamente posibles que podrían, eventualmente, actualizarse para convertirse en lo que comúnmente se conoce como “mundo actual”.

Con este análisis no se pretende afirmar que la elucidación filosófica se deba circunscribir a meramente recoger y analizar la práctica científica. En cambio, se sostiene el trabajo debe ser “de ida y vuelta”. En este caso en particular se mostrará que algunos elementos presentes en la práctica científica –desatendidos por el Enfoque Estándar– serían de suma utilidad para alcanzar una noción filosóficamente más refinada de posibilidad física. Entonces, ¿qué elementos y/o criterios presentes en la práctica científica ayudarían a alcanzar una noción más adecuada de posibilidad física? Mediante los siguientes casos de estudio se intenta dilucidar estos elementos y/o criterios.

6.5.1. El problema de la acción a distancia

La mecánica newtoniana, desde su formulación, ha sido una fuente inagotable de investigaciones tanto en física como en filosofía. Uno de los problemas abordados por estas investigaciones, y quizás el más emblemático, es el llamado “problema de la acción a distancia”. De acuerdo con la mecánica de Newton, dos cuerpos masivos ejercen entre sí una fuerza de atracción mutua inversamente proporcional al cuadrado de la distancia entre ellos y directamente proporcional al producto de la masa de los cuerpos. Esta fuerza de atracción se produciría, de acuerdo con la teoría, de manera instantánea y sin ningún tipo de mediación física entre ambos cuerpos. Una de sus implicaciones más misteriosas y problemáticas es que esta “acción a distancia” supondría algún tipo de propagación instantánea de señales. El problema central es explicar físicamente esta fuerza y sus implicaciones.

Si bien es posible abordar el problema de la acción a distancia desde diversas perspectivas, en esta subsección se lo abordará en el marco de nuestra elucidación del concepto de posibilidad física. De acuerdo con el Enfoque Estándar, la acción a distancia (o

la propagación de señales a velocidad infinita) sería una genuina posibilidad física, ya que es una consecuencia directa del formalismo de la mecánica newtoniana. Una manera de evitar el problema sería evaluar este resultado desde una teoría más general: de acuerdo con la relatividad especial no es posible propagar señales a una velocidad mayor que la velocidad de la luz. Por lo tanto, en el marco de la relatividad especial, la acción a distancia no es una posibilidad física en absoluto. El Enfoque Estándar contempla la relativización de la posibilidad física al marco de una teoría determinada: lo que es físicamente posible en una teoría, podría no serlo en otra. Y la acción a distancia es un caso ejemplar de ello. Esto podría suponer que el problema automáticamente desaparece porque la mecánica newtoniana pasa a ser una teoría secundaria que aproxima ciertos resultados de la relatividad especial en contextos determinados (bajas velocidades, por ejemplo). La manera correcta de evaluar la posibilidad física sería tomar la teoría más general posible que describa (o incluya) el mismo rango de fenómenos y, a partir de ella, evaluar qué es lo físicamente posible.

Sin embargo, el punto que se quiere hacer notar es sutilmente diferente: independientemente de que la acción a distancia sea o no una posibilidad física en general dado el contexto de las mejores teorías aceptadas en la actualidad, persiste el problema de cómo se llega a identificar y evaluar un caso de posibilidad física. Y este problema aún persiste en cada teoría en la cual se quiera llevar a cabo la definición. El problema no es si x es una posibilidad física en general o conforme a nuestras mejores teorías físicas actualmente vigentes, sino cómo se logra tal definición *en cada teoría*. Y es en este punto donde se advierten las divergencias entre el Enfoque Estándar y la práctica científica. Por lo tanto, no es necesario evaluar casos extraños o anti-intuitivos de posibilidad física desde un marco teórico más general, sino que éstos deben ser evaluados y contrastados con la práctica científica en el marco de la teoría en discusión.

Volvamos ahora con el problema de la acción a distancia para aclarar el punto que se pretende destacar. Es posible rastrear, en el propio Newton, algunos indicios de la tesis que se propone aquí. Mediante un análisis de los escritos de Newton, por ejemplo, en el propio Escolio de los *Principia*, es posible inferir que la razón de Newton para no ofrecer una explicación científica de la acción a distancia como una posibilidad física corresponde a criterios extra-nómicos que podrían ser considerados como de tipo metafísico y teológico.

En primer lugar, Newton apela a la noción metafísica de *causalidad* para delimitar la posibilidad física. Newton sostiene que debe existir una causa para todo efecto y que la aceptación de la acción a distancia permitiría la existencia de ciertos efectos sin una causa deducible de los propios fenómenos. Evidencia de esta postura se puede rastrear en el Escolio general al final de los *Principia*:

“Hasta ahora he explicado los fenómenos de los cielos y de nuestro mar por la fuerza de la gravedad, pero no he asignado todavía una causa a la gravedad. En efecto, esta fuerza surge de alguna causa que penetra hasta los centros del sol y los planetas sin ninguna disminución de su poder para actuar, y no lo hace en proporción a la cantidad de superficie de las partículas sobre las que actúa (como lo hacen las causas mecánicas) sino en proporción a la cantidad de materia sólida, y cuya acción se extiende por todas partes a distancias inmensas, siempre decreciendo como el cuadrado de la distancia [...] No he podido todavía deducir de los fenómenos la razón de estas propiedades de la gravedad, y no finjo hipótesis. Porque lo que no se deduce de los fenómenos debe llamarse hipótesis; y las hipótesis, sean metafísicas o físicas, o basadas en cualidades ocultas, o mecánicas, no tienen lugar en la filosofía experimental [...]. Unas pocas podrían ahora ser añadidas con respecto a un cierto espíritu muy sutil que impregna grandes cuerpos y que se encuentra oculto ellos; por su fuerza y acciones, las partículas de los cuerpos se atraen entre sí a distancias muy pequeñas y se cohesionan cuando se vuelven contiguas [...] no hay un número suficiente de experimentos para determinar y demostrar con precisión las leyes que gobiernan las acciones de este espíritu.” (Newton 1999 [1687], pp. 943–944).

Hacia el final de la cita anterior se puede observar que Newton desliza alguna de sus intuiciones respecto del problema de la acción a distancia al mencionar que no ha podido deducir de los propios fenómenos la causa de las propiedades de la gravedad y que estas propiedades pueden ser obra de alguna especie de espíritu universal que penetra a través de los sólidos y que permite que los cuerpos logren actuar a pequeñas y grandes distancias. La búsqueda de una causa, incluso fuera de los propios fenómenos, deja en evidencia que, para

Newton, es imprescindible la presencia de una causa que explique el fenómeno de la acción a distancia: no estaba dispuesto en modo alguno a renunciar a la causalidad como principio explicativo.

Si efectivamente se añaden criterios metafísicos para delimitar lo físicamente posible, la caracterización estándar de mundo posible como sistema jerarquizado (ver Sección 6.2) entra en tensión. Como se ha discutido más arriba, los diferentes niveles de posibilidad en la visión de mundos posibles son los siguientes: el conjunto mayor es el de los mundos lógicamente posibles, y el conjunto de los mundos metafísicamente posibles está incluido en él (o sea, todos los mundos metafísicamente posibles son mundos lógicamente posibles, pero no viceversa). A su vez, el conjunto de los mundos físicamente posibles estaría incluido en el conjunto de los mundos metafísica y lógicamente posibles. Sin embargo, al adoptar el principio de causalidad como un criterio para lo metafísicamente posible (como hace Newton), se debería seguir que la acción a distancia no podría ser físicamente posible. Caso contrario, existirían mundos físicamente posibles, que no serían metafísicamente posibles, lo cual sería un resultado bastante paradójico. En este estado de cosas, tenemos dos opciones: (a) cambiar los compromisos metafísicos, admitiendo que quizás la causalidad no es tan importante como creemos, o (b) modificar la noción de posibilidad física, reconociendo que, si bien hay situaciones que parecen seguirse del formalismo de la teoría y, por lo tanto, serían físicamente posibles según el Enfoque Estándar, sería un error considerarlas físicamente posibles porque contradicen algún principio metafísico previamente asumido.

En segundo lugar, se puede también encontrar un criterio de tipo teológico que se encuentra estrechamente ligado con el criterio metafísico expuesto en los párrafos anteriores. La interpretación más aceptada respecto de cuál sería la causa de la gravitación de la que Newton habla es que dicha causa vendría impuesta por Dios. Por ejemplo, Alfred Rupert Hall en su libro *La Revolución Científica. 1500–1750* sostiene que algunos newtonianos de la época opinaban que el significado de las afirmaciones de Newton es que la omnipresencia de Dios es la causa de que en todas partes funcione la fuerza gravitacional entre los cuerpos (Hall, 1985 [1954], p. 482). Este sería el espíritu universal del que Newton hablaba y que haría posible la acción a distancia. Por su parte, Alexandre Koyré en su libro *Newtonian Studies* sostiene una tesis similar a la de Hall respecto de cuál sería la

causa de la gravitación según Newton: “La atracción como acción a distancia a través del vacío sin mediación, le dijo a Bentley, era un completo absurdo en el que nadie podía creer, afirmando además con bastante claridad que esta mediación tenía que ser realizada por algo que no es material, es decir, por Dios” (Koyré, 1965, p. 149).

Este breve análisis muestra cómo la comunidad científico-filosófica de la época recurría a criterios extra-nómicos para establecer qué es lo físicamente posible. Y la tendencia, *mutatis mutandis*, persiste hasta nuestros días: Newton, como otros científicos de la época y actuales, lidian con el problema de la acción a distancia siguiendo esta estrategia. Los compromisos metafísicos previamente asumidos juegan un papel fundamental en esta tarea, ya que imponen principios más generales acerca de cómo circunscribir la noción de posibilidad física. En este sentido, el Enfoque Estándar resulta demasiado permisivo o incompleto, en tanto solo presta atención a las soluciones de las ecuaciones dinámicas de la teoría física en cuestión. Bajo este enfoque, Newton –y cualquier otro científico– debería *ciegamente* aceptar la acción a distancia (en el marco de la mecánica newtoniana) como una genuina posibilidad física. Más allá de que esto resulte incómodo o controversial, el punto que se pretende enfatizar es que los científicos mismos parecen ser más sutiles en esta tarea al suponer ciertos compromisos metafísicos y evaluar la posibilidad física a la luz de ellos. Por lo tanto, se afirma que una definición más refinada de la posibilidad física debería incluirlos explícitamente, cumpliendo el papel activo que parecen jugar en la práctica científica.

6.5.2. Problemas cosmológicos

El segundo ejemplo que se propone para poner de manifiesto la multiplicidad de criterios utilizados para circunscribir la noción de posibilidad física en diversas teorías físicas se refiere a los agujeros negros, agujeros blancos y agujeros de gusano en cosmología. En términos generales, un agujero negro corresponde a una región del espacio-tiempo dentro de la cual existe una singularidad y donde la gravedad es tan fuerte que nada puede escapar de esta región. En segundo lugar, un agujero blanco es un objeto teórico derivado de las ecuaciones de Einstein (al igual que los agujeros negros), correspondiente a una región del espacio-tiempo dentro de la cual existe una singularidad que deforma esta región y desde la cual emerge materia y energía. Por último, un agujero de gusano se puede definir como un

punto de vista, un puente entre dos regiones distantes del espacio-tiempo. Se denomina agujero de “gusano”, ya que la materia y energía podría viajar a través de este conducto, el cual conectaría dos regiones del universo¹³².

Como es sabido en cosmología, tanto los agujeros negros, así como también los agujeros blancos y de gusano son soluciones de las ecuaciones de campo de Einstein. Es decir, de acuerdo con el Enfoque Estándar, corresponden a genuinas posibilidades físicas en el contexto de la relatividad general. Sin embargo, la comunidad científica no parece haber aceptado esta conclusión de manera universal y sin mayores precisiones: un breve análisis de diversas fuentes bibliográficas muestra que se ha acudido a motivaciones extra-nómicas para aceptar o rechazar la posibilidad física de estos objetos cosmológicos. Esto fortalecería el argumento esgrimido aquí, al señalar una multiplicidad de criterios en lugar de recaer en el mero formalismo de las teorías científicas como criterio de posibilidad física.

Uno de los principales argumentos que se encuentran en la bibliografía, y que podría constreñir la noción de posibilidad física, está relacionado con algunos principios más generales de la física. En este sentido, un primer criterio que es posible rastrear es la imposibilidad de objetos con densidad infinita: la comunidad científica de la época consideró con reticencia la posibilidad del colapso gravitacional de las estrellas en singularidades de densidad infinita. Algunos de los físicos más renombrados de la época, como Arthur Eddington y Albert Einstein, se negaron a aceptar tal posibilidad. Stephen Hawking, en su libro *Historia del tiempo: del Big-Bang a los Agujeros Negros*, escribe:

“... Eddington se asombró tanto por esta conclusión que se rehusó a creer en el resultado de Chandrasekhar. Pensó que era simplemente imposible que una estrella pudiera colapsarse y convertirse en un punto. Éste fue el criterio de la mayoría de los científicos: el mismo Einstein escribió un artículo en el que sostenía que las estrellas no podrían encogerse hasta tener un tamaño nulo...”

(2011 [1988], p. 138).

¹³² El primer físico en proponer la posibilidad matemática y quizás física de los agujeros de gusano fue Ludwig Flamm en 1916, derivando algunos resultados de la relatividad general. Posteriormente, en 1935, Einstein y Rosen publicaron un artículo asociado a los agujeros negros en que postulaban la posibilidad matemática, derivada de la relatividad general, de los agujeros de gusano. Este resultado hasta el día de hoy se conoce como puente Einstein-Rosen. El término ‘agujero de gusano’ fue acuñado recién en 1957 por el físico John Wheeler.

La reticencia de Einstein y Eddington a aceptar la posibilidad física de los agujeros negros fue manifiesta aun cuando tal posibilidad se seguía de las ecuaciones de campo de Einstein: la relatividad general, desde la perspectiva de su formalismo puro, acepta la posibilidad de cuerpos de densidad infinita. El punto a subrayar aquí es que tanto Einstein como Eddington (y otros físicos de la época) veían la necesidad de apelar a criterios adicionales para circunscribir la posibilidad física y evitar leer su definición del puro formalismo. La plausibilidad de tales objetos, al parecer, entraría en conflicto con principios más generales de la física, y estos podrían ser utilizados como criterios adicionales para evaluar la posibilidad física de alguna solución peculiar.

Otro criterio físico que parece jugar un papel central en la conceptualización de la posibilidad física en una teoría es el principio de la conservación de la energía. Como se dijo anteriormente, un agujero blanco es un objeto teórico cuyo comportamiento es, en algún sentido, contrario al de un agujero negro. Es decir, estos objetos solo permiten que de ellos escape energía y materia, sin que nada (es decir, energía y/o materia) pueda entrar en ellos. Un argumento para negar la posibilidad física de los agujeros blancos recaería en que su existencia violaría el principio de conservación de la energía al producir materia-energía *ex nihilo*.

Principios como el de la conservación de la energía juegan un papel muy relevante en física, y parecen guiar muchas de nuestras intuiciones respecto de cómo delimitar la noción de posibilidad física. Sin embargo, no es completamente claro qué estatus tienen en el corpus de las teorías físicas. Si bien abordar cabalmente este problema excede los alcances de este trabajo, vale la pena mencionar brevemente el papel preponderante que los principios de esta índole podrían jugar en la delimitación de la posibilidad física. En general, es posible ofrecer dos enfoques: (a) los principios de conservación forman parte del marco general de una teoría física, o se deducen de este marco general y, por tanto, están incluidos en el Enfoque Estándar. En este sentido, serían principios de tipo nomológico en el seno de una teoría en particular. O, en su lugar, (b) son principios generales de validez inter-teórica independientes de las leyes de una teoría determinada. Este es un asunto de crucial importancia para este análisis, pues, si efectivamente estos principios forman parte de los elementos propios de una teoría física o se deducen de ellos, entonces cuando se hace referencia a estos principios para constreñir lo físicamente posible

no se está haciendo nada más que apelar al Enfoque Estándar. Por otra parte, si el principio de conservación de la energía tiene un carácter extra-nómico, actuando como un principio más general que constriñe la posibilidad física derivada de las leyes de una teoría en particular, entonces sí podría considerarse que este principio es un genuino criterio que permite caracterizar la modalidad física más allá del Enfoque Estándar¹³³.

Este análisis abre una serie de interrogantes que no han sido abordados por la bibliografía filosófica hasta el momento: ¿cuál es el alcance estricto del Enfoque Estándar? Al apelar puramente al aparato nomológico de una teoría para caracterizar la posibilidad física, ¿qué deberíamos incluir en tal aparato nomológico? Ya vimos que el estatus de los principios de conservación es ambiguo, pero el mismo razonamiento podría extenderse a otros principios, como las simetrías, por ejemplo. Si bien el estatus general de los principios de conservación o de las simetrías ha sido ampliamente discutido en la bibliografía pertinente (ver, por ejemplo, Lange, 2007), tal discusión no ha sido relacionada con la caracterización de la posibilidad física. Este nuevo elemento conduciría a complejizar más aún el panorama acerca de lo físicamente posible: o bien se considera que tales principios o simetrías son ajenas y constriñen el aparato nomológico, y por lo tanto, deberían considerarse como un criterio extra-nómico más para caracterizar la posibilidad física (mostrando la inadecuación del Enfoque Estándar), o bien se los considera como parte del aparato nomológico, no calificando como un criterio extra-nómico, pero sugiriendo que el Enfoque Estándar debería modificarse, o al menos ser más explícito respecto de estos compromisos y sus alcances.

Dejando por un momento de lado los principios generales de conservación y las simetrías, otro posible criterio extra-nómico que puede hallarse en la bibliografía es de tipo matemático: la presencia de *inestabilidad*. La estabilidad de una solución está asociada a cómo ésta varía cuando se modifican levemente las condiciones iniciales del sistema. No es difícil encontrar en la bibliografía técnica asociada a los agujeros de gusano menciones a la inestabilidad matemática de este tipo de soluciones. Por ejemplo, Hawking, al describir qué sucedería con un astronauta que cae en el interior de un agujero de gusano, hace referencia a la inestabilidad de estos objetos cosmológicos:

¹³³ En la Sección 6.6 se discutirá con más detalle el problema del estatus de las leyes de conservación.

“...él puede evitar chocar con la singularidad y, en vez de esto, caer a través de un agujero de gusano (“*wormhole*”), para salir en otra región del universo. Esto ofrecería grandes posibilidades de viajar en el espacio y en el tiempo, aunque desafortunadamente parece ser que estas soluciones son altamente inestables; la menor perturbación, como, por ejemplo, la presencia del astronauta, puede cambiarlas, de forma que el astronauta podría no ver la singularidad hasta que chocara con ella, momento en el que encontraría su final...” (2011 [1988], p 143-144).

La inestabilidad de la solución de agujero de gusano daría a entender que su posibilidad física resulta cuestionable, ya que pequeñas modificaciones en las condiciones iniciales modificarían sustancialmente el sistema físico, eventualmente haciéndolo evolucionar hacia otro objeto físico con características totalmente distintas a las originales.

