

Los orígenes de la Teoría de Distribuciones

Fernando Bombal
Seminario de Historia de la Matemática I
Universidad Complutense, Madrid. 1991. págs. 211-244

Introducción

Las ecuaciones fundamentales de la Física matemática vienen dadas por ecuaciones diferenciales y, por tanto, son aplicables solamente (al menos en principio) a fenómenos en los que las variables físicas (como el desplazamiento elástico o la intensidad eléctrica) sean funciones muy regulares del espacio y el tiempo. Desde el comienzo mismo de la aplicación del cálculo diferencial a la Física, se vio que había una marcada diferencia entre los fenómenos estáticos (descritos usualmente por ecuaciones elípticas) y los dinámicos (regidos generalmente por ecuaciones hiperbólicas). Las variables físicas en los fenómenos estáticos podían suponerse sin dificultad de clase infinito, sin discrepancia con los datos observables; pero las variables físicas en los problemas dinámicos exhibían a menudo discontinuidades esenciales en la práctica. Por ejemplo, las funciones potenciales son de clase infinito en el espacio vacío, pero las funciones de onda pueden exhibir discontinuidades incluso cuando el frente de onda atraviesa una región sin perturbaciones. Así, una cuerda de violín pulsada en su punto medio vibra inicialmente de acuerdo con una ley de la forma

$$g(x, t) = a - \frac{1}{2} \left(\frac{a}{t} \right) (|x - ct| + |x + ct|), \quad 0 \leq ct \leq 1$$

y la derivada del desplazamiento es discontinua en $x = \pm ct$. Como la ecuación de ondas, que rige el movimiento de la cuerda, es de orden 2, la función anterior no puede interpretarse como solución de la misma en el marco tradicional de la teoría de ecuaciones diferenciales.

Por otro lado, dentro de la matemática tradicional, las hipótesis de validez de muchos teoremas son a veces de difícil comprobación en la práctica. Los ingenieros y técnicos vienen usando desde el pasado siglo diferentes «cálculos operacionales», en los que «todo vale»; el problema es que estos métodos, que adolecen de falta de rigor matemático, conducen a resultados satisfactorios. Puede decirse, como veremos en seguida, que desde el comienzo del cálculo diferencial se intuye la necesidad de extender estas nociones de modo que las operaciones fundamentales del análisis se puedan realizar siempre y sin hipótesis complicadas de validez. Se trata, pues, de alguna manera, de obtener una *noción generalizada de diferenciación* que permita derivar funciones (en principio) que no lo son en sentido ordinario. Este será, junto

con la noción paralela de *solución generalizada* de una ecuación diferencial, el primer origen de la teoría de distribuciones (cfr, prólogo de [6]).

El otro antecedente principal del origen de las distribuciones está relacionado con el anterior, aunque tiene unas raíces más próximas: se trata de las «funciones generalizadas» que fueron apareciendo cada vez con más frecuencia en diversas áreas del análisis, ligadas bien a las ecuaciones diferenciales (función de Green, teoría del potencial) o a los distintos cálculos operacionales que fueron surgiendo y usándose especialmente por los ingenieros (transformación de Fourier, transformación de Laplace) y los físicos teóricos (cálculo de Heaviside y, sobre todo, de Dirac). Entre estas funciones generalizadas primitivas, destaca, con luz propia, la ubicua «función δ », que hizo su primera aparición en sociedad de la mano de G. Kirchoff en 1882, en un estudio sobre la ecuación de ondas (aunque algunos resultados de Fourier podrían reinterpretarse por medio de la δ), y fue usada por una pléyade de matemáticos, ingenieros y físicos con distintos maquillajes, que muchas veces la hicieron irreconocible de unos autores a otros.

Dividiremos la exposición en dos grandes bloques: el primero sobre la generalización del concepto de función derivable y de solución de una ecuación diferencial, y el segundo sobre las «funciones singulares» y los cálculos operacionales, aunque es evidente que hay fuertes interrelaciones entre ambos grupos.

1. Diferenciación generalizada y soluciones generalizadas de E. D.

1. El período antiguo: 1750-1840

Para los matemáticos del siglo XVIII, no estaba muy claro el concepto de función, y menos aún el de función diferenciable. Los matemáticos de este período, en lugar de con una jerarquía bien establecida de funciones, trabajaban con una clase no muy bien definida: las «funciones continuas» (que llamaremos a partir de ahora *E*-continuas). Una función *E*-continua venía dada por una *expresión analítica*, entendiéndose por tal cualquier combinación de constantes y variables construida a partir de las operaciones matemáticas conocidas, algebraicas y trascendentes, incluyendo sumas infinitas, diferen-

ciación e integración. Se admitía que tales funciones se podían representar en el entorno de cada punto por una serie de potencias y, por tanto, en lenguaje moderno, serían funciones localmente analíticas. Después venían las «funciones discontinuas» (E -discontinuas) o «mecánicas», que aparecían definidas por expresiones analíticas distintas en distintos intervalos (o ¡incluso variando de punto a punto!). En cualquier caso, estos conceptos no estaban expresados ni aceptados de una manera universal, sino que variaban de un autor a otro.

¿Cuán regular ha de ser la solución de una ecuación diferencial?