Un trabajo que analiza la inestabilidad de la solución de agujero de gusano, en particular del tipo Ellis o Morris-Thorne (un tipo de agujero de gusano “atravesable”) es el de los físicos Takashi Torii y Hisa-aki Shinkai (2013). Los autores sostienen que “[l]a inestabilidad de los agujeros de gusano requiere técnicas adicionales de ciencia ficción para su conservación. No solo eso, sino que esto indica que una construcción de agujero de gusano tan simple no puede estar disponible como un objeto astrofísico con la configuración actual.” (p. 6). Respecto de este argumento, la inestabilidad de estos objetos cosmológicos no solo imposibilitaría utilizarlos como medios de viaje, sino también haría imposible conocerlos efectivamente. Si bien su posibilidad física, de acuerdo con el Enfoque Estándar, estaría dada por su aparición en el formalismo de la relatividad general, hay fuertes motivos para sospechar que serían objetos *imposibles* de conocer dada su inestabilidad. Si bien esto no es un argumento contundente en torno a su posibilidad física *per se*, sí es un argumento en contra de cualquier aspiración de utilizar o conocer sus propiedades. ¿Estarían los físicos dispuestos a aceptar la posibilidad física de objetos cosmológicos imposibles de conocer por su propia naturaleza? Una respuesta afirmativa a tal pregunta corre el riesgo de extraer conclusiones ontológicas únicamente sobre la base de los aspectos matemáticos de una teoría.

Un tercer tipo de argumento esgrimido recurrentemente por los científicos para descartar ciertas posibilidades físicas está relacionado con la ausencia, en muchos casos, de corroboración empírica. Este criterio rechaza la posibilidad física de un objeto o fenómeno si no existe evidencia experimental que lo respalde, aunque su posibilidad se derive de las ecuaciones de una teoría física en particular. Este argumento se ha esgrimido para rechazar la posibilidad física de los agujeros negros en sus primeros estadios de desarrollo teórico, así como también para rechazar la existencia de los agujeros blancos y agujeros de gusano. Se sostiene que la omnipresencia de este argumento en contra de todos estos objetos matemáticos se debe a que estos han sido, en sus inicios, desarrollos puramente teórico-matemáticos alejados de cualquier corroboración empírica. Todos estos casos surgen como soluciones matemáticas de las ecuaciones de Einstein que podrían o no tener un significado físico.

Hawking, manifiesta que, efectivamente, muchos hombres de ciencia apelan a este criterio empírico para desestimar soluciones “extrañas” de las ecuaciones de Einstein, en particular, los agujeros negros (en el momento en que no se contaba con la evidencia que actualmente se tiene de estos objetos):

“Los agujeros negros son un caso, entre unos pocos en la historia de la ciencia, en el que la teoría se desarrolla en gran detalle como un modelo matemático, antes de que haya ninguna evidencia a través de las observaciones de que aquella es correcta. En realidad, esto constituía el principal argumento de los oponentes de los agujeros negros: ¿cómo podría uno creer en objetos cuya única evidencia eran cálculos basados en la dudosa teoría de la relatividad general?” (2011 [1988], p. 149)

Apelar a este criterio empírico ha permitido a los físicos descartar soluciones exóticas de las ecuaciones de Einstein, como en su momento fueron los agujeros negros o como hoy en día sucede con la solución de agujero blanco o agujero de gusano.

Este tipo de argumentación, sin embargo, va en contra de una visión positiva de la posibilidad en sí misma. Es decir, es un argumento puramente actualista, cuasi-eliminativista de la posibilidad: todo lo que puede existir, existe; y lo que no existe, no puede existir. Esto es algo contrario al Enfoque Estándar, pero que también se opone a

cualquier intención de brindar una teoría más positiva de la posibilidad. La posibilidad física no es acerca del mundo actual, sino acerca de los mundos posibles entre los cuales se encuentra el mundo actual. Afirmar que algo no es posible porque no es actual es adoptar una postura eliminativista respecto de la posibilidad. Por lo tanto, este tipo de argumento debería ser descartado como candidato a especificar la posibilidad física.

Por último, es posible encontrar un criterio de tipo metafísico que apela nuevamente a la noción de causalidad para argumentar en contra de la posibilidad física de los agujeros de gusano. Este criterio sostiene que los agujeros de gusano harían posibles los viajes al pasado y, por lo tanto, podrían darse casos en los que el efecto precede a la causa. Respecto de este problema, en un artículo de 1988 Michael Morris, Kip Thorne y Ulvi Yurtsever sostienen lo siguiente: "...si las leyes de la física permiten agujeros de gusano atravesables, entonces probablemente también permiten que un agujero de gusano se transforme en una "máquina del tiempo" con la cual podría violarse la causalidad." (p.1446). En esta cita es evidente que los autores se preocupan por la posibilidad de la existencia de los agujeros de gusano y la consecuente ruptura de la causalidad.

Ya dejando de lado los objetos cosmológicos, otro interesante caso que se puede mencionar brevemente para poner en evidencia cómo operan los criterios metafísicos en la delimitación de lo físicamente posible es el llamado "Universo de Gödel": una solución exacta de las ecuaciones de Einstein propuesta por Kurt Gödel en 1949. Una de las características más importantes de la solución de Gödel es la posibilidad matemática de que en todo punto del espacio-tiempo existan curvas temporales cerradas ("*closed timelikes curves*", CTCs). Una CTC es básicamente una cadena causal cerrada. Es decir, sobre la base de estas CTCs es posible obtener una situación en la que el efecto ocurre antes que la causa sin necesidad de agujeros de gusano. Verbigracia, moviéndose a través de una curva de este tipo sería posible viajar al pasado y visitar, por ejemplo, a nuestro abuelo. Para muchos físicos parece ser claro que este tipo de soluciones no son posibles por las extrañas propiedades causales que presenta. Una vez más, al igual que con el problema de la acción a distancia, el criterio metafísico de la causalidad parece perfilarse como un serio candidato para ser incluido como factor extra-nómico que permite delinear un enfoque más adecuado de la posibilidad física.

6.6. Leyes de conservación y posibilidad física

En esta sección haremos uso del argumento defendido en las secciones anteriores acerca de la noción modal de “posibilidad física” y, en particular, de las nociones de “mundo físicamente posible” y “mundo físicamente imposible”, para postular una condición que, consideramos, sería deseable imponer a toda teoría de la causación que pretenda fundamentarse sobre cantidades físicas conservadas y sobre (o hacer uso de) nociones modales como las mencionadas anteriormente. Tomando como referencia esta condición, evaluaremos las consecuencias para la compatibilidad (o incompatibilidad) de la teoría de la causación de Dowe en la RG.

Como hemos argumentado en las secciones precedentes, parece razonable, de acuerdo con la práctica y la teorización científica, que ciertos factores, adicionales a las leyes dinámicas de una teoría, sean incluidos en la caracterización de la posibilidad física. Como mostraremos más abajo, resulta particularmente relevante para el argumento de esta sección la inclusión de las leyes de conservación como un elemento que permita circunscribir lo físicamente posible. De este modo, y considerando la fundamentación ofrecida en las secciones anteriores, parece deseable que una teoría de la causación física formulada en términos de cantidades conservadas, como la TCC de Dowe, adhiera a la siguiente condición respecto a la noción de posibilidad física:

Cond2: La noción de posibilidad física debe incluir ciertos elementos, adicionales a las ecuaciones dinámicas, que se encuentren en concordancia con la práctica y teorización física. En particular, debe incluir, al menos, a las leyes de conservación (en su forma diferencial o en su forma diferencial e integral).

De esta manera, la *Cond2* exige que la posibilidad física resulte determinada no solo por aquello que es permitido por las leyes dinámicas de una teoría (es decir, el Enfoque Estándar), sino además por el cumplimiento de las leyes de conservación. Dicho de otro modo, en todo mundo físicamente posible se cumplen las leyes de conservación. Aquellos mundos donde no se cumplen las leyes de conservación serán mundos físicamente imposibles aun cuando respondan a las leyes de la teoría. En consecuencia, bajo la *Cond2*,

el conjunto de los mundos físicamente posibles posee una cardinalidad menor que el conjunto de mundos físicamente posibles dados por aquella noción de posibilidad física que no cumpla la *Cond2* (ver Figura 6.2).

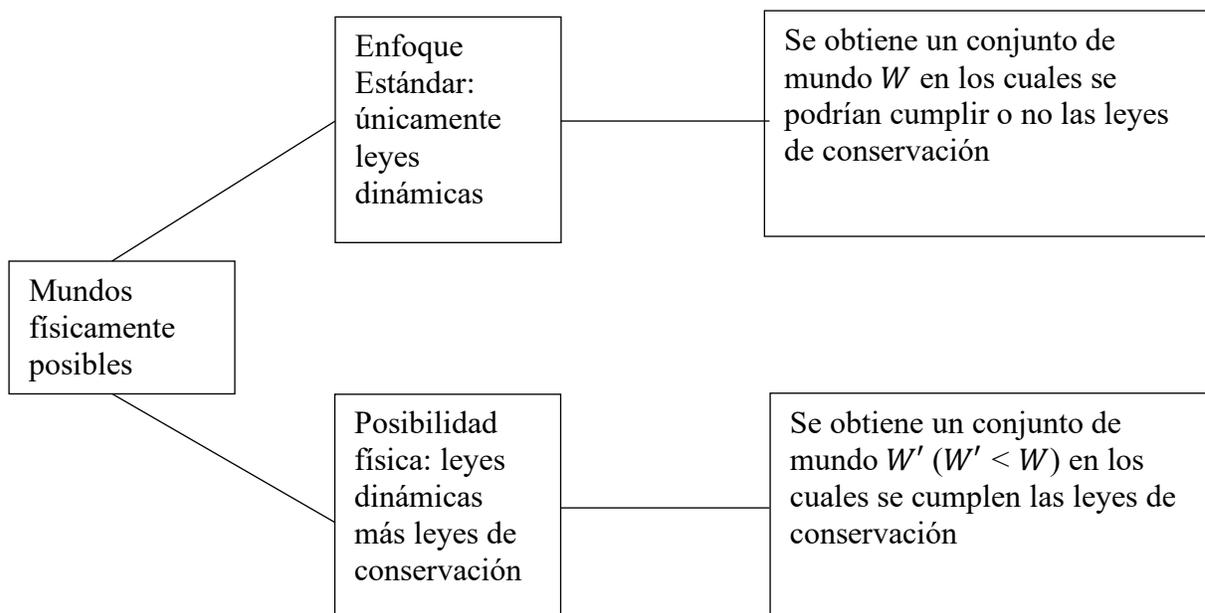


Figura 6.2: Noción de posibilidad física de acuerdo a si se agregan o no las leyes de conservación en su caracterización.

Veamos cómo funciona esta condición en el caso particular de la RG, el caso que nos interesa en esta investigación. Con este objetivo en mente, primero recordaremos algunas ideas importantes mencionadas en el capítulo anterior que nos serán de utilidad para la presente discusión. En primer lugar, en el capítulo precedente hemos señalado que la RG, a través de sus ecuaciones de campo, permite obtener diversos modelos o espacio-tiempos, los cuales dependen de la variedad M , la métrica g y el tensor energía-momentum T que se considere. Recordemos, en segundo lugar, que cada uno de los modelos (M, g, T) que cumpla con las ecuaciones de campo de Einstein es denominado un modelo dinámicamente posible (MDP). Cada MDP puede ser considerado un “mundo posible” o, más específicamente, un “mundo físicamente posible” de la RG en concordancia con el Enfoque Estándar (recuérdese Capítulo 5, Sección 5.5). Por último, hemos también argumentado que las leyes de conservación en forma integral, las relevantes para la TCC, se cumplen en una cierta clase de MDP’s de la RG. Este hecho lo hemos considerado a través de la *Cond1*.

Pues bien, dicho lo anterior, ¿cómo influye entonces la *Cond2* en nuestra consideración del concepto de posibilidad física en el marco de la RG? Para responder a esta pregunta tenemos dos escenarios posibles: aceptar o rechazar la *Cond2*. Veamos, en primer lugar, qué sucede si desechamos esta condición. En este caso, aquello que consideremos como físicamente posible estará dado por lo que dicta el Enfoque Estándar. En efecto, tendríamos un conjunto W de mundos físicamente posibles dados por todos los MDP's de la teoría. Desde esta perspectiva, el conjunto de mundos físicamente posibles W puede cumplir o no con las leyes de conservación, sean estas diferenciales o integrales. Por otro lado, si adherimos a la *Cond2*, el conjunto de mundos posibles W se verá reducido. El conjunto resultante W' estará compuesto únicamente por aquellos MDP's que presenten además leyes de conservación correctamente definidas. En este nuevo escenario, los mundos físicamente posibles de la RG serán definidos por su pertenencia a W' y no a W . Aquellos MDP's que no cumplan con las leyes de conservación (ni diferenciales ni integrales) serán considerados mundos físicamente imposibles.

Debemos, sin embargo, precisar aún más las consecuencias que impone la *Cond2* a la noción de posibilidad física en el marco de la RG. En esta condición aludimos al término “leyes de conservación” en un sentido amplio, incluyendo a las leyes de conservación tanto en su forma diferencial como en su forma integral. Hemos señalado en el Capítulo 5, Sección 5.5, que la *Cond1* trae consigo un ordenamiento de los modelos de la RG en la que tenemos un conjunto de ellos en el que únicamente son válidas las expresiones diferenciales de conservación y un subconjunto del conjunto anterior en que son válidas tanto las expresiones diferenciales como integrales de conservación. Recuérdese, de hecho, que la ecuación integral de conservación de la energía-momentum de los campos de materia más la energía-momentum del campo gravitacional, la ecuación (5.27), se deriva de la ecuación diferencial (5.24) bajo restricciones particulares. Esto trae como consecuencia que esta última ecuación sea válida en una mayor cantidad de modelos de la RG que la primera ecuación antes mencionada.

Por lo tanto, una imagen completa de la noción de posibilidad física de acuerdo a si se incluyen o no las leyes de conservación como una restricción para delimitar aquello considerar como físicamente posible es la siguiente. Por un lado, si se rechaza la *Cond2*, tenemos el caso más general, es decir, una noción de posibilidad física que incluye a *todos*

los modelos que son soluciones a las ecuaciones de campo de Einstein, el conjunto W mencionado más arriba. Por otro lado, si se acepta la *Cond2* se obtiene el conjunto W' que representa a los mundos físicamente posibles dados por los modelos de la RG que cumplen con las ecuaciones diferenciales de conservación, con independencia de si a partir de estas expresiones diferenciales es posible derivar alguna ecuación integral de conservación. Ahora bien, si en algunos modelos se puede además derivar las leyes de conservación integrales a partir de las leyes diferenciales (por ejemplo, bajo algunas restricciones particulares), se obtiene un conjunto de mundos físicamente posibles, llamémoslo W'' , que será un subconjunto de W' . Esta imagen que nos ofrece la *Cond2* acerca de la noción de posibilidad física en la RG queda ilustrada en la Figura 6.3.

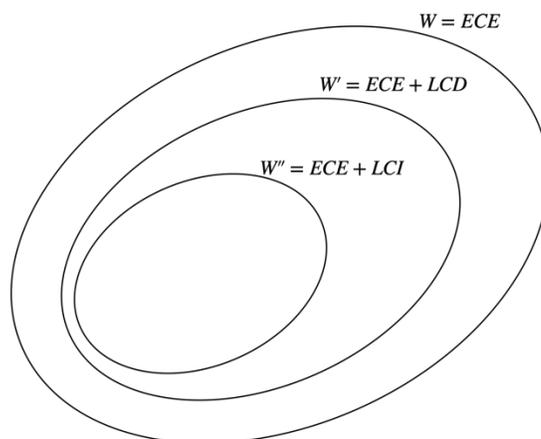


Figura 6.3: Noción de posibilidad física en el marco de la RG. En la imagen, ECE: ecuaciones de campo de Einstein; LCD: leyes de conservación diferenciales y LCI: leyes de conservación integrales.

Ahora que ya sabemos cómo opera la *Cond2*, veamos qué consecuencias trae consigo esta condición para la TCC en el marco de la teoría de la gravitación de Einstein. Una primera consecuencia es que, al aplicar las condiciones *Cond1* y *Cond2* simultáneamente, obtendremos un conjunto de mundos físicamente posibles (el conjunto W'') en los cuales existen genuinas leyes de conservación tanto diferenciales como integrales; por lo tanto, podemos afirmar que las proposiciones *CC1* y *CC2* de Dowe pueden ser aplicadas sin dificultades en esos mundos posibles. Esta sería una consecuencia deseable para aquellos y aquellas que pretendan defender ciertas afirmaciones causales en el marco de la RG. Sin

embargo, una segunda consecuencia es que la imposición de ambas condiciones a la vez permite la existencia de mundos físicamente posibles en los que solo se cumplen las leyes de conservación diferenciales, pero no así las integrales, las necesarias para la teoría de la causación de Dowe (recuérdese Capítulo 5, Sección 5.5).

¿Qué hacer, entonces, con la condición *Cond2*? ¿No es exigible para una teoría de la causación física en el marco de la RG? En particular, ¿qué sucede con la TCC en aquellos mundos físicamente posibles en los que solo se cumplen las leyes diferenciales de conservación? ¿En estos casos la TCC sería falsa o diríamos que en ellos no hay causación? Por ahora, suspenderemos el juicio respecto de estas preguntas, pero las retomaremos en el siguiente capítulo, cuando contemos con más elementos para analizar estas dificultades.

Sin embargo, antes de pasar al siguiente capítulo, no puede dejar de mencionarse un asunto importante relacionado con la condición *Cond2*, que se encuentra asociado con el estatus de las leyes de conservación respecto de los principios de simetría y las leyes dinámicas de una teoría. Un análisis detallado de este problema escapa a los alcances de esta tesis; sin embargo, es oportuno mencionar, brevemente, algunos aspectos relevantes al respecto. Existe una visión ampliamente aceptada en la física teórica que presenta a las leyes de conservación como derivadas de las leyes físicas (ecuaciones dinámicas) de la teoría considerada a través de las propiedades de simetría de estas últimas. Este punto de vista se enraizó fuertemente en el trabajo de los físicos a partir de los aportes de Noether (1918) al desarrollo de la física¹³⁴. A partir del notable trabajo de esta matemática, el análisis se concentró en la relación existente entre la invariancia de las leyes dinámicas de un sistema físico, es decir, sus propiedades de simetría, y los principios de conservación. En términos generales, el primer teorema de Noether sostiene que cada propiedad de simetría del lagrangiano de un sistema físico trae como consecuencia la aparición de una cantidad conservada. Esta forma de abordar el problema de los principios de conservación se ha difundido ampliamente y es la manera en que se los enseña a los estudiantes en los departamentos de física de las universidades.

¹³⁴ De hecho, el aporte de los teoremas de Noether fue fundamental en las discusiones tempranas de la RG (ver por ejemplo, Brading, 2005; De Haro, 2021).