En 1747, D'Alembert, a partir de un análisis de una ecuación en «diferenciales», encontró como solución del desplazamiento de una cuerda vibrante fija en los extremos, en función del tiempo t , la expresión:

$$y = f(x, t) = \psi(x + t) + \varphi(x - t)$$

donde las funciones «arbitrarias» ψ y φ pueden determinarse a partir de las condiciones iniciales. Euler, en 1753, exhibió la ecuación diferencial que rige el fenómeno que, en lenguaje moderno, sería equivalente a

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial t^2} \quad (1.1)$$

(recuérdese que la noción moderna de derivada, como límite de un cociente, se debe a Cauchy; en este período la noción de derivada parcial tenía interpretaciones esencialmente geométricas). Euler llega a la misma solución que D'Alembert, pero con interpretaciones esencialmente distintas. Para D'Alembert, las funciones «arbitrarias» ψ y φ deberían ser E -continuas, a fin de que se les pudiera aplicar las reglas del «análisis conocido». En una argumentación posterior (1761), D'Alembert usó el análogo geométrico de la expresión

$$f''(z) = \frac{f(z) + f(z + 2r) - 2f(z + r)}{r^2} \quad (r > 0, \text{ infinitamente pequeño})$$

obteniendo para la solución $f(x, t) = \psi(x+t) + \varphi(x-t)$ de la ecuación (1.1) la expresión

$$\begin{aligned} \frac{\partial^2 f(x, t)}{\partial x^2} &= \frac{\psi(x-t) + \psi(x+2r+t) - 2\psi(x+r+t)}{r^2} + \\ &\quad \frac{\varphi(x-t) + \varphi(x+2r-t) - 2\varphi(x+r-t)}{r^2} \\ \frac{\partial^2 f(x, t)}{\partial t^2} &= \frac{\psi(x+t) + \psi(x+t+2r) - 2\psi(x+t+r)}{r^2} + \\ &\quad \frac{\varphi(x-t) + \varphi(x-t-2r) - 2\varphi(x-t-r)}{r^2} \end{aligned} \quad (*)$$

Para que f fuera solución de la ecuación (1.1), ambas expresiones deberían ser iguales, según observó D'Alembert, lo que implicaba que los radios de curvatura en cualquier valor de $(x-t)$ deberían coincidir (no podían tener saltos, en palabras de D'Alembert). Como tales saltos del radio de curvatura eran posibles para funciones E -discontinuas, esas soluciones deberían desecharse. [Un lector moderno puede apreciar que el argumento de D'Alembert sustenta la idea de que las soluciones de (1.1) han de ser de clase \mathcal{C}^2 , y no localmente analíticas.]

Sin embargo, Euler mantenía la posición de que ψ y φ podían ser completamente arbitrarias (incluso E -discontinuas), con cualquier curva que pudiéramos dibujar como posible estado inicial. Naturalmente, Euler contestó a los argumentos de D'Alembert. Respecto a las críticas basadas en las reglas usuales del cálculo, se mostró de acuerdo en que el análisis conocido no podía solventar el problema, pero replicó que había que extender estas reglas para poder tratar las funciones E -discontinuas, ya que estas funciones aparecían de modo *obligatorio* en matemáticas como «constantes» de integración en la teoría de ecuaciones diferenciales en dos o más variables.

Pero, además de estas consideraciones generales, presentó algunos interesantes argumentos contra las objeciones de D'Alembert. Así, considerando el ejemplo de una solución construido a partir de dos expresiones analíticas distintas (dos parábolas, concretamente), de modo que el radio de curvatura tenga un salto en un punto B , Euler adujo las siguientes razones respecto al argumento de D'Alembert de que $\frac{\partial^2 f}{\partial x^2} \neq \frac{\partial^2 f}{\partial t^2}$:

- 1) El error afecta exclusivamente al comportamiento de la solución

en un punto y, por tanto, es de naturaleza infinitamente pequeña, comparado con el resto de los infinitos puntos del intervalo donde la solución es correcta desde el punto de vista ortodoxo (este argumento refleja las ideas de los matemáticos de la época sobre el análisis global y la indiferencia hacia las singularidades, y apunta hacia la dirección de considerar soluciones regulares a trozos para las ecuaciones diferenciales).

- II) Pero, además, basta alterar de modo infinitesimal las dos curvas en el punto B para que la f' sea regular en ese punto y el radio de curvatura no tenga saltos; esta alteración infinitesimal no alterará el movimiento de la cuerda. Este argumento es mucho más interesante desde el punto de vista de las distribuciones, pues anticipa el llamado «método de las sucesiones» para definir soluciones generalizadas de ecuaciones diferenciales, que sería ampliamente utilizado en el siglo XX. En efecto, si en lugar de una alteración infinitesimal pensamos en una sucesión (f_n) que converja a f en alguna topología, el argumento de Euler toma la forma siguiente: si las f_n son soluciones «clásicas» de la E.D. y convergen a f , entonces f es solución «generalizada» de la misma E.D.