No obstante, la relación entre leyes de conservación, leyes dinámicas y propiedades de simetría ha sido largamente debatido en filosofía de la ciencia, en particular en filosofía de la física. En este ámbito, se discute la presunta prioridad explicativa de las simetrías por sobre las leyes de conservación (ver, por ejemplo, Brading, 2002; Brading y Castellani, 2003, 2006; Brown y Holland, 2004; Romero-Maltrana, 2015), así como también el estatus que tendrían las leyes de conservación en comparación con las leyes dinámicas y las propiedades de simetría (ver, por ejemplo, Lange, 2007, 2009, 2011; Wolff, 2013). En general, las discusiones se centran en tratar de determinar la relación explicativa/causal entre las simetrías y las leyes de conservación, así como también entre las leyes dinámicas y las leyes de conservación. De este modo, se pone en tela de juicio la visión generalmente asumida en física teórica, que acepta la prioridad de las simetrías/leyes dinámicas por sobre las leyes de conservación.

En particular, de las discusiones en torno al estatus que tendría cada una de las clases de leyes mencionadas en el párrafo anterior se desprende que, respecto de la relación entre las leyes dinámicas de una teoría y las leyes de conservación, es posible reconocer dos posturas claramente definidas. De acuerdo con la primera, las leyes físicas o dinámicas (mecánica clásica, relatividad especial, relatividad general, etc) son principios básicos, de carácter fundamental, a partir de los cuales deben desprenderse o derivarse los principios de conservación respectivos con la ayuda de las propiedades de simetría. Es decir, en esta primera postura, se adjudica a las leyes dinámicas un estatus superior o más fundante que a los principios de conservación. De acuerdo con la segunda postura, las leyes de conservación se conciben como principios de mayor fuerza o más fundamentales que las leyes físicas consideradas. Desde esta perspectiva, las leyes de conservación poseerían validez inter-teórica, serían principios fundamentales independientes de –y lógicamente previos a– las leyes físicas consideradas.

En el primer caso mencionado en el párrafo anterior, es decir, si las leyes de conservación se consideran derivadas de las leyes dinámicas de una teoría (o en el caso que sean consideradas como partes del corpus de una teoría), entonces el Enfoque Estándar de la modalidad física simplemente ya incluiría en su caracterización de la posibilidad física las leyes de conservación. Si, por el contrario, se considera que las leyes de conservación poseen validez inter-teórica y un carácter más fundamental que las leyes dinámicas de una

teoría, entonces al apelar al Enfoque Estándar no necesariamente deben incluirse las leyes de conservación para constreñir la posibilidad física. Por lo tanto, el debate acerca del estatus de las leyes de conservación respecto de las leyes dinámicas, tal como fue expuesto más arriba, resulta jugar un papel crucial en aquello que caracteriza al Enfoque Estándar. Escapa a los objetivos de esta tesis ir más allá en este debate, pero consideramos que es un tema relevante que debe ser considerado en con mayor detenimiento en las actuales discusiones en filosofía de la ciencia y, en particular, en filosofía de la física. Desde un punto de vista práctico, podemos omitir las discusiones en torno al problema del estatus de las diferentes clases de leyes sin que esto afecte las conclusiones a las que podamos arribar.

Capítulo 7. *Metafísica de la causación II:* *el problema de la definición*

Resumen

En el capítulo anterior hemos abordado *el problema de la posibilidad física*. En particular, hemos argumentado que la inclusión de las leyes de conservación en la noción de posibilidad física no solo se encuentra en concordancia con la práctica y la teorización física, sino que además es filosóficamente relevante, entre otras razones, para evaluar los alcances y limitaciones de la TCC en contextos relativistas generales. En este capítulo abordamos el tercer y último problema que nos permitirá completar el estudio de la compatibilidad de la TCC con la RG: *el problema de la definición*.

El enfoque asumido por Dowe en su TCC puede ser resumido mediante tres tesis: (a) la TCC es producto de un análisis empírico, (b) la TCC es una teoría metafísicamente contingente y (c) la TCC es una teoría refutable. Como mostraremos, no solo cada una de estas tesis es conflictiva en sí misma, sino que también todas ellas resultan incompatibles entre sí. Para afrontar estos problemas, proponemos una solución que no solo resuelve los inconvenientes internos que presenta la TCC, provenientes de la adopción de las tesis antes mencionadas, sino que además permite que esta teoría de la causación recupere cierto grado de compatibilidad con los escenarios relativistas generales. Esta solución defiende la idea de que las proposiciones *CC1* y *CC2*, que forman el núcleo de la teoría de Dowe, sean asumidas como *definiciones reales* y no como meras caracterizaciones o criterios cuya utilidad se limita al mundo actual. Esta solución nos permite fijar la siguiente condición (*Cond3*):

Cond3: Los conceptos nucleares de una teoría de la causación física como la TCC, en particular las proposiciones *CC1* y *CC2*, deben presentarse como definiciones reales y, por tanto, metafísicamente necesarias e irrefutables.

Algunas consecuencias para la TCC en el marco de la RG, derivadas de esta condición, son las siguientes. En primer lugar, en nuestro mundo actual y en todos los mundos físicamente posibles, serán procesos causales aquellas líneas de mundo de objetos que poseen las cantidades conservadas según los principios de conservación de nuestra física. En segundo lugar, en mundos físicamente imposibles, donde se cumplen principios de conservación diferentes a los del mundo actual, serán procesos causales las líneas de mundo de objetos que poseen las cantidades conservadas según los principios de conservación de ese propio mundo. Y, por último, en los mundos físicamente imposibles donde no existen principios de conservación para ninguna magnitud física, sencillamente no habrá procesos causales¹³⁵.

7.1. El problema general de la definición

De los trabajos de Dowe acerca de su teoría de la causación física (1992a, 1992b, 1992c, 1995a, 1995b, 1999, 2000, 2000a, 2001, 2009) es posible extraer ciertas conclusiones que develan las principales características de la TCC. En particular, el autor defiende explícitamente tres tesis:

- (a) La TCC es un análisis empírico y no conceptual
- (b) La TCC es una teoría metafísicamente contingente
- (c) La TCC es una teoría refutable

Es posible mostrar que estas tesis descansan sobre el siguiente supuesto que Dowe parece defender respecto de la relación entre la metafísica y la física:

- (S): Las ciencias particulares (específicamente la física) y sus métodos representan la única forma de acceso a lo real.

En consecuencia, el proyecto de Dowe con su teoría de la causación parece descansar en un naturalismo metafísico extremo, que deja poco o nada de espacio para una metafísica que

¹³⁵ La discusión llevada a cabo en este capítulo se encuentra basada, en parte, en el siguiente trabajo: Herrera, M. (2017). “Teoría de las Cantidades Conservadas: una tensión interna”. *Revista Colombiana de Filosofía de la Ciencia*, 17(35), 91–117.

preserve cierta autonomía en sus métodos respecto de la física. En general, cuando se busca establecer una teoría acerca de la naturaleza de un cierto ítem filosófico, el punto de vista brindado por el supuesto (S) representa una perspectiva filosófica válida. Sin embargo, argumentaremos que, cuando el ítem filosófico a investigar es la causación, es decir, cuando a través de una teoría se intenta responder a la pregunta acerca de qué es la causación en el mundo, las tesis (a), (b) y (c) no solo son problemáticas en sí mismas, sino que además son incompatibles entre sí y, por lo tanto, el supuesto (S) se muestra inapropiado.

En efecto, la tesis (a) afirma que la TCC es el resultado de un análisis que pretende responder a la pregunta acerca de qué es la causación en el mundo. Al considerar que esta respuesta puede ser encontrada en la física, Dowe asume implícitamente que en dicha disciplina se cuenta con una noción técnico-científica del concepto “causa”; sin embargo, como se sabe, esta idea es fuertemente discutible. Por otro lado, la tesis (b) es doblemente problemática. En primer lugar, a partir de esta tesis se infiere que la TCC debe ser falsa en algún mundo metafísicamente posible; sin embargo, Dowe no provee ningún criterio para determinar cuándo su teoría sería falsa. En segundo lugar, si la TCC es una teoría metafísicamente contingente, entonces solo ofrece una caracterización o criterio para identificar casos de causación en el mundo actual y no una teoría acerca de *qué* es la causación en el mundo físico (que excede lo meramente actual), lo cual se contrapone con la tesis (a). Por último, respecto de la tesis (c), se argumentará que esta resulta inapropiada por dos motivos: en primer lugar, porque debido a las propias limitaciones que Dowe impone a su análisis, no existe un medio independiente para determinar cuándo la TCC sería refutada; en segundo lugar, porque la característica de refutabilidad que Dowe atribuye a su teoría descansa en una visión distorsionada de la ciencia. Cada uno de estos problemas afecta a los propios fundamentos de la TCC. Para resolver estos problemas, en este capítulo propondremos una solución, según la cual las proposiciones *CCI* y *CC2*, que forman el núcleo de la teoría de Dowe, son asumidas como *definiciones reales*. Esta solución tiene su origen en la sustitución del supuesto (S), adoptado por Dowe, por un nuevo supuesto que preserva la autonomía de la metafísica respecto de la física. Esta solución nos permitirá formular una condición (*Cond3*). Quienes pretendan defender una teoría de la causación como la de Dowe, deberían adherir a esta condición, pues su

imposición no solo permite disolver ciertas incoherencias internas de la TCC, sino que además permite que esta exhiba algún grado de compatibilidad con los escenarios relativistas generales.

Sobre esta base, el resto del presente capítulo se estructura del siguiente modo. En la Sección 7.2 presentaremos en detalle las tesis (a), (b) y (c) de Dowe. Luego, en las Secciones 7.3, 7.4 y 7.5 analizamos cada una de las tesis antes mencionadas y argumentamos que estas no solo son incompatibles entre sí, sino que además cada una de ellas representa un serio problema que afecta a los propios fundamentos de la TCC. Finalmente, en la Sección 7.6 propondremos una solución a los problemas antes señalados y postularemos la condición *Cond3*.

7.2. Las tesis de Dowe

Como ya fue señalado, la propuesta general de Dowe en su TCC descansa sobre tres tesis explícitamente formuladas por el autor. A continuación, se considerará cada una de ellas.

7.2.1. (a) TCC: un análisis empírico

La empresa general que acomete Dowe con su análisis de la causación ya ha sido brevemente discutida en el Capítulo 3, Sección 3.1. Sin embargo, es absolutamente pertinente regresar a esta discusión aquí dada su relevancia para el argumento que se pretende desarrollar en este capítulo.

Como ha sido señalado en la Introducción y en el Capítulo 3 de esta tesis, la TCC, cuya versión final se presenta en el libro *Physical Causation* (2000), es el resultado del trabajo de casi una década por parte de su autor, quien publicó una primera versión de esta teoría en el año 1992 como una reacción al trabajo de Salmon (1984) sobre explicaciones de tipo causal. Ya en ese primer artículo es posible rastrear la posición general del autor, heredada de Salmon, respecto de la forma en que debería ser abordado el problema de la causación. En ese primer trabajo, el autor sostiene: “la TCC retiene el espíritu del trabajo de Salmon en que es una teoría de ‘proceso’ (esto es, considera a los procesos e interacciones

como más fundamentales que los eventos), y en que *aspira a los mismos ideales empiristas.*” (Dowe, 1992c, p. 126; énfasis agregado con fines expositivos)^{136,137}.

Sin embargo, es en su libro del año 2000 donde el autor expresa con mayor claridad la naturaleza de su TCC y el proyecto general en el que esta teoría de la causación se enmarca. En aquella obra, Dowe dedica varias páginas a estos fines (el Capítulo 1 completo). Allí se distinguen dos tipos de análisis que pueden ser llevados a cabo para analizar la causación, cada uno de ellos, de acuerdo con el autor, legítimo por derecho propio. Por un lado, Dowe afirma que es posible llevar a cabo un *análisis conceptual*, entendido como “un análisis del significado que comienza con nuestra comprensión cotidiana y de sentido común de los conceptos relevantes. Esto es, la manera en la cual comúnmente hablamos y pensamos provee la información primaria para el análisis” (Dowe, 2000, p. 2)¹³⁸. Según el autor, este tipo de análisis es a priori y, de ser verdadero, será necesariamente verdadero (Dowe, 2000, p. 2). Por otro lado, sostiene que un análisis filosófico de un cierto ítem, en particular de la causación, también puede ser llevado a cabo mediante un *análisis empírico*, el cual “busca establecer qué *es* la causación en el mundo actual. El análisis empírico tiene como objetivo mapear el mundo objetivo, no nuestros conceptos. Tal análisis solo puede proceder a posteriori” (Dowe, 2000, p. 3)¹³⁹. Dowe enfatiza la naturaleza de este tipo de análisis afirmando lo siguiente:

“Todos estos filósofos [los que históricamente han llevado a cabo, según Dowe, análisis empíricos acerca de la causación] están de acuerdo en que para comprender la causación necesitamos ir más allá de las palabras, mirar el

¹³⁶ “*The CQ theory retains the spirit of Salmon's work in that it is a 'process' theory (that is, takes processes and interactions as more fundamental than events), and in that it aspires to the same empiricist ideals.*”

¹³⁷ Dada la relevancia de las citas presentadas en esta sección para el argumento de esta tesis, de ahora en adelante toda cita será también presentada en el idioma original como una nota a pie de página.

¹³⁸ “*conceptual analysis is a meaning analysis that begins with our everyday, common sense understanding of the relevant concepts. That is, the way in which we commonly speak and think provides the primary data for the analysis.*”

¹³⁹ “*...empirical analysis seeks to establish what causation in fact is in the actual world. Empirical analysis aims to map the objective world, not our concepts. Such an analysis can only proceed a posteriori.*” (énfasis en el original)

mundo. Esto es lo que se pretende con la etiqueta ‘análisis empírico’” (Dowe, 2000, p. 4)¹⁴⁰.

Además, el autor afirma que este tipo de análisis puede tener diferente estatus dependiendo de si se busca establecer una verdad contingente o una verdad necesaria (al estilo de Kripke) (Dowe, 2000, p. 4). Dada esta distinción entre análisis de tipo conceptual y de tipo empírico, Dowe sostiene que el proyecto que desarrolla se enmarca en uno del segundo tipo: “el proyecto de este libro [*Physical Causation* (2000)] es desarrollar un análisis empírico de la causación” (Dowe, 2000, p. 2)¹⁴¹. Ahora bien, como ha sido señalado, un análisis empírico, según Dowe, puede ser contingente o necesario. En este sentido el autor afirma que la TCC es un análisis de la causación de tipo contingente y enfatiza en que sería justamente esta característica de su teoría la que le brinda el sentido de “empírico” a su proyecto (Dowe, 2000, p. 4).

7.2.2. (b) TCC: una teoría no metafísicamente necesaria sino contingente

Como ya ha sido discutido en el Capítulo 4, Sección 4.7, la TCC enfrenta una serie de críticas referidas a la compatibilidad de esta teoría con los escenarios relativistas generales, pues en esta teoría física no se contaría con genuinas leyes de conservación (leyes de conservación integrales), las cuales son esenciales para la TCC y para cualquier teoría de la causación física que pretenda caracterizar las relaciones causales mediante cantidades físicas conservadas. En efecto, los conceptos de “posesión”, “intercambio” y “cantidad conservada”, fundamentales para la TCC, dependen de leyes integrales de conservación. En particular, la identidad enunciada por Dowe en su proposición *CCI* (y también la identidad contenida en la proposición *CC2*) entre “proceso causal” y “línea de mundo de un objeto que posee una cantidad conservada” dejaría de tener sentido si no se cuenta con genuinas leyes de conservación.

Por supuesto, Dowe se hace eco de estas objeciones. En el Capítulo 5 de su libro *Physical Causation* (2000), dedica algunas páginas a responder, en particular, a la crítica

¹⁴⁰ “All these philosophers agree that to understand causation we need to go beyond words, to look at the world. This is what is intended by the label ‘empirical analysis’”.

¹⁴¹ “The project of this book is to develop an empirical analysis of causation”.

formulada por Rueger (1998), la cual, recuérdese, se relaciona con el carácter supuestamente local que Dowe predica de su teoría. Rueger sostiene que, como no en todos los espacio-tiempos relativistas generales es posible enunciar leyes de conservación integrales, sino únicamente en aquellos que presenten las simetrías adecuadas, entonces la TCC depende fuertemente del espacio-tiempo de fondo, lo cual le impide ser una teoría local como Dowe señala. Para Rueger, en consecuencia, existen dos opciones para la TCC: o bien insistir que la relación causal caracterizada por la TCC es local y concluir, por lo tanto, que no existen relaciones causales genuinas en el mundo, o bien abandonar el carácter local de la TCC. Más allá del problema de la localidad de la TCC, en esta sección se presentan las respuestas de Dowe a Rueger, puesto que es en el marco de esta defensa que la posición del autor frente a la necesidad o contingencia de la teoría que postula se manifiesta más claramente. Se revisan a continuación, detalladamente, los contraargumentos del autor.

En primer lugar, para dar una respuesta a la crítica formulada por Rueger, Dowe afirma que la identidad entre “proceso causal” y “línea de mundo de un objeto que posee una cantidad conservada”, tal como se establece en la proposición *CC1* (y, se deduce, también la identidad que se establece en la proposición *CC2*), es una identidad *contingente* y no metafísicamente necesaria. Además, enfatiza, por un lado, que esta identidad se sostiene en el mundo actual y en aquellos mundos que presenten las mismas leyes que el actual, y que esta característica de su teoría se encuentra alineada con el carácter “empírico” del análisis que lleva a cabo; por otro lado, sostiene que el carácter contingente de la identidad depende de las leyes de la naturaleza (del mundo actual). En palabras del propio autor:

“La identidad de ‘proceso causal’ con ‘línea de mundo de un objeto que posee una cantidad conservada’ es contingente, y no metafísicamente necesaria. La hipótesis es que en nuestro mundo, y en mundos suficientemente cercanos, como la mayoría de los que obedecen nuestras leyes, un proceso causal es la línea de mundo de un objeto que posee una cantidad conservada. Dejamos de lado la pregunta de hasta donde podemos alejarnos de lo actual antes que nuestra hipótesis deje de tener sentido. Al llamar a esto un análisis empírico

(ver Capítulo 1), enfatizamos la prioridad de la afirmación de que la identidad se sostiene en lo actual. Al llamar al análisis una identidad contingente, queremos decir que es contingente respecto de las leyes de la naturaleza y quizás incluso de las condiciones de borde.” (Dowe, 2000, p. 95)¹⁴².

En relación con esta afirmación, Dowe ofrece el siguiente ejemplo. Supongamos que $\{q_a, q_b, q_c, q_d\}$ es un conjunto de cantidades conservadas en el mundo actual W_a , y consideremos un mundo W_e donde ninguna de estas cantidades se conserva y donde existe una cantidad conservada q_e . ¿Es la línea de mundo de un objeto en W_e , que posee la cantidad conservada q_e , un proceso causal en W_a ? ¿O es en W_e , la línea de mundo de un objeto que posee q_a , por ejemplo, un proceso causal? La respuesta del autor a estas preguntas es, simplemente, que su teoría no nos puede decir nada al respecto (Dowe, 2000, p. 96).

De acuerdo con Dowe, lo afirmado en los párrafos anteriores le permite responder a la crítica de Rueger ofreciendo una tercera alternativa a las dos propuestas por este último autor. En efecto, Dowe sostiene que la salida a la encrucijada de Rueger consiste en afirmar que la TCC es una hipótesis contingente (en particular, las identidades enunciadas en las proposiciones *CC1* y *CC2*) respecto de las leyes de la naturaleza. De este modo, según el autor, si se da el caso de no poder contar con leyes de conservación, entonces la TCC sería refutada¹⁴³. En sus palabras:

“Sin embargo, hay una tercera opción, que se desprende de lo que ya dije. La teoría de las Cantidades Conservadas es una hipótesis contingente, contingente respecto de las leyes de la naturaleza, por ejemplo. Esto significa que si las

¹⁴² “*The identity of 'causal process' with 'the world line of an object that possesses a conserved quantity' is contingent, and not metaphysically necessary. The hypothesis is that in our world, and in close enough worlds, such as most of those that obey our laws, a causal process is the world line of an object that possesses a conserved quantity. We leave aside the question of how far we can stray from actuality before this hypothesis stops making sense. In calling this an empirical analysis (see Chapter 1), we emphasize the priority of the claim that the identity holds in actuality. In calling the analysis a contingent identity, we mean that it is contingent on the laws of nature and perhaps even on boundary conditions.*”

¹⁴³ La supuesta refutabilidad de la teoría será presentada un poco más abajo en esta misma sección y analizada en la Sección 7.5.

leyes resultaran ser de cierta manera, la teoría sería refutada. Este puede ser el caso si resulta que en realidad no hay leyes de conservación. (Dowe, 2000, p. 97)^{144,145}.

De este modo, Dowe argumenta que, si hay espacio-tiempos relativistas generales donde no se cumplen los principios de conservación, esto no implica que tales principios fallen en nuestro mundo: el espacio-tiempo exhibe las simetrías correctas y, por lo tanto, los principios de conservación se cumplen en nuestro mundo:

“Pero el hecho de que haya espacio-tiempos relativistas generales en los cuales las leyes de conservación globales no se cumplen no implica que tales leyes fallen en nuestro mundo. Si lo hacen o no depende de la estructura *actual* del espacio-tiempo, y en particular, si ciertas simetrías se cumplen. Tal como yo lo entiendo, nuestro espacio-tiempo exhibe las simetrías correctas; las leyes globales de conservación sí se cumplen en nuestro universo hasta donde sabemos. Considero, entonces, que la teoría de las cantidades conservadas no está refutada.” (Dowe, 2000, p. 97)¹⁴⁶.

Dowe avanza incluso más allá cuando sostiene que no se ha mostrado que los principios de conservación fallen en todos los mundos posibles. Por ello, sugiere que la identidad entre “proceso causal” y “línea de mundo de un objeto que posee una cantidad conservada” podría cumplirse en todos los mundos físicamente posibles, donde los mundos físicamente posibles, para el autor, son aquellos donde se cumplen las mismas leyes de la

¹⁴⁴ “However, there is a third option, which follows from what I have already said. The Conserved Quantity theory is a contingent hypothesis, contingent on the laws of nature, for example. This means if the laws turned out to be a certain way, the theory would be refuted. This may be the case if it turns out that there actually are no conservation laws”.

¹⁴⁵ Un aspecto desconcertante en la presentación de Dowe en relación a las características y alcance de su TCC es que, por un lado, el autor afirma que su teoría nada puede decir respecto de la causación en aquellos mundos donde se cumplen leyes de conservación distintas a las del mundo actual, pero, por otro lado, afirma que su teoría sería refutada en aquellos mundos donde resultara no haber leyes de conservación. Este aspecto será analizado en las secciones siguientes.

¹⁴⁶ “But the fact that there are general relativistic spacetimes in which global conservation laws do not hold does not entail that global conservation laws fail in our world. Whether they do or not depends on the actual structure of spacetime, and in particular whether certain symmetries hold. As I understand it, our spacetime does exhibit the right symmetry; global conservation laws do hold in our universe as far as we know. I take it, then, that the conserved quantity theory is not refuted”.

naturaleza que en el mundo actual (Dowe, 2000, p. 97). De esta forma, Dowe argumenta que, al afirmar que existen espacio-tiempos no simétricos en RG, solo se está diciendo que esa es una solución de las ecuaciones de campo de Einstein, pero no que representan un mundo físicamente posible en el sentido en que él lo define:

“He sugerido que esta explicación debería probablemente cumplirse en todos los mundos físicamente posibles, esto es, en todos los mundos que tiene las mismas leyes de la naturaleza que el nuestro. ¿Ha mostrado Rueger que esto no es así? De ninguna manera. Decir, por ejemplo, que los espacio-tiempos no simétricos son posibles puede ser erróneo. Esto significa simplemente que es una solución de las ecuaciones de la Teoría General de la Relatividad. Pero esto no significa que un mundo así sea físicamente posible en el sentido dado aquí. Si tal mundo viola otras leyes que se mantienen en el mundo actual, entonces ese mundo no es físicamente posible.” (Dowe, 2000, p. 97)¹⁴⁷.

Es importante subrayar que, como quizás se podría inferir de la discusión de los párrafos anteriores, el carácter contingente de la TCC no es una característica que el autor adjudique a su teoría a raíz de las críticas que se le formulan respecto de los problemas que presenta en la RG; por el contrario, los postulados asumidos por Dowe en su proyecto para una teoría de la causación se remontan a la propia génesis de su teoría. En efecto, como se ha señalado más arriba, Dowe presenta su TCC como una alternativa a la teoría de transmisión de marca de Salmon (ver Capítulo 2), que pretendía resolver algunos de los problemas que esta última presentaba. Ahora bien, como el propio Dowe declara, en la formulación de la TCC adopta gran parte de los supuestos asumidos por Salmon en su proyecto, y entre ellos asume el carácter contingente de su teoría de la causación. Efectivamente, en relación a este punto, Dowe afirma:

¹⁴⁷ “I have suggested that the account should probably hold in all physically possible worlds, that is, in all worlds that have the same laws of nature as ours. Has Rueger shown that this is not so? Not at all. To say, for example, that nonsymmetric spacetimes are possible can be misleading. It means simply that it is a solution to the equations of the General Theory of Relativity. But this doesn't mean that such a world is a physically possible world in the sense given here. If such a world violates other laws that hold in the actual world, then that world is not physically possible.”

“El objetivo general de Salmon es ofrecer una teoría que sea consistente con las siguientes afirmaciones: (i) la causalidad es una característica objetiva del mundo; (ii) *la causalidad es una característica contingente del mundo*; (iii) una teoría de causalidad debe ser consistente con la posibilidad de indeterminismo, (iv) la teoría debe ser (en principio) independiente del tiempo para que sea consistente con una teoría causal del tiempo; (v) la teoría no debe violar las restricciones de Hume respecto de "poderes ocultos". Para los propósitos de este documento [aquí Dowe presenta, por primera vez, su TCC], *estos objetivos generales serán aceptados sin crítica*, aunque si Salmon ha logrado (v) estará sujeto a escrutinio.” (Dowe, 1992a, p. 195; énfasis incluido con fines expositivos)¹⁴⁸.

Antes de pasar a exponer la última de la tesis de Dowe, resulta necesario destacar que, si bien el autor no menciona explícitamente que su análisis de la causación es *metafísicamente* contingente, esto se infiere directamente del planteamiento acerca de cuál es su proyecto general acerca de la causación. En efecto, ya que la pretensión de Dowe es establecer una teoría acerca de qué es la causación en el mundo (no elucidar los usos de la palabra “causa” en lenguaje cotidiano, lo cual podría ser falso por motivos empíricos), entonces parece perfectamente razonable sostener que la contingencia de su teoría es de carácter metafísico y no de otro tipo. Por otra parte, la contingencia no podría interpretarse en un sentido *físico* puesto que, como Dowe afirma en una de las citas anteriores, su teoría se presenta como físicamente necesaria, es decir, como verdadera en todos los mundos donde se cumplen las leyes del mundo actual.

¹⁴⁸ “Salmon's broad objective is to offer a theory which is consistent with the following assertions: (i) causality is an objective feature of the world; (ii) causality is a contingent feature of the world; (iii) a theory of causality must be consistent with the possibility of indeterminism; (iv) the theory should be (in principle) time-independent so that it is consistent with a causal theory of time; (v) the theory should not violate Hume's strictures concerning "hidden powers". For the purposes of this paper these broad objectives will be accepted uncritically, although whether Salmon has achieved (v) will be subject to scrutiny.”

7.2.3. (c) TCC: una teoría refutable

Como queda claro a partir de la discusión de la tesis (b), la TCC presenta una tercera característica: es una teoría que puede ser refutada. Según Dowe, esto es lo que hace que la TCC, y en particular las identidades establecidas en las proposiciones *CC1* y *CC2*, posean contenido empírico. En dos de las citas presentadas más arriba ya se ha podido observar esta característica que Dowe atribuye a su teoría. No obstante, con el objetivo de señalar claramente el aspecto que se pretende destacar aquí, se repite una parte de cada una de aquellas dos citas, enfatizando el fragmento del texto que es de interés en esta ocasión:

“Tal como yo lo entiendo, nuestro espacio-tiempo exhibe las simetrías correctas; las leyes globales de conservación sí se cumplen en nuestro universo hasta donde sabemos. Considero, entonces, que la *teoría de las cantidades conservadas no está refutada.*” (Dowe, 2000, p. 97; énfasis incluido con fines expositivos).

La característica de refutabilidad de la TCC que Dowe le imputa a su teoría vuelve a ser mencionada en el párrafo siguiente:

La teoría de las cantidades conservada es una hipótesis contingente, contingente respecto de las leyes de la naturaleza, por ejemplo. *Esto significa que si las leyes resultaran ser de cierta manera, la teoría sería refutada. Este puede ser el caso si resulta que en realidad no hay leyes de conservación.* (Dowe, 2000, p. 97; énfasis incluido con fines expositivos).

En las siguientes secciones se analizará cada una de estas tesis y se argumentará que no solo resultan inconsistentes entre sí, sino que también son incompatibles con el enfoque general que Dowe asume en su teoría de la causación.

7.3. ¿Análisis empírico?

Respecto de la tesis (a) de Dowe, surge un primer problema de tipo epistemológico: dilucidar con toda claridad el tipo de análisis de la causación que el autor lleva a cabo. En este sentido, la etiqueta utilizada por Dowe de “análisis empírico” para designar el tipo de

análisis realizado en su TCC es un tanto engañosa y lleva a confusión con relativa facilidad. En efecto, como se ha señalado más arriba, el objetivo del autor es intentar establecer *qué* es la causación en el mundo, es decir, pretende descubrir la naturaleza de la relación causal como una característica objetiva de lo real; en consecuencia, esta tarea quedaría mejor caracterizada por la etiqueta “análisis metafísico” o “análisis ontológico” que por la utilizada por el autor. Al denominar “análisis empírico” a la tarea que emprende con su teoría de la causación, Dowe parece sugerir que *todo* análisis metafísico es un análisis empírico en el sentido por él definido. Sin embargo, esto no tiene porqué ser siempre de ese modo. Por lo tanto, para evitar cualquier tipo de confusión, parece más adecuado que un análisis que aspira a develar la naturaleza de un cierto ítem con contenido filosófico, cuya existencia en el plano de lo real se considera independiente de los sujetos, se denomine “análisis metafísico” y no “análisis empírico”. El propio Dowe ofrece una clara pista en favor de esta interpretación. En una nota a pie de página en su libro *Physical Causation* (2000), afirma que la distinción entre “análisis conceptual” y “análisis empírico” que él introduce es presentada por otros autores como distinción entre “análisis semántico” y “análisis metafísico” (Bigelow y Pargetter, 1990) o “análisis conceptual” y “análisis metafísico” (Jackson, 1994), respectivamente (Dowe, 2000, p. 1).

Como ya se ha insinuado en el párrafo anterior, este aparente embrollo conceptual no responde meramente a una confusión terminológica inocua, sino que, por el contrario, pone de manifiesto la posición filosófica del autor respecto de la relación entre la física y la metafísica, relación que juega un papel central al momento de encontrar una salida al problema de la definición abordado en este capítulo. Efectivamente, detrás del hecho de denominar “análisis empírico” aquello que claramente corresponde a un análisis metafísico parece esconderse el siguiente supuesto:

(S): Las ciencias particulares (específicamente la física) y sus métodos representan la única forma de acceso a lo real.

Matteo Morganti y Tuomas Tahko (2017) ofrecen una clasificación de los diferentes enfoques respecto de la relación entre ciencias particulares o empíricas y metafísica sobre la base de los métodos que emplean y los campos de estudio que abordan. Según los autores, se pueden reconocer cuatro posiciones diferentes respecto de la relación entre

ambos ámbitos. Según la primera posición, (i) las ciencias particulares y la metafísica difieren tanto en sus métodos como en sus campos de estudio. Bajo esta mirada, difieren en sus métodos pues mientras la metafísica es una actividad puramente a priori, las ciencias especiales se desarrollan bajo métodos puramente a posteriori. Además, difieren en sus campos de estudio, puesto que la metafísica no puede o no debe preocuparse de investigar la naturaleza y su estructura (por ejemplo, la metafísica se concibe como un análisis puramente conceptual). De acuerdo con la segunda posición, (ii) la metafísica es una disciplina desde la cual sí es posible conocer la naturaleza o la estructura de la “realidad”, al igual que la física (o en conjunción con ella), pero con métodos independientes a los de la ciencia. Bajo esta visión, si bien la física sigue siendo una ciencia fundamentalmente a posteriori y la metafísica una a priori, coinciden en que a partir de ambas disciplinas es posible conocer, en algún sentido, la estructura de lo real. Es decir, este enfoque piensa a la metafísica como una actividad independiente de la física desde la cual es posible investigar el mundo. Desde la tercera posición, (iii) la metafísica, a diferencia de las ciencias particulares, se ocupa de cuestiones puramente conceptuales, pero utilizando los mismos métodos que la primera (por ejemplo, la filosofía experimental). Por último, desde una cuarta posición, (iv) la metafísica se concibe como una disciplina completamente naturalista que solo tiene valor si provee hipótesis comprobables por medios empíricos y/o si se encuentra basada en la observación y teorización que se lleva a cabo en las ciencias particulares. Bajo esta perspectiva, se piensa a la metafísica como de la misma naturaleza que las ciencias particulares, pues ambas coincidirían tanto en sus métodos como en sus campos de estudio. La metafísica queda así reducida a la teorización científica.

Si bien se han propuesto diferentes versiones del naturalismo metafísico –como el naturalismo extremo, que considera que la *única* forma de conocer lo real es mediante las ciencias particulares, o como el naturalismo moderado, que considera que la *mejor* (en el sentido de objetividad) forma de conocer la realidad es la que provee la ciencia empírica–, en general, las posiciones naturalistas quedan bien caracterizadas por el enfoque (iv). En este sentido, el supuesto (S), sobre el cual descansa el tipo de análisis que Dowe lleva a cabo en su teoría de la causación, encaja bien bajo un enfoque de este tipo. Dowe parece entender que la única manera de hacer metafísica es brindando hipótesis comprobables por medios empíricos y, en consecuencia, espera que el análisis de la causación que propone

posea las mismas características que las obtenidas mediante la teorización y la práctica científica. Esto explica muy bien las tesis (b) y (c) enunciadas más arriba, que se derivan o dependen de la tesis (a).

Sin embargo, como se ha mostrado anteriormente, la posición que abraza Dowe respecto al estatus de la metafísica respecto de la física no es la única posible. Como se argumentará en las secciones siguientes, una teoría de la causación en términos de cantidades físicas conservadas (como la TCC) que descansa en un enfoque del tipo (ii) no solo se muestra compatible con la RG, sino que evita muchos de los problemas que presenta la TCC tal como es formulada por Dowe. Es importante destacar que no por el hecho de concebir la metafísica como una disciplina independiente de la física (como en el caso del enfoque (ii)) sus resultados entrarán en contradicción con los resultados de las ciencias empíricas. En efecto, tal como muestran Morganti y Tahko (2017), es posible concebir a la metafísica como una disciplina autónoma respecto de las ciencias especiales pero que puede ser considerada como naturalista y, por lo tanto, científicamente informada. De todos modos, para los fines de esta investigación no se requiere que un enfoque de este tipo (el enfoque (ii)) acerca de la naturaleza de la metafísica sea subsumido bajo el rótulo de “naturalista”; es suficiente con el requerimiento más débil de que un análisis metafísico se encuentre científicamente bien informado y que, en consecuencia, sus resultados no entren en contradicción con los postulados de las ciencias empíricas, en particular con los de la física. Bajo esta premisa, si se pretende encontrar una solución al problema de la definición abordado en esta sección y, en consecuencia, al problema general de esta tesis, un primer paso consiste en remplazar la tesis (a) (La TCC es un análisis empírico y no conceptual) por la siguiente tesis alternativa:

(a’): Una investigación acerca de qué es la causación en el mundo debe desarrollarse bajo un análisis de tipo metafísico.

Por supuesto, en la tesis (a’) se entiende por “análisis metafísico” a aquel caracterizado por el enfoque (ii) expuesto previamente. Es decir, se asume que la metafísica, que se ocupa, entre otras cosas, justamente de conceptos como el de causación, puede proveer una adecuada definición de la naturaleza de la causación tal como se manifiesta en el mundo. Este enfoque, como ya fue señalado, reconoce la autonomía e

independencia de la metafísica respecto de la física en tanto disciplina a priori, a diferencia de las ciencias empíricas; sin embargo, se trata de análisis metafísicos científicamente informados. La tesis (a'), en consecuencia, descansa sobre el siguiente supuesto (S'), que reemplazaría al supuesto (S):

(S'): La metafísica goza de independencia y autonomía, en relación a sus métodos, respecto de la física, en su acceso a la estructura de la realidad.

Asumir la tesis (a') bajo el supuesto (S') conforma el primer paso hacia una solución al problema que se aborda en este capítulo y hacia lograr que una teoría de la causación física, como la TCC, resulte compatible con los escenarios relativistas generales. Los pasos restantes para arribar finalmente a una solución satisfactoria serán abordados en la Sección 7.6. Antes de ello, en lo que resta de esta sección y en las siguientes, se analizará la tesis (a) (y el supuesto (S)) de Dowe y se expondrán las consecuencias problemáticas a las que se arriba al asumirla.

Si se pretende analizar la causación bajo la tesis (a), tal como lo hace Dowe, entonces el supuesto (S) obliga a que todo lo que se diga acerca del mundo debe encontrarse en la estructura, leyes y conceptos fundamentales de la física. Para ser más precisos, bajo el alero de una posición naturalista extrema o cuasi-eliminativista de la metafísica (en la que, como se ha visto, se posiciona Dowe), el ítem filosófico a analizar debería encontrarse en el cuerpo teórico o conceptual de las ciencias particulares, en este caso, de la física. Sin embargo, y aquí aparece un importante escollo para Dowe, existen buenas razones para poner en duda que el concepto “causa” sea parte de alguna teoría física¹⁴⁹. Como es bien conocido, Russell fue uno de los primeros autores en cuestionar explícitamente el papel que desempeña la causación en la ciencia física. En su reconocido artículo “On the notion cause” el filósofo británico afirma:

“En los movimientos de cuerpos mutuamente gravitantes, no hay nada que pueda ser llamado causa y nada que pueda ser llamado efecto; hay simplemente una fórmula. Se pueden encontrar ciertas ecuaciones diferenciales, las cuales se mantienen en cada instante para cada partícula del sistema, y que, dada la

¹⁴⁹ Para una crítica en este mismo sentido, ver Miguel y Paruelo (2004).

configuración y las velocidades en un instante, o las configuraciones en dos instantes, hacen que la configuración en cualquier otro anterior o posterior sea instantáneamente calculable teóricamente [...] Esta afirmación es válida en toda la física, y no solo en el caso especial de la gravitación. Pero no hay nada que pueda llamarse adecuadamente "causa" y nada que pueda llamarse adecuadamente "efecto" en dicho sistema." (Russell, 1912/1913, p. 14).