Otra aportación interesante al tema, precursora también de otros métodos de definición de soluciones generalizadas, se debe a Lagrange (1736-1813). En su artículo *Recherches sur la nature et la propagation du son* (1759), Lagrange se mostró de acuerdo con los argumentos de D'Alembert sobre la solución de la ecuación diferencial de la cuerda vibrante, pero también pensó que la solución de Euler era correcta para describir el movimiento real de la cuerda. Para justificar esta opinión, obtuvo una nueva derivación del resultado de Euler que (en su opinión) no usaba el cálculo diferencial o el integral, origen de las dificultades. Su método consistía en considerar primero una cuerda sin peso con un número finito de masas puntuales, calcular el movimiento de esas masas, y luego hacer tender el número de puntos a infinito. Es decir, se trata de sustituir la ecuación diferencial por otra descripción matemática más adecuada para describir el fenómeno físico considerado; esta actitud fue seguida por un gran número de físico-matemáticos de finales del siglo XIX.

Sin embargo, en su *Nouvelles recherches sur la nature et propagation du son* (1760-1761), Lagrange volvió de nuevo a considerar la ecuación diferencial $\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial t^2}$ como eje central de su análisis, pero esta vez utilizando un interesante argumento: si se multiplica la ecuación por una función $M(x)$ y se integra parcialmente sobre $[0, a]$, se obtiene

$$\begin{aligned} \int_0^a \frac{\partial^2 f(x, t)}{\partial t^2} M(x) dx &= \int_0^a \frac{\partial^2 f(x, t)}{\partial x^2} M(x) dx = \\ &M(x) \frac{\partial f(x, t)}{\partial x} \Big|_{x=0}^{x=a} - \int_0^a \frac{\partial f(x, t)}{\partial x} M'(x) dx = \\ &\left[M(x) \frac{\partial f(x, t)}{\partial x} - f(x, t) M'(x) \right] \Big|_{x=0}^{x=a} + \int_0^a f(x, t) M''(x) dx \end{aligned} \quad (1.3)$$

Si se supone que f es 0 en los puntos $x = 0$ y $x = a$ de la cuerda, y se consideran sólo funciones M que se anulen en esos dos puntos, la ecuación (1.3) se transforma en

$$\int_0^a \frac{\partial^2 f(x, t)}{\partial t^2} M(x) dx = \int_0^a f(x, t) M''(x) dx \quad (1.4)$$

Imponiendo ahora que

$$M''(x) = -k^2 M(x) \quad (1.5)$$

(donde k es una constante), Lagrange obtiene de (1.4)

$$\int_0^a \frac{\partial^2 f(x, t)}{\partial t^2} M(x) dx = -k^2 \int_0^a f(x, t) M(x) dx \quad (1.6)$$

Si se introduce una nueva variable por la expresión

$$s = \int_0^a \frac{\partial^2 f(x, t)}{\partial t^2} M(x) dx \quad (1.7)$$

la ecuación (1.6) se transforma en la

$$\frac{d^2 s}{dt^2} = -k^2 s \quad (1.8)$$

Por tanto, la integración de la ecuación en derivadas parciales (1.1) se reduce a la integración del par de ecuaciones diferenciales ordinarias (1.5) y (1.8), ligadas ambas por (1.7), con lo cual «las diferenciales de f respec-

to de x desaparecen», por lo que, en palabras de Lagrange «on n'a point á craindre d'introduire par lá dans notre calcul aucune loi de continuité entre les différentes valeu de f ». Después Lagrange procedió a resolver (1.6) con unos extraños argumentos con integrales trigonométricas. La fórmula (1.8) contiene todavía las derivadas de f respecto de t , pero Lagrange trata de convencer al lector de que la dependencia respecto a t puede tratarse según las reglas usuales del cálculo, pues «la ley de continuidad se verifica en esta ocasión» (es obvio que no es éste el caso: la solución final presenta el mismo tipo de discontinuidades en x que en t). El descubrimiento importante de Lagrange es que el uso de la integración por partes permite trasladar el operador diferencial de la función incógnita a una cierta «función test» adecuada M . Como veremos, ésta es la idea básica en el cálculo con distribuciones. Naturalmente, con esta actitud Lagrange caía dentro de las críticas de D'Alembert. Sin embargo, rechaza estas críticas con el siguiente razonamiento: D'Alembert había calculado las derivadas segundas a partir de los valores de la función en los puntos $(z; z + r; z + 2r)$. Lagrange creía que debían tomarse los puntos $(z - r; z; z + r)$, es decir, tomar

$$f''(z) = \frac{f(z - r) + f(z + r) - 2f(z)}{r^2}, \quad (r > 0, \text{ inf. pequeño})$$

De esta forma, Lagrange encontraba que las dos expresiones de (*) eran idénticas, incluso en los puntos en los que el radio de curvatura tenía un salto (pero manteniéndose f de clase \mathcal{C}^1). De esta forma, se restablecía la validez de la solución de Euler para la ecuación de la cuerda vibrante. Este argumento de cambiar la definición de derivada, anticipa las consideraciones de Riemann y muchos matemáticos posteriores.

Durante finales del siglo XVIII y comienzos del XIX, este tipo de controversias fue decayendo, en parte porque habían desaparecido las dos figuras principales, Euler y D'Alembert (curiosamente, los dos murieron el mismo año: 1703), y también porque los matemáticos se habían dado cuenta de que era necesaria una fundamentación más rigurosa del cálculo para obtener una respuesta sin ambigüedades.

El movimiento hacia el rigor en el Análisis comenzó alrededor de 1820 con Cauchy y culminó en los años 1870, después del impulso decisivo de Weierstrass y su escuela. En este período se definen claramente conceptos como

«función n veces diferenciable» o «función de clase n », y se va imponiendo la idea de que sólo las funciones diferenciables podían ser (obviamente) soluciones de las ecuaciones diferenciales.

2. La edad del rigor: 1840-1870

En este período, la cuestión de la regularidad de las «funciones arbitrarias» que aparecían como constantes de integración en las E.D., fue completamente dejada de lado, yendo las preocupaciones de los especialistas por otros derroteros, como la existencia y unicidad de las soluciones para el problema de valores iniciales.

Sin embargo, cuando se abordaban problemas de la Física Matemática en donde aparecían condiciones iniciales no regulares (por ejemplo, una cuerda sometida a un fuerte plegamiento, tratado por Christoffel, o el problema de la propagación de ondas planas, en donde incluso si las condiciones iniciales son muy regulares, aparecen inevitablemente «ondas de choque» con singularidades, tratado por Riemann, Harnack y Christoffel, entre otros), la consideración de singularidades era inevitable. La idea usada es que la ecuación diferencial que rige el fenómeno no puede utilizarse en los puntos singulares, y debe sustituirse por otros modelos del sistema físico que eviten la consideración de «soluciones generalizadas».

Sin embargo, fuera del marco de la teoría de ecuaciones diferenciales, aparecen en esta época dos ejemplos conspicuos de generalización de la noción de diferenciación. El primero fue propuesto por Riemann en su *Habilitationarbeit* sobre las series trigonométricas (1854): dada una serie trigonométrica, no necesariamente de Fourier

$$\Omega(x) \sim \frac{1}{2} b_0 + \sum_{n=1}^{\infty} (a_n \operatorname{sen}(nx) + b_n \operatorname{cos}(nx))$$

si se integra formalmente dos veces, se obtiene otra

$$F(x) = C + C_1 x + b_0 x^2 - \sum_{n=1}^{\infty} \frac{1}{n^2} (a_n \operatorname{sen}(nx) + b_n \operatorname{cos}(nx))$$

que (suponiendo los a_n y b_n acotados, al menos) es uniformemente conver-

gente. El problema planteado por Riemann es cómo recuperar $\Omega(x)$ a partir de $F(x)$, cuando $\Omega(x)$ está definido (es decir, la serie trigonométrica converge en x). Riemann, en efecto, probó que si $\Omega(x)$ está definido, la fracción

$$\frac{F(x+r+s) - F(x+r-s) - F(x-r+s) + F(x-r-s)}{4rs}$$

converge a $\Omega(x)$ cuando r y s tienden a 0, siempre que r/s se mantenga finito y positivo. El límite de la expresión anterior generaliza la noción usual de derivada segunda de F en x , como es fácil ver, aunque Riemann no hace ningún comentario al respecto (nótese que cuando $r = s$, esta noción coincide con la derivada segunda generalizada de Lagrange).

La otra generalización del concepto de derivada, está ligada al desarrollo del cálculo integral durante este siglo, y los llamados «Teoremas fundamentales del cálculo integral» que, para una función continua en el intervalo $[a, b]$ se pueden enunciar:

- I) $F(x) = \int_a^x f(t) dt$ es derivable en $[a, b]$ y $F'(x) = f(x)$.
- II) Si f' es continua en $[a, b]$, $\int_a^x f'(t) dt = f(x) - f(a)$ para todo $x \in [a, b]$.

Las sucesivas extensiones de la noción de integral destruyeron la validez de (I) y (II) para funciones integrables, y se hicieron gran número de intentos para recuperar la simetría y belleza de estos teoremas en los distintos contextos (puede consultarse la interesante obra de T. Hawkins: *Lebesgue's theory of integration: its origin and development*, Chelsea, 1979, o Fernando Bombal: *La teoría de la medida: 1875-1925*).

Quizá la contribución más importante a este tema, antes de la invención de la integral de Lebesgue, la hizo el italiano U. Dini, quien en 1878 definió lo que hoy conocemos como derivadas de Dini (simplemente, se trata de sustituir el límite ordinario en la definición de derivada por los límites superior e inferior, a la derecha o a la izquierda del punto en consideración), y probó que si f era una función integrable Riemann, las cuatro derivadas de Dini de la función

$$F(x) = \int_a^x f(t) dt$$

existían y diferían de f en una función con integral nula, restableciendo así el teorema (I) en la integración Riemann.

La aparición de la integración de Lebesgue, condujo a la introducción de los conceptos de derivada en casi todo punto. En 1905, Vitali demostró rigurosamente un teorema anunciado por Lebesgue en su Tesis, probando que las funciones para las que (II) es cierto en el marco de la integración de Lebesgue son las que él llamó «absolutamente continuas». Concretamente, probó que una función F sobre $[a, b]$ es absolutamente continua si y sólo si existe una función f integrable Lebesgue en $[a, b]$, tal que $f = F'$ en casi todo punto y $F(x) - F(a) = \int_a^x f(t) dt$, para todo x de $[a, b]$. Vitali extendió su teorema al plano, después de generalizar adecuadamente la noción de función absolutamente continua. Otra definición alternativa fue dada por Tonelli en 1926. Tonelli quería caracterizar las superficies dadas por

$$z = f(x, y), \quad (x, y) \in Q = [0, 1] \times [0, 1], \quad f \text{ continua}$$

que tuvieran «área» según la definición dada por Lebesgue en 1902, y que además ese área viniera dada por la fórmula clásica

$$S = \iint_Q \sqrt{1 + p^2 + q^2}, \quad \text{con } p = \frac{\partial f}{\partial x} \quad q = \frac{\partial f}{\partial y}$$

Por tanto, necesitaba alguna condición que garantizase la existencia de las derivadas parciales p y q en algún sentido generalizado, de modo que existiera la integral y coincidiera con el área. La noción encontrada fue precisamente la de función absolutamente continua de dos variables.