En la actualidad, existe una posición más o menos estándar respecto al lugar que ocupa la causación en las ciencias empíricas y, en particular en la física, influenciada en gran medida por el trabajo de Russell. De acuerdo con este punto de vista, si bien la noción de causa puede encontrar un lugar (secundario) en la física madura actual mediante la imposición de ciertos modelos causales externos a las teorías, este concepto no formaría parte de la estructura fundamental de las teorías físicas (ver, por ejemplo, Price y Corry, 2007; Frisch, 2012, 2014). Por ejemplo, Norton (2007) afirma que:

"...insisto en que los conceptos de causa y efecto no son los conceptos fundamentales de nuestra ciencia y que la ciencia no se rige por una ley o principio de causalidad. [...] Lo que sí niego es que la tarea de la ciencia es encontrar las expresiones particulares de algún principio causal fundamental en el dominio de cada una de las ciencias. [...esta forma de escepticismo causal] Está motivado por tomar el contenido de nuestras teorías científicas maduras seriamente." (Norton, 2007, p. 12).

No obstante, es pertinente mencionar que algunos filósofos de la física han argumentado en contra de esta posición, afirmando que en la física fundamental sí hay un lugar para la causación. En general, los argumentos de estos autores descansan en la idea de que la causación tiene sentido en la física fundamental, o puede ser "anclada" a ella a través del concepto de "asimetría" de la causación, el cual sí puede ser representado por los conceptos y herramientas formales de la física (ver Bartels y Wohlfarth, 2014; Hubert y Poellinger, 2014). Sin embargo, aun cuando estos trabajos estuvieran en la dirección adecuada, resultan totalmente insuficientes para argumentar que la física puede informar adecuadamente, únicamente a través del concepto de asimetría causal, acerca de la *naturaleza misma* de la causación. En otras palabras, no brindan una base suficiente para

demostrar que, a partir del concepto de asimetría causal, se puede desarrollar una teoría que permita establecer adecuadamente qué es la causación en el mundo, tal como pretende Dowe. Por lo tanto, este es un problema central para quienes pretendan estudiar la causación desde un enfoque naturalista que reduce la metafísica a la física. Se retomará el problema de la ausencia del concepto de causa en física y los problemas que esto conlleva para la TCC en la Sección 7.5.

En las dos secciones siguientes, se argumentará que, incluso si el problema anterior fuera superado, la TCC bajo la tesis (a) y el supuesto (S) presenta problemas adicionales aún más graves que el presentado aquí.

7.4. Definición de causación: ¿real o nominal?

De acuerdo con lo presentado en la Sección 7.2, resulta claro que la intención de Dowe con su TCC es establecer una teoría que capture la noción de causación tal como se manifiesta en lo real y no en el lenguaje. Dentro de este marco general, el autor ofrece dos proposiciones para responder a la pregunta “¿qué son los procesos e interacciones causales?” (Dowe, 2000, p. 89), las cuales forman el núcleo de su teoría. Si bien estas proposiciones ya fueron presentadas en el Capítulo 3, se enuncian nuevamente aquí dada su relevancia para la argumentación de esta sección:

CC1: Un proceso causal (*causal process*) es una línea de mundo de un objeto que posee una cantidad conservada.

CC2: Una interacción causal (*causal interaction*) es una intersección de líneas de mundo que involucra intercambio de una cantidad conservada.

Si bien Dowe presenta estas proposiciones como *definiciones* (Dowe, 1992a, p. 210; Dowe 1992c, p. 126; Dowe, 2000, p. 199) resulta pertinente, dado que se están analizando los fundamentos de la TCC, preguntarse qué tipo de definición dichas proposiciones suministran.

En general, las definiciones se clasifican en dos tipos: *reales* (o *de re*) o *nominales* (o *de dicto*). En las del primer tipo se busca establecer la esencia de la cosa que se pretende definir, es decir, aquello que hace que la cosa sea lo que es y posea ciertas propiedades. Por

lo tanto, para establecer una definición real se debe investigar la cosa (o cosas) denotada por el término que se pretende definir. En cambio, en las del segundo tipo se busca definir la idea que los hablantes asocian a la palabra que se está definiendo. En consecuencia, para brindar una definición nominal de un término se debe investigar su significado y su uso en el lenguajes (Gupta, 2019). Por lo tanto, si Dowe pretende brindar una teoría de la causación que establezca *qué es* la causación en la realidad, entonces, con las proposiciones *CC1* y *CC2*, el autor ofrece una *definición real* de “proceso causal” e “interacción causal”. Ello se encuentra en resonancia con su explícito propósito de alejarse del “análisis conceptual, que “es un análisis de significado que comienza con nuestra comprensión cotidiana, de sentido común, del concepto relevante” (Dowe, 2000, p. 2).

No obstante, cuando se considera la argumentación general del autor, estas definiciones no parecen desempeñarse como tales de acuerdo con las características que Dowe le atribuye a su teoría de la causación. Esta inconsistencia constituye un segundo problema de la TCC, que se suma al presentado en la sección anterior. A continuación se presenta en detalle este segundo inconveniente.

Recuérdese de la Sección 7.2 que la tesis (b) de Dowe afirma que su teoría tiene un carácter contingente y, en consecuencia, la identidad entre “proceso causal” y “línea de mundo de un objeto que posee una cantidad conservada” (en adelante, identidad *p*)¹⁵⁰ es una *identidad metafísicamente contingente* y no metafísicamente necesaria. Puesto que Dowe, como se señaló más arriba, pretende ofrecer *definiciones reales* de “proceso causal” y de “interacción causal”, entonces ser una línea de mundo de un objeto que posee una cantidad conservada constituye la esencia misma de la causación: es lo que hace que una línea de mundo sea un proceso causal. En otras palabras, lo que hace que una identidad (como *p*) sea una definición real es el hecho de ser *metafísicamente necesaria*: si bien pueden existir mundos lógicamente posibles donde la esencia de proceso causal no sea la recogida en la definición, lo que la definición real expresa es aquello que hace que la causación sea lo que es y no otra cosa, y en ello radica su necesidad metafísica. Sin

¹⁵⁰ En adelante, y durante todo lo que resta de este capítulo, en el análisis que se lleve a cabo se hará referencia a la identidad *p*, la cual tiene su origen en la proposición *CC1*, sin apelar a la proposición *CC2*. Sin embargo, este hecho responde única y exclusivamente a una cuestión de claridad expositiva, pues exactamente el mismo análisis y las mismas conclusiones pueden ser aplicados a la identidad correspondiente a la proposición *CC2*.

embargo, como se ha señalado, Dowe ha puntualizado, inequívocamente, que la identidad p es metafísicamente contingente y no metafísicamente necesaria; por lo tanto, el autor, en los hechos, no brinda una definición real de “proceso causal” e “interacción causal”, esto último, en absoluta contraposición con sus lineamientos generales.

Pero los problemas con la teoría de la causación de Dowe no acaban aquí. Si p no es una definición real de proceso causal, lo que resta es suponer que p constituye una *mera caracterización* de los procesos causales que permite identificar, en el entorno en el que vivimos, qué línea de mundo es un proceso causal y cuál no lo es, pero sin explicar el motivo para ello; en otras situaciones, el criterio suministrado podría dejar de ser válido o adecuado. Para expresar esta idea en el lenguaje de los mundos posibles, diríamos que p es una identidad *metafísicamente contingente*: si bien verdadera en nuestro mundo actual, existen mundos metafísicamente posibles donde la identidad p es falsa. Todo parece indicar que esta es la interpretación que Dowe hace de sus proposiciones *CC1* y *CC2* en su respuesta a las críticas provenientes de las discusiones acerca de la validez de los principios de conservación en RG.

Si bien una posición en la que se pretende suministrar únicamente una caracterización para identificar un cierto ítem filosófico en el mundo actual es, en sí misma, una empresa filosóficamente legítima, no parece serlo en el caso de las pretensiones de la TCC. Bajo las tesis asumidas por Dowe, una estrategia como esta es problemática, al menos, por dos razones. En primer lugar, la TCC no suministra criterios definicionales para determinar cuándo la identidad p sería falsa. En efecto, puesto que el autor explícitamente afirma que p es una identidad *metafísicamente contingente* (contingente respecto de las leyes de la naturaleza), parece razonable suponer que p será *falsa* en ciertos mundos metafísicamente posibles, pero físicamente imposibles, es decir, en aquellos mundos donde no se cumplen las leyes físicas del mundo actual. En particular, según Dowe, esto sucede en aquellos mundos metafísicamente posibles que *no poseen leyes de conservación*. Ahora bien, para que una identidad como p sea falsa, es necesario que los dos términos de la identidad refieran a clases distintas; en este caso, sería necesario que la clase de los “procesos causales” y la clase de las “líneas de mundo de objeto que poseen una cantidad conservada” no fueran la misma clase. Pero esto supone que existen criterios definicionales que permiten identificar las clases así relacionadas independientemente de la identidad cuya

verdad se pretende establecer. Para el caso que se está analizando, se debería contar con elementos para identificar la clase de los procesos causales y la clase de las líneas de mundo de objetos que poseen una cantidad conservada con independencia de p . Pero estos criterios definicionales no son suministrados por la TCC, pues es el propio enunciado p el que caracteriza los procesos causales.

Ahora bien, en el marco del pensamiento de Dowe, existe otra forma de que un mundo sea metafísicamente posible pero físicamente imposible: es el caso de aquellos mundos donde *no se cumplen las mismas leyes de conservación* que en el mundo actual. En este caso (recuérdese el ejemplo del propio Dowe presentado en la Sección 7.2) Dowe afirma que su teoría no permite hacer afirmación alguna. Esta elección de Dowe también es problemática. Efectivamente, no resulta claro cómo comprender este agnosticismo por parte del autor; ya que, o bien este hecho puede ser interpretado como una incapacidad de la TCC de capturar la naturaleza de los procesos causales en ciertas situaciones y, por lo tanto, como un indicio de una teoría de la causación “incompleta”, o bien, esta salida nos fuerza a sostener la incómoda situación de asumir que en los mundos físicamente imposibles p carece de valor de verdad.

La segunda razón por la cual suponer que Dowe proporciona una mera caracterización contingente de la causación es problemática se debe a que ello conduce a una contradicción con los objetivos iniciales del autor, esto es, llevar a cabo un análisis metafísico que conlleve dilucidar la naturaleza de la causación. En efecto, si una teoría de la causación es contingente en el sentido que propone Dowe, esto es, solo se aplica al mundo actual, entonces no es una teoría que pretenda develar qué es la causación. Sería más bien una teoría que únicamente brinda un criterio para identificar procesos causales e interacciones causales en nuestra realidad contingente, como sucede, por ejemplo, en el caso de la teoría intervencionista de la causación de Woodward (2003). Acerca de su propia teoría de la causación, Woodward afirma que la naturaleza de su proyecto se aleja tanto de un análisis empírico en el sentido definido por Dowe como de un análisis conceptual (2003, p. 7). Por esta misma razón, Michael Strevens (2007) argumenta: “sería mucho mejor que Woodward, sugiero, presentara su enfoque no como una metafísica o una definición de la causación general [*type level causation*], sino como describiendo algo como un estereotipo al estilo de Putnam de nuestros conceptos causales, esto es, como una teoría de un tipo de

‘contenido’ que no fija condiciones de verdad” (p. 246). Claramente, el proyecto de Woodward se distancia de las pretensiones de ser una teoría metafísica de la causación que ofrezca la definición de qué es la causación en el mundo, y se acerca a una teoría que más bien ofrece un criterio para diferenciar, en el uso ordinario y científico, aquellas declaraciones que calificarían como causales de aquellas que no calificarían como tales. En este sentido, si atendemos al carácter contingente de la TCC, ésta parece calificar más como una teoría que ofrece meramente un criterio para identificar relaciones causales en el mundo actual, al estilo de Woodward, que como una teoría metafísica que revela la verdadera naturaleza de la causación, como pretende su autor.

7.5. La TCC: ¿una teoría refutable?

Como fue señalado en la Sección 7.2, Dowe presenta su teoría de la causación como una teoría con contenido empírico y, por lo tanto, refutable (tesis (c)). A continuación, argumentaremos que esta tesis también resulta como problemática, y ello por dos razones.

En primer lugar, para que la teoría fuera refutable, se debería contar con un medio independiente de la propia teoría para identificar los procesos causales. Muchas teorías de la causación encuentran ese criterio en los supuestos de sentido común que se expresan en el lenguaje cotidiano: si la teoría identifica como procesos causales los casos que el sentido común identifica como tales, y solo ellos, entonces la teoría será verdadera; en caso de alguna discordancia, la teoría será falsa. Pero esto es precisamente lo que Dowe rechaza, ya que busca una teoría que acuerde, no con el sentido común, sino con la física actual. Si se sigue a Dowe, entonces, la teoría sería refutable, pues el medio independiente para identificar los procesos causales es el que brinda la física: la teoría será falsa si no recoge el concepto de causa de las ciencias físicas. El problema de esta estrategia es que la física no cuenta con una definición técnica del concepto de proceso causal (ver Sección 7.3). Pero aún si se admitiera que, si no en la teoría, en la práctica de la física se utiliza el concepto de causación, no puede negarse que se lo hace de un modo informal e intuitivo, no demasiado lejano al uso que le otorga el sentido común expresado en el lenguaje ordinario. En definitiva, si una teoría de la causación es refutable, termina siéndolo porque se la contrasta con los supuestos pre-teóricos que están a la base del uso del término “causa” en el lenguaje

ordinario. En este sentido, la teoría no logra determinar qué es la causación en el mundo, sino que elucida el modo en que el sentido común utiliza el concepto de causa.

En segundo lugar, la presunta refutabilidad de la teoría de Dowe descansa sobre una visión distorsionada de la ciencia. En efecto, en su argumentación Dowe compara el concepto de causación con el de energía, afirmando que, si bien con origen en el lenguaje ordinario, la ciencia brindó una definición del concepto de energía cuyo origen se encontraría en la aplicación del método científico: “podemos decir que la aplicación del método científico de teorización y experimentación produjo un 'análisis empírico' de energía” (Dowe, 2000, p. 7). Sin embargo, esta afirmación es altamente discutible. Si bien es cierto que diversos términos científicos tenían un uso pre-científico en el lenguaje ordinario antes de su incorporación a una teoría científica particular –por ejemplo, en el lenguaje cotidiano los términos “fuerza”, “trabajo” y “energía” se encuentran aún hoy ligados a la idea de esfuerzo biológico para la realización de una tarea física–, cuando esos conceptos ingresan a la física a través de la mecánica clásica pierden sus connotaciones antropocéntricas o biológicas y adquieren un significado técnico preciso a través de una *definición científica*.

En otras palabras, no es adecuado afirmar que la aplicación del método científico produjo un análisis empírico del concepto de energía que ya se encontraba en el uso cotidiano del término: el método científico, en todo caso, *definió* un ítem físico y produjo un *análisis empírico* de sus propiedades y comportamiento. Cuando la física actual brinda una definición del término “energía”, el enunciado que la expresa ya no posee contenido empírico en el sentido de refutabilidad: la definición será más o menos útil, más o menos fructífera, pero no puede ser falsa; en la práctica de la física no importa ya cuál es la relación del término científico con su homónimo pre-científico, y no existen laboratorios destinados a testear empíricamente la definición de “energía”. Por otra parte, es importante recordar que no todos los términos científicos tienen una génesis como la indicada: si bien los términos de las teorías macroscópicas pueden encontrar su origen en el lenguaje cotidiano (además de la mecánica clásica, este es el caso de la termodinámica, con los términos “calor” y “temperatura”), la mayor parte de los conceptos de las teorías fundamentales de la física son nombrados mediante neologismos (por ejemplo, “electrón”,

“neutrino”), o palabras existentes pero utilizadas de un modo completamente diferente del original (“spin”, “color” aplicado a los quarks).

Sobre la base de los argumentos presentados en estas últimas tres secciones, en la siguiente sección se abordará la solución que se ofrece en el presente trabajo al problema de la definición tal como se presenta en la TCC.

7.6. Una propuesta de solución al *problema de la definición*

En las secciones anteriores se han analizado los problemas a los que la TCC se ve enfrentada bajo las tesis (a), (b) y (c) asumidas por Dowe. Se ha argumentado, en la Sección 7.3, que la adopción de la tesis (a) obliga a Dowe a suponer que es posible encontrar una definición técnica o científica de causa en la física actual, lo cual resulta altamente discutible de acuerdo con las discusiones actuales en filosofía de la física. En la Sección 7.4 se ha afirmado que la tesis (b) es problemática por dos razones. En primer lugar, debido a que el autor no provee los criterios definicionales necesarios para determinar cuándo las proposiciones *CC1* y *CC2* serán falsas. En segundo lugar, debido a que asumir que la TCC es una teoría contingente va en directa contradicción con las pretensiones iniciales del mismo autor de ofrecer una elucidación acerca de la naturaleza de la causación tal como se presenta en la realidad. Por último, en la Sección 7.4 se ha mostrado que la tesis (c) es igualmente problemática. Por un lado, puesto que deberíamos contar con medios independientes a la propia teoría para evaluar bajo qué circunstancia esta teoría sería falsa, pero tales medios no se encuentran disponibles debido, en parte, a las propias limitaciones que Dowe impone a su análisis de la causación. Por otro lado, debido a que la aparente refutabilidad de la TCC descansa sobre una visión imprecisa o distorsionada de la teorización y la práctica científica.

Al inicio de esta sección hemos afirmado que el origen de los problemas mencionados más arriba se encuentra en la tesis (a) asumida por Dowe, la cual esconde el supuesto (S) que el autor parece defender. Recordemos que el supuesto (S) afirma que las ciencias empíricas, en particular la física, representan la única forma de acceder a lo real. Bajo este supuesto, Dowe considera que la investigación metafísica exige la adopción de los métodos y las herramientas que brinda la física. Por esta razón, como argumentamos

más arriba, Dowe denomina “análisis empírico” a aquello que, de acuerdo con la terminología filosófica habitual, se denomina “análisis metafísico”. De este modo, la tesis (a) bajo el supuesto (S) implica la adopción de las tesis (b) y (c), pues éstas se encuentran en concordancia con el supuesto que convierte a la física en metafísica.