Evidentemente, la generalización de los operadores diferenciales a funciones absolutamente continuas o derivables en casi todo punto era tan obvia que de nuevo se puso sobre el tapete, esta vez con suficiente bagaje matemático, la noción de «solución generalizada» de un operador diferencial, como veremos en la sección siguiente.

3. La protohistoria de las distribuciones: 1870 a 1950

Entre los muchos ejemplos que durante este período fueron forzando a la comunidad matemática a considerar «funciones generalizadas» como

soluciones de ecuaciones diferenciales, vamos a tratar, en primer lugar, uno de los más antiguos y que más trascendencia han tenido en el desarrollo del análisis: el problema de Dirichlet, es decir, dada una función f definida en la frontera de un dominio Ω , determinar una función u en $\bar{\Omega}$, tal que $\Delta u = 0$ en Ω y $u = f$ en $\partial\Omega$. El *principio de Dirichlet* establece que el problema de Dirichlet tiene una solución única, a saber, la función $u \in \mathcal{C}^1(\Omega) \cap \mathcal{C}^0(\bar{\Omega})$ que satisface la condición de contorno y minimiza la integral de Dirichlet

$$D(u) = \int_{\Omega} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx dy$$

Este principio, «probado» por Sir W. Thomson (1847) y Dirichlet (1876), fue invalidado por Weierstrass dentro de su programa de rigorización del análisis en general y del cálculo de variaciones en particular. La dificultad del principio es demostrar que el mínimo de $D(u)$ existe y se alcanza para una función u de clase \mathcal{C}^2 , pues en este caso se prueba sin dificultad que esta función es solución del problema de Dirichlet. Este principio influyó decisivamente en el desarrollo del análisis funcional y la topología, particularmente en la consideración de espacios funcionales y la decantación del concepto de compacidad.

En un principio, la clase de funciones considerada como admisible para minimizar $D(u)$ era $\{u \in \mathcal{C}^1(\Omega) \cap \mathcal{C}^0(\bar{\Omega}) : u|_{\partial\Omega} = f\}$ y la integral era la integral de Riemann. El matemático italiano Beppo Levi fue el primero en considerar la integral de Lebesgue en lugar de la de Riemann (1906), y no por *amore di generalità*, sino por *necessità di cose*, pues esta generalización de la integral le permitía usar una clase más amplia de funciones admisibles para buscar el mínimo [nótese que para que $D(u)$ tenga sentido basta que las derivadas primeras de u existan en casi todo punto y sean integrables]. La clase de funciones admisibles considerada por Beppo Levi es (salvo la condición de que tomen los valores prefijados en la frontera) esencialmente lo que hoy conocemos como espacio de Sobolev H_2^1 (es decir, funciones f integrables, tales que su derivada en el sentido de las distribuciones, pertenece también a L^2). Beppo Levi demostró que dada una sucesión minimal en este espacio, se podía extraer una subsucesión uniformemente convergente, en cuyo límite $D(u)$ alcanzaba el mínimo. También probó que esta extensión no introducía nuevas soluciones al problema, pues la función minimal era de

hecho una función \mathcal{C}^2 .

El trabajo de B. Levi no fue olvidado, pues Otto Nikodym retomó sus ideas en una comunicación *Sobre el principio del mínimo* en 1932, en relación con el estudio de una clase de E.D. que incluía la de Laplace, introduciendo las funciones que él llamó (BL) (de Beppo Levi) en un rectángulo Q de \mathbb{R}^3 por las condiciones:

- 1) F está definida en casi todo punto de Q .
- 2) $F(x_0, y_0, z)$ es absolutamente continua en z , para casi todos los valores de (x_0, y_0) , y lo mismo al permutar las otras variables.
- 3) Las derivadas primeras de F están en $L^2(Q)$.

En lugar de operar con la convergencia puntual o la uniforme, usó la norma

$$\|f\|_Q = \iint_Q \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 + \left(\frac{\partial f}{\partial z} \right)^2$$

lo que le dio un instrumento analítico muy potente para el estudio del problema de Dirichlet y otros con él relacionados. Exhibiendo una tendencia cada vez más extendida en este período, Nikodym presentó también una serie de teoremas sobre las propiedades de las funciones (BL), independientemente de las aplicaciones, que habían sido su razón de ser.

El método más utilizado para extender la noción de solución de una E.D. se basa en una idea muy simple: dada una ecuación diferencial A de orden n , se trata de encontrar otra ecuación o condición B que para funciones de clase n sea equivalente a la original, pero que tenga sentido para funciones más generales. Los objetos que satisfagan B , se llamarán entonces *soluciones generalizadas* de A . Como regla general, la nueva condición B se obtiene por alguna forma de integración por partes. Por ello, a menudo la condición a satisfacer por las soluciones generalizadas es de la forma «la ecuación B se cumple para todos los objetos de una cierta clase». Estos objetos pueden llamarse, pues, objetos de prueba («objetos test»).