Como se adelantó al finalizar la Sección 7.3, las dificultades que aparecen al adoptar las tesis (a), (b) y (c) se disuelven si la tesis (a) es reemplazada por la tesis (a’), que sostiene que una teoría de la causación que pretenda establecer qué es la causación en el mundo se fundamenta en un análisis de tipo metafísico en toda regla, esto es, considerando la metafísica no como una actividad que puede ser casi completamente reducida a la física, sino como una disciplina independiente que permite, mediante sus propios métodos, acceder a la estructura de lo real (supuesto (S’)). Desde la perspectiva de esta nueva tesis, una teoría de la causación que pretenda develar aquello que es sustancial o esencial de las relaciones causales, tal como las encontramos en el mundo, debe ofrecernos *definiciones reales* de aquellas características que son significativas para caracterizar la naturaleza de la causación. Como tal, una teoría de la causación de este tipo tendrá el carácter de necesaria, pues una definición real, bajo un análisis metafísico, es a priori y, en consecuencia, necesaria. En una teoría *metafísica* de la causación, no hay lugar para la contingencia o refutabilidad de las definiciones ofrecidas.

Por lo tanto, una solución al *problema de la definición* consiste en aceptar el carácter *real* de la definición de causación, en particular de las proposiciones *CC1* y *CC2*, y dejar de lado las afirmaciones de Dowe acerca de la contingencia y la refutabilidad de la teoría. En otras palabras, un enfoque adecuado de la causación física, compatible con las teorías físicas del espacio-tiempo y en particular con la RG, debe ofrecer una *definición real, metafísicamente necesaria* y no una mera hipótesis contingente y refutable.

De acuerdo con la argumentación ofrecida en esta sección y en las precedentes, sería absolutamente razonable asumir que las proposiciones *CC1* y *CC2* sean consideradas como definiciones reales y no como meras caracterizaciones de los conceptos de “proceso causal” e “interacción causal” involucrados en ellas. En particular, este giro argumentativo permite a la TCC de Dowe adquirir un cierto grado de compatibilidad con los escenarios físicos que propone la RG. En consecuencia, quienes pretendan defender una teoría de la causación

física en contextos relativistas generales desearían adherir a la siguiente condición metafísica:

Cond3: Los conceptos nucleares de una teoría de la causación física como la TCC, en particular las proposiciones *CC1* y *CC2*, deben presentarse como definiciones reales y, por tanto, metafísicamente necesarias e irrefutables.

Por una parte, la imposición de esta condición trae consigo algunas consecuencias importantes para la TCC. En efecto, en nuestro mundo actual y en todos los mundos físicamente posibles, serán procesos causales aquellas líneas de mundo de objetos que poseen las cantidades conservadas según los principios de conservación (integrales) de nuestra física. En mundos físicamente imposibles, donde se cumplen principios de conservación (integrales) diferentes a los del mundo actual, serán procesos causales las líneas de mundo de objetos que poseen las cantidades conservadas según los principios de conservación de ese propio mundo. Y en los mundos físicamente imposibles donde no existen principios de conservación (integrales) para ninguna magnitud física, sencillamente no habrá procesos causales.

Por otra parte, en tanto definición real de proceso causal, la identidad p (recuérdese la discusión de la Sección 7.4) no puede ser refutada (como así tampoco la proposición *CC2*). Sin duda, esto no significa que estemos forzados a admitirlo. Una definición real, en tanto metafísica, si bien no posee contenido empírico de tal modo que podría ser falsa, puede ser abandonada por motivos filosóficos si se mostrara que no es útil, fructífera o suficientemente amplia, que atenta contra la unificación de diferentes enfoques teóricos o, en general, que no tiene la fecundidad teórica que se esperaba al momento de su formulación.

En el próximo capítulo, que cierra la tesis, se presentará una reconsideración general de las discusiones desarrolladas a lo largo de esta investigación, la cual permitirá identificar más claramente el alcance y las limitaciones de la TCC en la teoría de la gravitación de Einstein.

Capítulo 8. *Revisión general, conclusiones y perspectivas*

Con el objetivo de obtener una visión panorámica acerca del alcance y las limitaciones de la teoría de la causación de Dowe en nuestra mejor teoría física del espacio-tiempo, la RG, en este capítulo llevamos a cabo una revisión general de las condiciones *Cond1*, *Cond2* y *Cond3* y evaluamos sus consecuencias para la TCC. Antes de ello, sin embargo, es pertinente que recordemos algunas cuestiones importantes abordadas en los capítulos anteriores, poniendo especial énfasis en los Capítulos 5, 6 y 7 donde fueron presentadas las condiciones antes mencionadas.

En los Capítulos 1 y 2 hemos presentado el marco general de esta tesis, esto es, las teorías contemporáneas de la causación. En particular, en el Capítulo 1 hemos examinado las teorías de la causación etiquetadas bajo el rótulo de teorías *difference-making*. En esta categorización hemos incluido a las teorías regularistas, contrafácticas, probabilistas y de agencia e intervención. Por otro lado, en el Capítulo 2 hemos analizado las teorías fisicalistas de la causación, incluyendo las denominadas teorías de transferencia y teorías de procesos. A continuación, en el Capítulo 3 hemos profundizado en la principal teoría de procesos, y objeto de estudio de esta investigación, la TCC de Dowe. Luego, en el Capítulo 4 hemos hecho una revisión detallada de las principales críticas formuladas a la TCC y hemos puesto especial acento en la crítica que ha motivado esta investigación: la supuesta incompatibilidad de esta teoría de la causación en escenarios relativistas generales. Hemos mencionado que, como resultado de un análisis detallado de esta crítica a través del abordaje de tres problemas independientes derivados de ella, es posible realizar una evaluación general acerca de la aplicabilidad de la TCC en la RG. Los tres problemas analizados fueron el *problema de las leyes de conservación*, el *problema de la posibilidad física* y el *problema de la definición*.

En el *Problema de las leyes de conservación*, abordado en el Capítulo 5, hemos argumentado que las leyes de conservación en su forma integral, las relevantes para la TCC, se cumplen en una cierta clase de modelos o mundos físicamente posibles de la RG. Dada la crítica a la TCC respecto a su incompatibilidad con los “universos” relativistas generales, por presuntamente no contar estos últimos con leyes integrales de conservación,

el resultado anterior juega un papel central para evaluar el alcance de la TCC en la RG. En efecto, debido a que las proposiciones *CCI* y *CC2* se fundamentan en la noción de “cantidad conservada”, se hace necesario evaluar el estatus de las leyes de conservación integrales, las cuales permiten definir estas cantidades. La ausencia de este tipo de expresiones de conservación trae como consecuencia la inexistencia de este tipo de cantidades físicas. Sobre la base de la argumentación ofrecida en este capítulo, hemos postulado la siguiente condición:

Condl: En las teorías físicas del espacio-tiempo, en particular en la teoría general de la relatividad, las leyes de conservación integrales se cumplen en una cierta clase de sus modelos.

Las consecuencias para la TCC de la adhesión o rechazo de la condición anterior dependen del carácter de las proposiciones *CCI* y *CC2*, tal como argumentamos en el Capítulo 7.

A continuación, en el Capítulo 6, hemos examinado el *Problema de la posibilidad física*. Por un lado, hemos señalado que los mundos físicamente posibles, desde el punto de vista del Enfoque Estándar, son considerados como aquellas soluciones o modelos de las ecuaciones diferenciales que gobiernan los sistemas físicos en el contexto de una teoría determinada, con independencia de si en estos modelos se cumplen o no las leyes de conservación. Sin embargo, por otro lado, hemos argumentado que es más adecuado considerar, desde el punto de vista de la práctica y teorización física, que los mundos físicamente posibles son aquellos modelos que son soluciones de las ecuaciones dinámicas y que además cumplen con poseer leyes de conservación correctamente definidas. En efecto, una noción adecuada de “posibilidad física” es aquella que incluye en la concepción de lo “físicamente posible” ciertos factores extra-nómicos, entre los que se cuentan las leyes de conservación. Por lo tanto, un mundo físicamente posible será no solo aquel que resulta solución de las ecuaciones dinámicas, sino que debe cumplir además con poseer genuinas leyes de conservación, sean estas únicamente diferenciales o diferenciales e integrales. En consecuencia, desde el punto de vista del debate acerca de la modalidad en física, sería deseable imponer la siguiente condición a toda teoría de la causación física en términos de cantidades conservadas que apele a estas nociones modales:

Cond2: La noción de posibilidad física debe incluir ciertos elementos, adicionales a las ecuaciones dinámicas, que estén en concordancia con la práctica y teorización física. En particular, debe incluir, al menos, las leyes de conservación (en su forma diferencial o en su forma diferencial e integral).

Recordemos que la aplicación de la condición anterior (en conjunción con la *Cond1*) a la RG implica que obtengamos un conjunto de mundos físicamente posibles en los que solo se cumplen las leyes diferenciales de conservación y otro conjunto de mundos físicamente posibles, subconjunto del conjunto anterior, en que se cumplen tanto las leyes diferenciales como integrales de conservación.

Por último, en el Capítulo 7, se ha abordado el *Problema de la definición*. Se ha argumentado que, si bien Dowe es bastante explícito al afirmar que con su teoría busca establecer qué es la causación en el mundo, no es igualmente claro acerca de si las proposiciones centrales de su teoría (las proposiciones *CCI* y *CC2*) representan una definición de los términos involucrados y menos aún a qué tipo de definición corresponde. Todo parece indicar, de acuerdo con el discurso del propio autor, que estas son más bien caracterizaciones contingentes y refutables. Si bien este es un inconveniente que aparece en escena producto del análisis del problema de la presunta incompatibilidad de la TCC en la RG, afecta los propios fundamentos de la teoría con independencia de esta problemática. Hacia el final de ese capítulo hemos argumentado que los problemas internos que presenta la TCC pueden encontrar una solución si se asume que las proposiciones *CCI* y *CC2* representan definiciones reales de los conceptos de “proceso causal” e “interacción causal”. Sobre esta base argumentativa, por consiguiente, es un requisito deseable para la TCC y para cualquier teoría de la causación física, la imposición de la siguiente condición:

Cond3: Los conceptos nucleares de una teoría de la causación física como la TCC, en particular las proposiciones *CCI* y *CC2*, deben presentarse como definiciones reales y, por lo tanto, metafísicamente necesarias e irrefutables.

Después de este breve repaso por las principales ideas involucradas en los capítulos anteriores, volvamos ahora nuestra atención hacia el análisis del alcance y las limitaciones de la TCC de Dowe en la teoría de la gravitación de Einstein, tomando como referencia para este análisis las condiciones antes mencionadas. Para ello, esquematicemos en primer lugar las diferentes opciones que surgen de acuerdo con las condiciones sugeridas para cada problema.

(a) Respecto del *Problema de las leyes de conservación*:

(c1): Las leyes de conservación integrales no se cumplen en ningún modelo (o solución) de las ecuaciones de campo de Einstein.

(c2): Las leyes de conservación integrales se cumplen en algunos modelos de las ecuaciones de campo de Einstein.

(b) Respecto del *Problema de la posibilidad física*:

(b1): Lo físicamente posibles viene dado por las soluciones de las ecuaciones de campo de Einstein con independencia de si en ellas se cumplen o no las leyes de conservación.

(b2): Lo físicamente posible viene dado por las soluciones de las ecuaciones de campo de Einstein en las cuales se cumplen las leyes de conservación (en su forma diferencial o en su forma diferencial e integral).

(c) Respecto del *Problema de la definición*:

(a1): CC1 y CC2 representan una caracterización contingente (no una definición) de “proceso causal” e “interacción causal”, respectivamente.

(b1): CC1 y CC2 representan una definición real de “proceso causal” e “interacción causal”, respectivamente.

Dadas las opciones anteriores, existen ocho alternativas lógicamente posibles que ofrecen una panorámica general acerca del alcance y las limitaciones de la TCC de Dowe en el marco de los escenarios relativistas generales:

(i) (a1) – (b1) – (c1). De acuerdo con esta alternativa, los mundos físicamente posibles de la RG están dados, según (b1), por las soluciones a las ecuaciones de Einstein,

independientemente de si en estas soluciones es posible enunciar leyes de conservación bien definidas, ya sea únicamente diferenciales o diferenciales e integrales. Además, por la opción (a1), ninguno de estos mundos posibles cuenta con leyes de conservación integrales¹⁵¹. Ahora bien, consideremos que las proposiciones *CCI* y *CC2* podrían ser falsas, por (c1). En adición a lo anterior, debido a que no existen leyes de conservación integrales, tampoco contamos con cantidades físicas conservadas. Por lo tanto, la TCC no resulta aplicable a la RG, pues esta resultaría ser falsa en todos los mundos físicamente posibles de esta teoría física. En efecto, puesto que la ausencia de leyes de conservación integrales implica la ausencia de cantidades físicas que pueden ser calificadas como conservadas¹⁵², las nociones de “poseer” e “intercambiar” (las cuales, recordemos, forman parte de las proposiciones *CCI* y *CC2*) una cantidad conservada, no se encontrarían bien definidas en los contextos relativistas generales. Desde el punto de vista de la RG, la TCC pierde sentido. Este es el escenario en que ubican a la TCC sus críticos.

- (ii) (a1) – (b1) – (c2). Al igual que en el caso anterior, en este escenario los mundos físicamente posibles de la RG estarían dados por los modelos que son soluciones a las ecuaciones de Einstein y, además, en ninguno de ellos contamos con las leyes integrales de conservación. La diferencia con el caso (i) anterior, sin embargo, es que ahora las proposiciones *CCI* y *CC2* son consideradas definiciones, y no solo caracterizaciones contingentes, de los conceptos involucrados en ellas; por lo tanto, estas asumen el carácter de necesarias. Este último hecho marca una diferencia respecto de las consecuencias a las que arribamos en relación a la opción (i) anterior, pues debemos concluir que en ningún mundo físicamente posible de la RG hay causación. En efecto, debido a que no contamos con cantidades físicas conservadas,

¹⁵¹ Aun cuando las leyes de conservación integrales no existan, podrían existir las leyes de conservación en su forma diferencial (recuérdese Capítulo 5, Sección 5.5). De hecho, para el presente análisis asumiremos que las leyes de conservación diferenciales sí se encuentran correctamente definidas en el marco de la RG. Fijar esta variable no altera ni nubla las conclusiones a las que podamos arribar, sino que por el contrario, nos permitirá un análisis más claro de la situación general.

¹⁵² Se podría argumentar que aún contamos con leyes de conservación diferenciales con las cuales determinar qué cantidades físicas podemos considerar como conservadas. Sin embargo, recordemos que hemos asumido que son las leyes de conservación integrales las que poseen el significado físico adecuado en el marco de la TCC.

puesto que no existen las herramientas necesarias (las leyes integrales) para definir las, tampoco tendremos procesos e interacciones causales de acuerdo a la TCC, pues estos conceptos se fundamentan en la noción de conservación de una cantidad.

- (iii) (a1) – (b2) – (c1). En este caso, debido a que adoptamos la opción (b2), los mundos físicamente posibles de la RG se reducen al conjunto de modelos que poseen leyes de conservación diferenciales bien definidas, donde ninguno de ellos presenta leyes de conservación integrales. Además, por (c1), tanto *CCI* como *CC2* son caracterizaciones contingentes. En consecuencia, en este escenario debe concluirse, al igual que en la opción (i), que la TCC resulta ser falsa desde el punto de vista de la RG. La única diferencia entre la alternativa (i) y esta alternativa se encuentra en la noción de mundo físicamente posible. Mientras que en la alternativa (i) esta noción se encuentra determinada por los modelos de la RG que son soluciones a las ecuaciones de Einstein, en este caso la noción de mundo físicamente posibles se reduce a los modelos que cumplen con las leyes diferenciales de conservación.
- (iv) (a1) – (b2) – (c2). En esta opción, los mundos físicamente posibles se encuentran determinados por aquellos modelos de la RG que cuenten con genuinas leyes de conservación diferenciales, pero en ninguno de ellos contamos con leyes de conservación en su forma integral y, por lo tanto, tampoco tenemos cantidades físicas conservadas. Considerando lo anterior, y debido a que en este escenario las proposiciones *CCI* y *CC2* de Dowe son consideradas como definiciones, debemos concluir que no hay espacio para relatos causales, al estilo de la TCC, en el marco de la RG. Las conclusiones a las que hemos arribado aquí son las mismas que las presentadas en la alternativa (ii). Nuevamente, la única diferencia entre ambos escenarios se encuentra en nuestra comprensión del concepto de “mundo físicamente posible”.
- (v) (a2) – (b1) – (c1). En este caso tendríamos mundos físicamente posibles de la RG donde no hay leyes integrales de conservación y, por lo tanto, en ellos la caracterización dada por Dowe sería falsa. En consecuencia, la TCC solo tendría un rango de aplicación acotado a aquellos modelos de la RG que cumplan con las leyes

integrales. Esta es la posición que parece defender Dowe frente a las críticas a su teoría. Como hemos visto en el Capítulo 7, el autor hace afirmaciones contradictorias respecto de las conclusiones que debemos extraer de un escenario como este. Por un lado, el autor afirma que su teoría resultará ser falsa si, en efecto, no contáramos con leyes de conservación en la RG. Sin embargo, por otro lado, también afirma que la TCC simplemente no se pronuncia en aquellos mundos donde no haya este tipo de leyes. De todos modos, ambas alternativas son problemáticas. Por una parte, para saber si la TCC es falsa necesitamos contar con criterios definicionales para ello, por otra parte, si la teoría no se pronuncia en aquellos mundos donde no hay leyes de conservación integrales, la siguiente pregunta queda sin responder: ¿cómo se identifican los procesos e interacciones causales en los mundos físicamente posibles donde no se cumplen las leyes de conservación? ¿existen procesos e interacciones causales en tales mundos? Estas son preguntas que deberían ser contestadas por una teoría que aspira a explicar qué es la causación en el mundo, tal como pretende Dowe (recuérdese la discusión del Capítulo 7).

- (vi) (a2) – (b1) – (c2). Esta elección implica, por un lado, que hay mundos físicamente posibles, en los cuales se cuenta con leyes de conservación integrales, donde sí tenemos causación y, por otro lado, que hay mundos físicamente posibles donde no habría causación. En efecto, en este último caso, existirían mundos físicamente posibles donde no hay leyes integrales y, por lo tanto, las definiciones dadas por *CCI* y *CC2* no serían suficientes para determinar si en aquellos mundos habría causación o no.
- (vii) (a2) – (b2) – (c1). Las conclusiones a las que arribamos bajo esta elección son las mismas que hemos mencionado en la alternativa (v) más arriba. La TCC sería aplicable únicamente en aquellos mundos físicamente posibles de la RG que cumplen con las leyes integrales de conservación, pero sería falsa en aquellos mundos físicamente posibles en que únicamente contamos con las diferenciales. La única diferencia de esta alternativa con la alternativa (v) es que en este caso los mundos posibles de la RG se ven reducidos a aquellos modelos de la teoría en que se cumplen las leyes diferenciales, pero no necesariamente las integrales.

(viii) (a2) – (b2) – (c2). Esta elección es muy similar a la alternativa (vi) que hemos analizado más arriba. En efecto, en este caso debemos concluir que, de acuerdo con la TCC, habría causación en los mundos físicamente posibles donde existen leyes de conservación integrales bien definidas, pero no habría causación en aquellos mundos físicamente posibles donde solo contamos con las leyes diferenciales de conservación. Al igual de lo que sucede con (v) y (vii), esta alternativa y la alternativa (vi) solo difieren en la consideración de la noción de mundo físicamente posible.

En la siguiente figura se esquematizan las alternativas descritas en los puntos anteriores¹⁵³.

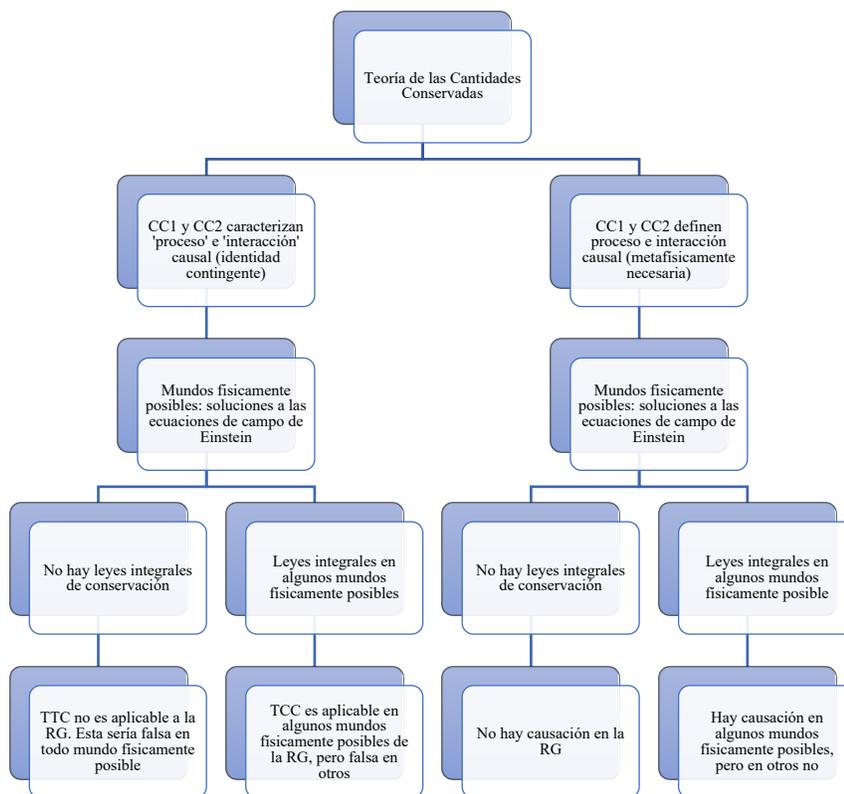


Figura 8.1: Representación esquemática de las alternativas posibles de acuerdo a los problemas abordados.

¹⁵³ De acuerdo con las alternativas analizadas anteriormente, tenemos como resultado cuatro escenarios posibles para la TCC. Los otros cuatro escenarios solo se diferencian en la noción de “mundo físicamente posible” asumida.

De acuerdo con el análisis anterior, hay básicamente cuatro escenarios posibles para la TCC en el marco de la RG (recordemos que los otros cuatro escenarios solo se diferencian en la noción de “posibilidad física” que es asumida). En un primer escenario la teoría de la causación de Dowe resulta no ser aplicable en los “universos” relativistas generales debido a que esta queda refutada por la RG. Como mencionamos más arriba, este parece ser el escenario en el que ubican a la TCC sus críticos. Una segunda alternativa es aquella en que la TCC es vista como una teoría metafísica que establece definiciones necesarias respecto de las nociones de proceso e interacción causal. En este caso, la TCC resulta en efecto aplicable a la RG, pero debemos concluir, sin embargo, que no hay causación en el marco de la teoría de la gravitación de Einstein. Un tercer escenario posible es aquel en el cual la TCC resulta aplicable a una cierta clase de mundos físicamente posibles de la RG (aquellos con leyes de conservación integrales), a la vez que resulta ser falsa en otra clase de mundos físicamente posibles. Esta es el escenario en que parece posicionarse Dowe en su defensa a las críticas formuladas a su teoría. Por último, un cuarto escenario posible es aquel en que tenemos causación en algunos mundos posible de la RG, pero no en otros.

Una primera conclusión que podemos extraer en relación a los escenarios antes mencionados es la siguiente. Los escenarios en los cuales la TCC se plantea como una teoría metafísica en toda regla, pero guiada por los resultados que nos brindan nuestras mejores teorías físicas, resultan ser más coherentes y menos problemáticos que aquellos escenarios en los que la TCC es vista como una caracterización contingente. En efecto, en este último caso debemos dar cuenta de los problemas que se han mencionado en las alternativas (v) y (vii), pero estos problemas no aparecen si la TCC se presenta como una teoría metafísica de la causación. Esta es la defensa que hemos presentado en el Capítulo 7 en cuanto a la propuesta de que las proposiciones *CCI* y *CC2* sean vistas como definiciones reales y no como meras caracterizaciones de los conceptos involucrados en ellas. En ambos casos solo podemos hablar de causación en una cierta clase acotada de mundos posibles de la RG, pero si la TCC se plantea como una definición necesaria, esta gana en coherencia debido a que se disuelven los problemas que se presentan si la TCC fuera una teoría contingente. Esto último es un hecho deseable desde el punto de vista de una defensa de la

TCC, pero que obliga a replantearse la visión que podamos tener acerca de la relación entre física y metafísica.

Como afirmamos en el párrafo anterior, cuando presentemos la TCC como resultado de un análisis metafísico, su aplicación en los escenarios relativistas generales no resulta problemática (al menos, no presentan los problemas mencionados en (v) más arriba). Ahora bien, una cuestión distinta es la conclusión a la que arribamos como resultado de la aplicación de la TCC en la RG: solo tendríamos causación en un rango muy acotado de modelos de la RG. En efecto, solo sería posible hablar de causación en los mundos posibles o modelos de la RG en que se cumplen las leyes de conservación integrales, es decir, en aquellos modelos que presentan ciertas simetrías del espacio-tiempo (caracterizados por la presencia de un campo vectorial de Killing), como son los espacio-tiempos asintóticamente planos. Esto último trae como consecuencia que en los mundos físicamente posibles o modelos más significativos en el contexto de la RG, como lo son aquellos que presentan un espacio-tiempo curvo, no haya espacio para los relatos de tipo causal.

Ahora bien, ¿cuáles son las perspectivas para una teoría de la causación fundamentada en cantidades físicas conservadas a la luz de nuestra mejor teoría física del espacio-tiempo? Creemos que hay tres estrategias que se pueden continuar con el objetivo de reevaluar la compatibilidad de la TCC en la teoría de la gravitación de Einstein. Una primera estrategia consistiría en argumentar que el contenido de las proposiciones *CC1* y *CC2* puede tener sentido con la sola existencia de leyes de conservación diferenciales, sin la necesidad de contar con sus homólogas integrales. Recordemos que las expresiones de conservación diferenciales se cumplen en un conjunto más amplio de modelos de la RG, que incluye aquellos que presentan un espacio-tiempo curvo. Por lo tanto, este movimiento permitiría la presencia de relatos o explicación causales en un rango más amplio de modelos de la RG. De hecho, si nos posicionamos en las alternativas (iv) u (viii) presentadas más arriba, podríamos concluir que hay causación en todo mundo físicamente posible de la RG. Sin embargo, este movimiento implica tres desafíos no menores. En primer lugar, implica una evaluación o sustitución de ciertas nociones utilizadas en el lenguaje causal de la TCC, tales como “posesión”, “intercambio”, “transferencia”, “propagación”, etc.; de modo que estas resulten compatibles con las expresiones diferenciales de conservación (para un análisis de los conceptos antes mencionados, ver

Curiel, 2000, Sección 4). En segundo lugar, implica argumentar en favor de una defensa de las leyes de conservación como poseedoras de contenido físico (ver De Haro, 2021, p. 4). Por último, esta estrategia también implica dar cuenta de los problemas que presentan las leyes de conservación diferenciales en el marco de la RG (recuérdese el Capítulo 5).

Una segunda estrategia posible para avanzar hacia una teoría de la causación compatible con los escenarios relativistas generales, y que queda pendiente en esta investigación, es apelar a los avances en las investigaciones sobre fundamentos de la RG. Recordemos que en el Capítulo 5 hemos presentado un análisis acerca de la validez de las expresiones de conservación integrales siguiendo el enfoque clásico pseudo-tensorial. Sin embargo, en ese mismo capítulo, también hemos mencionado que en las últimas décadas han proliferado variadas propuestas, que caen bajo la etiqueta de “enfoques cuasi-locales”, los cuales prometen encontrar expresiones integrales de conservación que tendrían la característica de ser menos problemáticas que sus análogas en el enfoque pseudo-tensorial. Para una defensa de la teoría de la causación de Dowe, podemos encontrar aquí una luz de esperanza.

Finalmente, una última estrategia, más radical que las anteriores, es asumir que la RG es una teoría incompleta (por ejemplo, por no considerar los efectos cuánticos) y que, por lo tanto, no debería ser problemático que en ella haya un espacio acotado, o que simplemente no haya lugar, para la causación. Sin embargo, este movimiento argumentativo contiene una suposición implícita: que hay un lugar para la causación, en el sentido de la TCC, en las teorías que pretenden sustituir a la RG y a la mecánica cuántica, tales como la teoría de gravedad cuántica. Lo cual, por supuesto, habría que probar.

Referencias

- Armstrong, D. (1978). *A Theory of Universals*. Cambridge: Cambridge University Press.
- Armstrong, D. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Armstrong, D. (1980). "Identity Through Time". En P. van Inwagen (Ed), *Time and Cause: Essays Presented to Richard Taylor* (pp. 67–78). Dordrecht: Reidel Publishing Company.
- Armstrong, D. (1983). *What Is a Law of Nature?* Cambridge: Cambridge University Press.
- Aronson, J. (1971a). "On the Grammar of 'Cause'". *Synthese*, 22(3-4), 414–430.
- Aronson, J. (1971b). "The Legacy of Hume's Analysis of Causation". *Studies in History and Philosophy of Science*, 2(2), 135–156.
- Baker, D. (2005). "Spacetime Substantivalism and Einstein's Cosmological Constant". *Philosophy of Science*, 72(5), 1299–1311.
- Bartels, A., y Wohlfarth, D. (2014). "How Fundamental Physics Represents Causality". En M. C. Galavotti, D. Dieks, W. Gonzalez, S. Hartmann, T. Uebel, y M. Weber (Eds.), *New Directions in the Philosophy of Science* (pp. 485–500). Dordrecht: Springer.
- Baumgartner, M. (2008). "Regularity Theories Reassessed". *Philosophia*, 36(3), 327–354.
- Belot, G. (2011). *Geometric Possibility*. Oxford: Oxford University Press.
- Ben-Menahem, Y. (2018). *Causation in Science*. Princeton: Princeton University Press.
- Bigelow, J., Ellis, B. y Pargetter, R. (1988). "Forces". *Philosophy of Science*, 55(4), 614–630.
- Bigelow, J. y Pargetter, R. (1990). *Science and Necessity*. Cambridge: Cambridge University Press.
- Bird, A. (2010). "Causation and Manifestation of Powers". En A. Marmodoro (Ed.), *The Metaphysics of Powers. Their Grounding and Their Manifestations* (pp. 160–168). New York: Routledge.

- Brading, K. (2002). “Which symmetry? Noether, Weyl, and conservation of electric charge”. *Studies in History and Philosophy of Modern Physics*, 33(1), 3–22.
- Brading, K. y Castellani, E. (2003). *Symmetries in Physics. Philosophical Reflections*. Cambridge: Cambridge University Press.
- Brading K. (2005). “A Note on General Relativity, Energy Conservation, and Noether’s Theorems”. En A. Kox y J. Eisenstaedt (Eds.), *The Universe of General Relativity. Einstein Studies (Volume 11)* (pp. 125–136). Boston: Birkhäuser.
- Brading, K. y Castellani, E. (2006). “Symmetries and Invariances in Classical Physics”. En J. Butterfield y J. Earman (Eds.), *Handbook of the Philosophy of Physics* (pp. 1331–1368). Amsterdam: North–Holland.
- Bradley, R. y Swartz, N. (1979). *Possible Worlds: An Introduction to Logic and its Philosophy*. Indianapolis: Hackett Publishing Company.
- Brown, H, y Holland, P. (2004). “Dynamical versus variational symmetries: understanding Noether's first theorem”. *Molecular Physics*, 102(11/12), 1133–1139.
- Brzezinski, J. y Nowak, L. (1992). *Idealization III: Approximation and Truth*. Amsterdam: Rodopi.
- Bueno, O. y Vickers, P. (Eds.). (2014). *Is science inconsistent? (Special Issue)*. *Synthese*, 191(13).
- Bunge, M. (2008). *Causality and Modern Science*. New Brunswick: Transaction Publishers.
- Carroll, S. (2004). *Spacetime and Geometry: An Introduction to General Relativity*. San Francisco: Addison Wesley.
- Carroll, J. (2020). “Laws of Nature”. En E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Edición invierno 2020)*. Recuperado de <https://plato.stanford.edu/archives/win2020/entries/laws-of-nature/>
- Cartwright, N. (1979). “Causal Laws and Effective Strategies”. *Noûs*, 13(4), 419–437.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. New York: Oxford University Press.

- Chakravartty, A. (2001). "The Semantic or Model-Theoretic View of Theories and Scientific Realism". *Synthese*, 127(3), 325–345.
- Chalmers, D. y Jackson, F. (2001). "Conceptual Analysis and Reductive Explanation". *Philosophical Review*, 110(3), 315–61.
- Choi, S. (2003). "The Conserved Quantity Theory of Causation and Closed Systems". *Philosophy of Science*, 70(3), 510–530.
- Cohen, J. y Callender, C. (2009). "A better best system account of lawhood". *Philosophical Studies*, 145(1), 1–34.
- Collingwood, R. (1940). *An Essay on Metaphysics*. Oxford: Clarendon Press.
- Craig, E. (1987). *The Mind of God and the Works of Man*. New York: Oxford University Press.
- Curiel, E. (2000). "The Constraints General Relativity Places on Physicalist Accounts of Causality". *Theoria*, 15(1), 33–58.
- Da Costa, N. y French, S. (2003). *Science and Partial Truth. A Unitary Approach to Models and Scientific Reasoning*. New York: Oxford University Press.
- De Haro, S. (2021). "Noether's Theorems and Energy in General Relativity". Recuperado de <https://arxiv.org/abs/2103.17160v1>
- Dieks, D. (1986). "Physics and the Direction of Causation". *Erkenntnis*, 25(1), 85–110.
- Dowe, P. (1992a). "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory". *Philosophy of Science*, 59(2), 195–216.
- Dowe, P. (1992b). "Process Causality and Asymmetry". *Erkenntnis*, 37(2), 179–196.
- Dowe, P. (1992c). "An Empiricist Defence of the Causal Account of Explanation". *International Studies in the Philosophy of Science*, 6(2), 123–128.
- Dowe, P. (1995a). "Causality and Conserved Quantities: A Reply to Salmon". *Philosophy of Science*, 62(2), 321–333.
- Dowe, P. (1995b). "What's Right and What's Wrong with Transference Theories". *Erkenntnis*, 42(3), 363–374.

- Dowe, P. (1999). “Good Connections: Causation and Causal Processes”. En H. Sankey (Ed.), *Causation and Laws of Nature* (pp. 247–264). Dordrecht: Kluwer Academic Publishers.
- Dowe, P. (2000). *Physical Causation*. Cambridge: Cambridge University Press.
- Dowe, P. (2000a). “The Conserved Quantity Theory Defended”. *Theoria*, 15(1), 11–31.
- Dowe, P. (2001). “A Counterfactual Theory of Prevention and ‘Causation’ by Omission”. *Australasian Journal of Philosophy*, 79(2), 216–226.
- Dowe, P. (2009). “Causal Process Theories”. En H. Beebe, C. Hitchcock y P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 213–233). New York: Oxford University Press.
- Dowe, P. (2014). “Localismo y Explicación Causal”. En H. Miguel (Ed.), *Causación, Explicación y Contrafácticos* (pp. 41–60). Buenos Aires: Prometeo Libros.
- Ducasse, C. (1976). “Causality: Critique of Hume’s Analysis”. En M. Brand (Ed.), *The Nature of Causation* (pp. 65–76). Urbana: University of Illinois Press.
- Dupré, J. (1984). “Probabilistic Causality Emancipated”. En P. French, T. Uehling y H. Wettstein (Eds.), *Midwest Studies in Philosophy IX* (pp. 169–175). Minneapolis: University of Minnesota Press.
- Dürr, P. (2019a). “Fantastic beasts and where (not) to find them: local gravitational energy and energy conservation in general relativity”. *Studies in History and Philosophy of Modern Physics*, 65, 1–14.
- Dürr, P. (2019b). “Against ‘Functional Gravitational Energy’: A Critical Note on Functionalism, Selective Realism, and Geometric Objects and Gravitational Energy”. *Synthese*. <https://doi.org/10.1007/s11229-019-02503-3>.
- Earman, J. (1986). *A Primer on Determinism*. Dordrecht: Reidel Publishing Company.
- Eells, E. (1988). “Probabilistic Causal Level”. En B. Skyrms y W. Harper (Eds.), *Causation, Chance and Credence (Volume 1)* (pp. 109–134). Dordrecht: Kluwer Academic Publishers.
- Eells, E. (1991). *Probabilistic Causality*. Cambridge: Cambridge University Press.

- Einstein, A. (1907). “On the Relativity Principle and the Conclusions Drawn for It”. En J. Stachel et.al. (Eds.), *The Collected Papers of Albert Einstein (Volume 2). The Swiss Years: Writings, 1900 – 1909* (pp. 252–311). Princeton: Princeton University Press.
- Einstein, A. y Rosen, N. (1935). “The Particle Problem in the General Theory of Relativity”. *Physical Review*, 48, 73–77.
- Ellis, B. (2001). *Scientific Essentialism*. Cambridge: Cambridge University Press.
- Fazekas, P., Gyenis, B., Hofer-Szabó, G. y Kertész, G. (2021). “A dynamical systems approach to causation”. *Synthese*, 198(7), 6065–6087.
- Fine, K. (1994). “Essence and Modality”. *Philosophical Perspectives*, 8, 1–16.
- Fleisch, D. (2012). *A Student’s Guide to Vectors and Tensors*. Cambridge: Cambridge University Press.
- Frigg, R. y Hartmann, S. (2020). “Models in Science”. *The Stanford Encyclopedia of Philosophy (Edición primavera 2020)*, E. Zalta (Ed.). Recuperado de <https://plato.stanford.edu/archives/spr2020/entries/models-science/>
- Frisch, M. (2012). “No place for causes? Causal skepticism in physics”. *European Journal for Philosophy of Science*, 2(3), 313–336.
- Frisch, M. (2014). *Causal Reasoning in Physics*. Cambridge: Cambridge University Press.
- García-Encinas, M. (2004). “Transference, or Identity Theories of Causation?”. *Theoria*, 19(1), 31–47.
- Gasking, D. (1955). “Causation and Recipes”. *Mind*, 64(256), 479–487.
- Glennan, S. (1996). “Mechanisms and the Nature of Causation”. *Erkenntnis*, 44(1), 49–71.
- Good, I. (1959). “A Theory of Causality”. *The British Journal for the Philosophy of Science*, 9(36), 307–310.
- Good, I. (1961a). “A Causal Calculus (I)”. *The British Journal for the Philosophy of Science*, 11(44), 305–318.
- Good, I. (1961b). “A Causal Calculus (II)”. *The British Journal for the Philosophy of Science*, 12(45), 43–51.