El primero en usar y sacar provecho a estas ideas fue el profesor de Harvard M. Bôcher en relación con la extensión de la noción de función armónica: usando la fórmula de Green

$$\iint_{\Omega} (u\Delta v - v\Delta u) = \int_{\partial\Omega} \left(v \frac{\partial u}{\partial n} - u \frac{\partial v}{\partial n} \right)$$

Bôcher observó que tomando $v = 1$, si u es armónica en Ω , entonces $\int_{\partial\Omega} \frac{\partial u}{\partial n} = 0$, condición que sólo involucra las derivadas primeras de u . Bôcher define entonces una función armónima generalizada en Ω como una función de clase 1, tal que para todo círculo Γ contenido en Ω , verifique

$$\int_{\partial\Gamma} \frac{\partial u}{\partial n} = 0$$

El trabajo de Bôcher se centra en probar que, de hecho, no se introducen nuevas funciones con esta generalización. Los trabajos de Bôcher fueron descubiertos independientemente en el mismo año (1906) por Koebe, para el caso de funciones en 3 variables, quien además caracterizó las funciones armónicas por la propiedad de la media superficial *e. d.*, una función continua en Ω es armónica si para todo $P \in \Omega$ y para toda bola B de centro P contenida en Ω

$$u(P) = \frac{1}{\text{área}(\partial B)} \int_{\partial B} u$$

Las ideas de Bôcher fueron continuadas por su colega G. C. Evans, en sus trabajos sobre potenciales de masas arbitrarias en el plano y el espacio. En 1914 dio la primera demostración rigurosa de que un problema en los límites podía admitir solución con condiciones iniciales de diferenciación más débiles que las clásicas, dando así apoyo a la postura de Euler sobre la cuerda vibrante. En su trabajo de 1914, Evans considera la ecuación

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f(x, y)$$

y la sustituye por la «ecuación generalizada»

$$\int_S \left[\frac{\partial u}{\partial x} dt + u dx \right] = \iint_{\sigma} f(x, t) dx dt$$

donde S es una curva cerrada arbitraria de una cierta clase Γ y σ la región que encierra S . La nueva ecuación tiene sentido para funciones que sean solamente de clase 1 en el dominio considerado, y Evans obtiene, efectivamente, soluciones no regulares.

En 1920 utilizó la misma técnica para transformar la ecuación de Poisson $\Delta u = F$ por medio de la fórmula de Green. Evans quiso eliminar incluso la exigencia de ser u de clase 1 en la ecuación generalizada, por lo que tuvo que extender la noción de derivada en \mathbb{R}^2 . Esto le permitió abordar problemas de potenciales de masas mucho más singulares que los tratados hasta entonces. Posteriormente, Evans analizó las relaciones entre sus «derivadas generalizadas» y la derivación en casi todo punto, probando que, en el caso de funciones continuas, su definición coincide con la noción de función absolutamente continua en \mathbb{R}^2 introducida por Tonelli en 1926, estableciendo así, por primera vez, la conexión entre dos definiciones de derivadas generalizadas que se habían originado en contextos muy diferentes.

Los métodos de Evans fueron extendidos a \mathbb{R}^n y desarrollados por otros dos americanos: J. W. Calkin y C. B. Morrey en 1940, quienes desarrollaron la teoría independientemente de sus orígenes (teoría del potencial). Las mismas ideas con distintas formulaciones y resultados fueron utilizadas por otros matemáticos de la época en relación con la teoría del potencial, como H. Werl (1940) y T. Rado (1937).

Las soluciones generalizadas para las E.D. hiperbólicas aparecieron mucho más tarde, a pesar de que fueron el origen del problema. Este hecho muestra claramente el diferente estado de desarrollo de ambas teorías. Los métodos empleados para resolver problemas de valores iniciales para este tipo de ecuaciones al principio del siglo XX, se basaban, más o menos, en técnicas operacionales inspiradas en el cálculo operacional de Heaviside, de difícil fundamento matemático. Una excepción a esta regla es el trabajo de N. Wiener *El cálculo operacional* (1926) en el que estudia soluciones generalizadas de E.D., que fue la primera exposición rigurosa del método de las funciones test, cuya idea básica es transferir por integración por partes el operador diferencial de la función incógnita a una función test suficientemente regular. Pero dejemos al mismo Wiener exponer esta idea. Dice Wiener:

Consideremos la ecuación lineal

$$L(u) = A \frac{\partial^2 u}{\partial x^2} + B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Fu = 0$$

donde, por simplicidad, supondremos que los coeficientes tienen

tantas derivadas como necesitemos. Si u satisface esta ecuación, obviamente debe poseer las derivadas que aparecen en la misma. Sin embargo, como es familiar en el caso de la ecuación de la cuerda vibrante, hay casos donde u puede considerarse como solución de nuestra ecuación diferencial en un sentido generalizado, sin poseer las derivadas de todos los órdenes indicados en la ecuación, e incluso sin ser diferenciable. Por tanto, tiene cierto interés precisar la manera en la que una función no diferenciable puede satisfacer una E.D. en un sentido generalizado.

Sea $G(x, y)$ una función positiva e indefinidamente diferenciable en una cierta región poligonal acotada R del plano XY , que se anula, junto a todas sus derivadas, en el borde de R y es 0 fuera de R . Entonces existe una función $G_1(x, y)$, tal que

$$\iint_R L(u)G(x, y) \, dx dy = \iint_R u(x, y)G_1(x, y) \, dx dy$$

para toda función u con derivadas hasta el orden 2 acotadas e integrables, como puede probarse por integración por partes. Por tanto, una condición necesaria y suficiente para que u satisfaga nuestra E.D. en casi todo punto es que

$$\iint_R u(x, y)G_1(x, y) \, dx dy = 0$$

para toda posible G (ya que las G 's forman un sistema completo sobre cualquier región), y que u posea las derivadas requeridas. Podemos, por tanto, considerar una función ortogonal a toda G_1 , como una solución de nuestra ecuación diferencial en un sentido generalizado. Este sentido es más general que el desarrollado por Bôcher, pues no se postula la existencia de las derivadas primeras. Si una sucesión de soluciones generalizadas de nuestra ecuación diferencial converge en media sobre un recinto dado a una función $\Phi(x, y)$, es claro que Φ es, a su vez, una solución generalizada de la ecuación diferencial sobre esa región.»

Es importante destacar la última observación de Wiener, pues esa propiedad era precisamente la que necesitaba en su estudio sobre la ecuación de los tele-

grafistas que realiza a continuación. Es el germen del llamado método de las sucesiones para definir soluciones generalizadas. Sin embargo, como sucede una y otra vez en esta época, esta definición de Wiener tiene un carácter específico para resolver el problema que le interesa en ese momento. Así, en su trabajo posterior de 1927 utiliza otra definición de solución generalizada, mejor adaptada a su problema concreto. Como se ve, las soluciones generalizadas de E.D. no constituían una teoría unificada, y tenían un carácter *ad hoc* en cada trabajo.

Antes de pasar a comentar los trabajos fundamentales de Sobolev, citaremos someramente otras dos contribuciones al tema. La primera, la del matemático francés J. Leray, profesor de L. Schwartz. Leray, estudiando el problema de Cauchy para las ecuaciones de Navier-Stokes, que regulan el movimiento de los líquidos viscosos, a fin de evitar incluso la hipótesis de existencia de derivadas parciales de la posible solución, introduce su noción de «casi-derivada»:

Dadas $U, U_i \in L^2(\mathbb{R}^3)$, se dice que U_i es la casi-derivada de U respecto de la variable y_i si

$$\int_{\mathbb{R}^3} \left[U(y) \frac{\partial a(y)}{\partial y_i} + U_i(y) a(y) \right] dy = 0$$

para toda función a de clase $\mathcal{C}^1(\mathbb{R}^3)$, tal que ella y sus derivadas primeras pertenezcan a L^2 .»

Esta definición, análoga a la de Wiener y coincidente con la de la teoría de distribuciones, está basada, como vemos, en el método de las funciones test. Como Wiener, Leray estaba muy interesado en la siguiente propiedad de esta generalización: toda $U \in L^2$ se puede aproximar por una sucesión $U_n = U * \rho_n$ de funciones suaves (tomando como (ρ_n) una aproximación suave de la identidad). Pues bien, si U tiene casi-derivada U_i ésta es el límite débil en L^2 de las derivadas (ordinarias) $\partial U_n / \partial y_i$. Esta propiedad, junto a su recíproca, permitieron a Leray probar la existencia de una solución generalizada para $t \geq 0$ de la ecuación en cuestión, por el método de resolver primero un problema próximo, más regular, y pasar al límite.

Finalmente, y sin pretender ser exhaustivos, señalemos que el método

de sucesiones y funciones test fue usado por Bochner en un trabajo sobre ecuaciones en derivadas parciales con coeficientes constantes, publicado en 1946, el mismo año en que Schwartz publicó su primer artículo sobre la teoría de distribuciones. Bochner definió una solución generalizada de la ecuación $Pf = 0$ en un dominio D como una función f definida en casi todo punto de D , localmente integrable y tal que para cada punto $x_0 \in D$ existe un entorno $U(x_0)$ en el que f es límite débil (en L^1) de soluciones estrictas. Después de dar esta definición por sucesiones, probó la equivalencia con una definición basada en el método de las funciones test y demostró varios interesantes teoremas (por ejemplo, una solución generalizada de la ecuación de Laplace es, salvo corrección en un conjunto de medida 0, una solución clásica; las soluciones generalizadas en un subdominio se pueden extender a soluciones generalizadas en el dominio total, etc.).

Por último, señalar que la primera definición de solución generalizada de una E.D.P. aparecida en un libro de texto fue la basada en el método de funciones test, incluida en la edición de 1937 del clásico *Methoden der Mathematischen Physik, II*, de Courant-Hilbert. En cualquier caso, el uso de soluciones generalizadas en la teoría de ecuaciones hiperbólicas es aún menos coherente y unificado que en la teoría del potencial.