- Grabhoff, G. and May, M. (2001). “Causal Regularities”. En W. Spohn, M. Ledwig y M. Esfeld (Eds.), *Current Issues in Causation* (pp. 85–114). Paderborn: Mentis.
- Gupta, A. (2019). “Definitions”. En E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Edición invierno 2019). Recuperado de <https://plato.stanford.edu/entries/definitions/>
- Hall, N. (2000). “Causation and the Price of Transitivity”. *The Journal of Philosophy*, 97(4), 198–222.
- Hall, R. (1985 [1954]). *La Revolución Científica. 1500–1750*. Barcelona: Editorial Crítica.
- Hausman, D. (1998). *Causal Asymmetries*. Cambridge: Cambridge University Press.
- Hawking, S. y Ellis, G. (1973). *The Large-Scale Structure of Space-Time*. Cambridge: Cambridge University Press.
- Hawking, S. (2011 [1988]). *Historia del Tiempo. Del Big-Bang a los Agujeros Negros*. Madrid: Alianza Editorial.
- Herrera, M. (2017). “Teoría de las Cantidades Conservadas: una tensión interna”. *Revista Colombiana de Filosofía de la Ciencia*, 17(35), 91–117.
- Herrera, M. y López, C. (2020). “¿Qué hace físicamente posible a un mundo posible?” *Principia: An International Journal of Epistemology*, 24(1), 65–88.
- Hitchcock, C. (1995). “Salmon on Explanatory Relevance”. *Philosophy of Science*, 62(2), 304–320.
- Hitchcock, C. (2018). “Probabilistic Causation”. *The Stanford Encyclopedia of Philosophy* (Edición primavera 2021), E. Zalta (Ed.). Recuperado de <https://plato.stanford.edu/archives/spr2021/entries/causation-probabilistic/>
- Hofer, C. (2000). “Energy Conservation in GTR”. *Studies in History and Philosophy of Modern Physics*, 31(2), 87–199.
- Hofer, C. (2009). “Causation in Spacetime Theories”. En H. Beebe, C. Hitchcock y P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 687–706). New York: Oxford University Press.

- Hubert, M., y Poellinger, R. (2014). “Anchoring Causal Connections in Physical Concepts”. En M. C. Galavotti, D. Dieks, W. Gonzalez, S. Hartmann, T. Uebel, y M. Weber (Eds.), *New Directions in the Philosophy of Science* (pp. 501–509). Dordrecht: Springer.
- Hume, D. (1978 [1739]). *A Treatise of Human Nature*. L. A. Selby-Bigge y P. H. Nidditch (Eds.). Oxford: Clarendon Press.
- Hume, D. (2007 [1748]). *An Enquiry Concerning Human Understanding*. P. Millican (Ed.). Oxford: Oxford University Press.
- Humberstone, L. (2015). *Philosophical Applications of Modal Logic*. London: College Publications.
- Hunt, I. (2005). “Omissions and Preventions as Cases of Genuine Causation”. *Philosophical Papers*, 34(2), 209–223.
- Jackson, F. (1994). “Armchair Metaphysics”. En M. Michael y J. O’Leary – Hawthorne (Eds.), *Philosophy of Mind* (pp. 23–42). Dordrecht: Kluwer Academic Publishers.
- Janssen, B. (2013). *Teoría de la Relatividad General*. Granada: Universidad de Granada.
- Jubien, M. (1996). “Actualism and Iterated Modalities”. *Philosophical Studies*, 84(2/3), 109–125.
- Jubien, M. (2009). *Possibility*. Oxford: Oxford University Press.
- Kistler, M. (2006). *Causation and Laws of Nature*. New York: Routledge.
- Kitcher, P. (1985). “Two Approaches to Explanation”. *The Journal of Philosophy*, 82(11), 632–639.
- Kitcher, P. (1993). *The Advancement of Science*. New York: Oxford University Press.
- Kment, B. (2017). “Varieties of Modality”. En E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Edición primavera 2017). Recuperado de <https://plato.stanford.edu/archives/spr2017/entries/modality-varieties/>
- Koyré, A. (1965). *Newtonian Studies*. London: Chapman and Hall.

- Kripke, S. (1959). "A Completeness Theorem in Modal Logic". *Journal of Symbolic Logic*, 24(1), 1–15.
- Kripke, S. (1963). "Semantical Considerations on Modal Logic". *Acta Philosophica Fennica*, 16, 83–94.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press.
- Lam, V. (2005). "Causation and Space-time". *History and Philosophy of the Life Science*, 27(3/4), 465–478.
- Lam, V. (2010). "Metaphysics of Causation and Physics of General Relativity". *Humana.Mente*, 4(13), 61–80.
- Lam, V. (2011). "Gravitational and Nongravitational Energy: The Need for Background Structures". *Philosophy of Science*, 78(5), 1012–1023.
- Lange, M. (2007). "Laws and meta-laws of nature: Conservation laws and symmetries". *Studies in History and Philosophy of Modern Physics*, 38(3), 457–481.
- Lange, M. (2009). *Laws and lawmakers*. Oxford: Oxford University Press.
- Lange, M. (2011). "Conservation Laws in Scientific Explanations: Constraints or Coincidences?". *Philosophy of Science*, 78(3), 333–352.
- Lewis, D. (1973a). *Counterfactuals*. Oxford: Blackwell Publishers.
- Lewis, D. (1973b). "Causation". *The Journal of Philosophy*, 70(17), 556–567.
- Lewis, D. (1986a). *Philosophical Papers Volume II*. New York: Oxford University Press.
- Lewis, D. (1986b). *On the Plurality of Worlds*. Oxford: Blackwell.
- Lewis, D. (2000). "Causation as Influence". *The Journal of Philosophy*, 97(4), 182 – 197.
- Loewer, B. (1996). "Humean Supervenience". *Philosophical Topics*, 24(1), 101–126.
- Lupher, T. (2009). "A Physical Critique of Physical Causation". *Synthese*, 167(1), 67–80.
- Mackie, J. (1980 [1974]). *The Cement of the Universe*. Oxford: Clarendon Press.

- Malament, D. (2012). “A remark about the ‘geodesic principle’ in general relativity. En M. Frapier, D. Brown y R. DiSalle (Eds.), *Analysis and interpretation in the exact sciences: Essays in honor of William Demopoulos* (pp. 245–252). Dordrecht: Springer.
- Maudlin, T. (2007). *The Metaphysics Within Physics*. New York: Oxford University Press.
- Maxwell, J. (1865). “VIII. A dynamical theory of the electromagnetic field”. *Philosophical Transactions of the Royal Society of London*, 155, 459–512.
<https://doi.org/10.1098/rstl.1865.0008>
- McMullin, E. (1985). “Galilean Idealization”. *Studies in History and Philosophy of Science*, 16(3), 247–273.
- Mellor, D. (1988). “On Raising the Chances of Effects”. En J. Fetzer (Ed.), *Probability and Causality. Essay in Honor of Wesley C. Salmon* (pp. 229–239). Dordrecht: D. Reidel Publishing Company.
- Menzies, P. y Price, H. (1993). “Causation as a Secondary Quality”. *The British Journal for the Philosophy of Science*, 44(2), 187–203.
- Miguel, H. y Paruelo, J. (2004). “Causación: Análisis Empírico vs. Análisis Conceptual”. *Epistemología e Historia de la Ciencia*, 10(10), 401–407.
- Miguel, H. y Paruelo, J. (2005). “Superposición de Interacciones Causales en la Teoría de Phil Dowe”. *Enrahonar*, 37, 123–139.
- Miguel, H. (2006). *Teorías Contrafácticas de la Causación y Teorías Físicas de la Causación (tesis doctoral)*. Universidad Nacional de la Plata, La Plata, Buenos Aires.
- Miguel, H. y Paruelo, J. (2007). “Causar o Dejar que Occurra”. *Andamios*, 4(7), 7–18.
- Mill, J. (2011 [1843]). *A System of Logic: Ratiocinative and Inductive*. Cambridge: Cambridge University Press.
- Misner, C., Thorne, K. y Wheeler, J. (1973). *Gravitation*. San Francisco: Freeman and Company.

- Morganti, M. y Tahko, T. (2017). “Moderately Naturalistic Metaphysics”. *Synthese*, 194(7), 2557–2580.
- Morris, M., Thorne, K. y Yurtsever, U. (1988). “Wormholes, Time Machines, and the Weak Energy Condition”. *Physical Review Letters*, 61(13), 1446–1449.
- Mumford, S. (2004). *Laws in Nature*. London: Routledge.
- Mumford, S. y Anjum, R. (2011). *Getting Causes form Powers*. New York: Oxford University Press.
- Neapolitan, R. (1990). *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. New York: John Wiley and Sons.
- Newton, I., Cohen, B., Whitman, A. (1999 [1687]). *The Principia: Mathematical Principles of Natural Philosophy*. Berkeley, California: University of California Press.
- Ney, A. (2009). “Physical Causation and Difference-Making”. *The British Journal for the Philosophy of Science*, 60(4), 737–764.
- Niiniluoto, I. (2007). “Idealization, Counterfactuals, and Truthlikeness”. En J. Brzezinski et.al. (Eds.), *The Courage of Doing Philosophy: Essays Dedicated to Leszek Nowak* (pp. 103–122). Amsterdam: Rodopi.
- Noether, E. (1971[1918]). “Invariant Variation Problems” (Trad. M. Tavel). *Transport Theory and Statistical Physics*, 1(3), 183–207.
- Norton, J. (2007). “Causation as Folk Science”. En P. Huw y R. Corry (Eds.), *Causality, Physics and the Constitution of Reality: Russell’s Republic Revisited* (pp. 11–44). Oxford: Oxford University Press.
- Norton, J. (2012). “Approximation and Idealization. Why the Difference Matters”. *Philosophy of Science*, 79(2), 207–232.
- Norton, J. (2015). “What Can We Learn about the Ontology of Space and Time from the Theory of Relativity?” En L. Sklar (Ed.), *Physical Theory. Method and Interpretation* (pp. 185–228). Oxford: Oxford University Press.

- Papineau, D. (1989). "Pure, Mixed and Spurious Probabilities and Their Significance for a Reductionist Theory of Causation". En P. Kitcher y W. Salmon (Eds.), *Scientific Explanation* (pp. 307–348). Minneapolis: University of Minnesota Press.
- Paul, L. y Hall, N. (2013). *Causation: A User's Guide*. Oxford: Oxford University Press.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann Publishers.
- Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
- Pitts, B. (2010). "Gauge-Invariant Localization of Infinitely Many Gravitational Energies from All Possible Auxiliary Structures". *General Relativity and Gravitation*, 42(3), 601–622.
- Plantinga, A. (1974). *The Nature of Necessity*. Oxford: Oxford University Press.
- Price, H. (1992). "Agency and Causal Asymmetry". *Mind*, 101(403), 501–520.
- Price, H. y Corry, R. (Eds.). (2007). *Causality, Physics and the Constitution of Reality: Russell's Republic Revisited*. Oxford: Oxford University Press.
- Priest, G. (2008). *An Introduction to Non-classical Logic*. Cambridge: Cambridge University Press.
- Psillos, S. (1996). "Scientific Realism and the 'Pessimistic Induction'". *Philosophy of Science*, 63, S306–S314.
- Psillos, S. (2002). *Causation and Explanation*. Stocksfield: Acumen.
- Psillos, S. (2009). "Regularity Theories". En H. Beebe, C. Hitchcock y P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 131–157). New York: Oxford University Press.
- Psillos, S. (2018). "Realism and Theory Change in Science". *The Stanford Encyclopedia of Philosophy* (Edición verano 2018), E. Zalta (Ed.). Recuperado de <https://plato.stanford.edu/archives/sum2018/entries/realism-theory-change/>.

- Putnam, H. (1975). "The Meaning of Meaning". *Minnesota Studies in the Philosophy of Science*, 7, 131 - 193.
- Quezada, W. (2002). "Causalidad Física: Procesos Causales y Cantidades Conservadas". *Revista de Filosofía, Universidad de Chile*, 58, 79–100.
- Quine, W. (1966). *The Ways of Paradox*. New York: Random House.
- Read, J. (2020). "Functional Gravitational Energy". *British Journal for the Philosophy of Science*, 71(1), 205–232.
- Reichenbach, H. (1959). "The Principle of Causality and the Possibility of Its Empirical Confirmation". En H. Reichenbach, *Modern Philosophy of Science* (pp. 109–134). London: Routledge and Kegan Paul.
- Reichenbach, H. (1971 [1956]). *The Direction of Time*. Berkeley: University of California Press.
- Romero–Maltrana, D. (2015). "Symmetries as by–products of conserved quantities". *Studies in History and Philosophy of Modern Physics*, 52, 358–368.
- Rueger, A. (1998). "Local Theories of Causation and the A Posteriori Identification of the Causal Relation". *Erkenntnis*, 48(1), 25–38.
- Ruetsche, L. (2013). *Interpreting Quantum Theories*. Oxford: Oxford University Press.
- Russell, B. (1912/1913). "On the Notion of Cause". *Proceedings of the Aristotelian Society*, 13(1), 1–26.
- Russell, B. (1948). *Human Knowledge. Its Scope and Limits*. London: George Allen and Unwin.
- Saatsi, J. (2016). "Models, Idealisations, and Realism". En E. Ippoliti, F. Sterpetti y T. Nickles (Eds.), *Models and Inferences in Science* (pp. 173–190). Springer.
- Salmon, W. (1977). "An 'At-At' Theory of Causal Influence". *Philosophy of Science*, 44(2), 215–224.
- Salmon, W. (1980). "Probabilistic Causality". *Pacific Philosophical Quarterly*, 61(1-2), 50–74.

- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Salmon, W. (1994). "Causality Without Counterfactuals". *Philosophy of Science*, 61(2), 297–312.
- Salmon, W. (1997). "Causality and Explanation: A Reply Two Critiques". *Philosophy of Science*, 64(3), 46–477.
- Salmon, W. (1998). *Causality and Explanation*. New York: Oxford University Press.
- Schaffer, J. (2000a). "Trumping Preemption". *The Journal of Philosophy*, 97(4), 165–181.
- Schaffer, J. (2000b). "Causation by Disconnection". *Philosophy of Science*, 67(2), 285–300.
- Schaffer, J. (2016). "The Metaphysics of Causation". En E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Edición otoño 2016)*. Recuperado de <https://plato.stanford.edu/entries/causation-metaphysics/>
- Schrenk, M. (2016). *Metaphysics of Science. A Systemic and Historical Introduction*. New York: Routledge.
- Schutz, B. (2009). *A First Course in General Relativity*. Cambridge: Cambridge University Press.
- Skyrms, B. (1980). *Causal Necessity. A Pragmatic Investigation of the Necessity of Laws*. New Haven: Yale University Press.
- Soames, S. (2005). *Reference and Description: The Case against Two-Dimensionalism*. Princeton: Princeton University Press.
- Soames, S. (2011). "Kripke on Epistemic and Metaphysical Possibility: Two Routes to the Necessary A posteriori". En A. Berger (Ed.), *Saul Kripke* (pp. 78–99). New York: Cambridge University Press.
- Sober, E. (1987). "Review Article: Explanation and Causation". *The British Journal for the Philosophy of Science*, 38(2), 243–257.

- Spirtes, P., Glymour, C. y Scheines, R. (1993). *Causation, Prediction, and Search*. New York: Springer-Verlag.
- Stanford, P. (2006). *Exceeding our Grasp*. Oxford: Oxford University Press.
- Strawson, G. (2014). *The Secret Connection. Causation, Realism, and David Hume*. New York: Oxford University Press.
- Strevens, M. (2007). “Review of Woodward: ‘Making things happen’”. *Philosophy and Phenomenological Research*, 74(1), 233–249.
- Suppe, F. (1989). *The Semantic Conception of Theories and Scientific Realism*. Chicago: University of Illinois Press.
- Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam: North Holland.
- Szabados, L. (2009). “Quasi-Local Energy-Momentum and Angular Momentum in General Relativity”. *Living Reviews in Relativity*, 12(4), 1–163.
- Torii, T. y Shinkai, H. (2013). “Wormholes in Higher Dimensional Space-Time: Exact Solutions and Their Linear Stability Analysis”. *Physical Review D*, 88(6), 064027.
- Torretti, R. (1983). *Relativity and Geometry*. Oxford: Pergamon Press.
- Torretti, R. (1994). *La Geometría del Universo y Otros Ensayos de Filosofía Natural*. Mérida: Universidad de los Andes.
- Uffink, J. (2001). “Bluff Your Way in the Second Law of Thermodynamics”. *Studies in History and Philosophy of Modern Physics*, 32(3), 305–394.
- Vaidya, A. (2007). “The Epistemology of Modality”. En E. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Edición invierno 2016). Recuperado de: <http://plato.stanford.edu/entries/modality-epistemology/>
- Van Inwagen, P. (1986). “Two Concepts of Possible Worlds”. *Midwest Studies in Philosophy*, 11(1), 185–213.
- Vicente, A. (2002). “The Localism of the Conserved Quantity Theory”. *Theoria*, 17(3), 563–571.

- Vickers, P. (2017). “Understanding the Selective Realist Defence against the PMI”. *Synthese*, 194(9), 3221–3232.
- von Wright, G. (1971). *Explanation and Understanding*. New York: Cornell University Press.
- Wald, R. (1984). *General Relativity*. Chicago: University of Chicago Press.
- Weatherall, J. (2018). “Geometry and Motion in General Relativity”. Recuperado de <http://arxiv.org/abs/1810.09046>.
- Weisberg, M. (2007). “Three Kinds of Idealization”. *Journal of Philosophy*, 104(12), 639–659.
- Whitehead, A. N. (1922). *The Principle of Relativity With Applications to Physical Science*. Cambridge: Cambridge University Press.
- Williamson, J. (2005). *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford: Oxford University Press.
- Williamson, J. (2009). “Probabilistic Theories”. En H. Beebe, C. Hitchcock y P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 185–214). New York: Oxford University Press.
- Wilson, A. (2021). “Counterpossible Reasoning in Physics”. *Philosophy of Science*, 88(5). <https://doi.org/10.1086/714706>
- Wolff, J. (2013). “Are Conservation Laws Metaphysically Necessary?”. *Philosophy of Science*, 80(5), 898–906.
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.
- Woodward, J. y Hitchcock, C. (2003). “Explanatory Generalizations, Part I: A Counterfactual Account”. *Noûs*, 37(1), 1–24.
- Woodward, J. (2009). “Agency and Interventionist Theories”. En H. Beebe, C. Hitchcock y P. Menzies (Eds.), *The Oxford Handbook of Causation* (pp. 234–262). New York: Oxford University Press.

Woodward, J. (2016). "Causation and Manipulability". *The Stanford Encyclopedia of Philosophy* (Edición invierno 2016), E. Zalta (Ed.). Recuperado de <https://plato.stanford.edu/archives/win2016/entries/causation-mani/>

Wright, J. (1983). *The Sceptical Realism of David Hume*. Manchester: Manchester University Press.