La contribución de Sobolev

Sergei Sobolev, nacido en 1908, estudió en San Petersburgo y dedicó la mayor parte de su vida profesional al estudio de las E.D.P. Su temprano interés por la ecuación de ondas y otras ecuaciones hiperbólicas refleja su empleo en el Instituto Sismológico durante los años 1929 a 1932. En 1933, inició una serie de trabajos sobre el problema de Cauchy que le llevaron, dos años más tarde, a la definición de función generalizada. En una nota corta de 1935, Sobolev esbozó cómo podía obtenerse un teorema muy general de existencia para el problema de Cauchy, extendiendo el problema a un cierto espacio de funcionales. Este esbozo fue desarrollado el año siguiente, en el artículo *Methode nouvelle á résoudre le problème de Cauchy pour les equations linéaires hyperboliques normales*. El problema que abordó Sobolev

fue resolver la ecuación

$$L(u) = \sum_{i=1}^{2k+1} \sum_{j=1}^{2k+1} A_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{2k+1} B_i \frac{\partial u}{\partial x_i} + Cu - \frac{\partial^2 u}{\partial t^2} = F \quad (1.9)$$

con las condiciones iniciales

$$u|_{t=0} = u^{(0)}(x, \dots, x_{2k+1}) \quad \left. \frac{\partial u}{\partial t} \right|_{t=0} = u^{(1)}(x_1, \dots, x_{2k+1}) \quad (1.10)$$

donde A_{ij} , B_i y C se suponen analíticas y la forma cuadrática

$$\sum_{i=1}^{2k+1} \sum_{j=1}^{2k+1} A_{ij} p_i p_j$$

definida positiva.

Suponiendo que existe una solución de (1.9) y (1.10) u con derivadas acotadas hasta el orden $k+1$ inclusive, Sobolev encontró una ecuación integral que debe satisfacer u que, por el método de las aproximaciones sucesivas le permitió obtener la expresión explícita $u = \sum_{n=0}^{\infty} J_n$, donde J_0 está unívocamente determinada por las condiciones iniciales (1.10) y J_n se definen recursivamente a partir de J_0 . Las hipótesis sobre u permitieron probar a Sobolev la convergencia de la serie $\sum_{n=0}^{\infty} J_n$. Sin embargo, desde un punto de vista matemático, esta solución era poco satisfactoria, pues en primer lugar las condiciones naturales a satisfacer por la solución es que u sea de clase 2, y en segundo lugar las hipótesis para obtener solución deberían hacerse sobre los datos $u^{(0)}$, $u^{(1)}$ y F , y no sobre u . Veamos cómo, para obtener este teorema de existencia, procedió Sobolev:

En primer lugar, definió los espacios Φ_s , de las funciones de clase s sobre \mathbb{R}^{2k+2} , con soporte compacto, y una convergencia en Φ_s dada por: $(\varphi_n) \rightarrow \varphi$ si y solo si:

- 1) Las φ_n y φ tienen soporte en un compacto fijo.
- 2) $(D^\alpha \varphi_n)$ converge uniformemente a $D^\alpha \varphi$ para todo n y todo índice de derivación α con $|\alpha| \leq s$.

(Este es el espacio que Schwartz llamó $D^s(\mathbb{R}^{2k+2})$, con la misma noción de

convergencia, aunque Schwartz probó después que esta convergencia era la asociada a una topología adecuada sobre el espacio.)

El espacio fundamental de funcionales Z_s , lo definió Sobolev como el conjunto de funcionales lineales sobre Φ_s que son sucesionalmente continuos. En Z_s se define una convergencia por

$$\rho_n \rightarrow \rho \stackrel{def}{\iff} (\rho_n, \varphi) \rightarrow (\rho, \varphi) \quad \forall \varphi \in \Phi_s$$

(misma definición que Schwartz para convergencia de sucesiones de distribuciones, cuando $s = \infty$). Toda función f con l derivadas localmente integrables en \mathbb{R}^{2k+2} define un funcional en Z_s , (funcional de grado l), por la fórmula

$$(\rho_f, \varphi) = \int f(t)\varphi(t) dt \quad \forall \varphi \in \Phi_s$$

(es la misma fórmula que usó Schwartz para sumergir las funciones localmente integrables en el espacio de distribuciones). Después, Sobolev probó que los funcionales de grado infinito son secuencialmente densos en Z_s , por un proceso standard de regularización, análogo al empleado por Schwartz, de modo que los elementos de Z_s son, en último término, límites de funciones de Φ_s . Posteriormente, Sobolev observó que si $L : \Phi_{s_1} \rightarrow \Phi_{s_2}$ es lineal y (sucesionalmente) continuo, el operador transpuesto, definido por $(L^*\rho, \varphi) = (\rho, L\varphi)$ aplica Z_{s_2} en Z_{s_1} , y así Sobolev pudo definir una multiplicación en $Z_{\min(s', s_2)}$ por elementos de $\mathcal{C}^{s'}$, como el adjunto del operador de multiplicación en Φ_{s_2} . Del mismo modo, definió el operador D_i , de derivación respecto a x_i , en Z_{s-1} como el adjunto de $-\frac{\partial}{\partial x_i}$ en Z_s , es decir:

$$(f\rho, \varphi) = (\rho, f\varphi), \quad \forall f \in \mathcal{C}^{s'}, \quad \forall \rho \in Z_{\min(s', s_2)}, \quad \forall \varphi \in \Phi_{s_2}$$

$$(D_i\rho, \varphi) = - \left(\rho, \frac{\partial \varphi}{\partial x_i} \right), \quad \forall \rho \in Z_{s-1}, \quad \forall \varphi \in \Phi_s$$

Por integración parcial se prueba que estas nuevas operaciones son consistentes con las antiguas cuando se aplican a operadores de grado l . De esta forma, el operador L se generaliza al espacio Z_s . Después, Sobolev volvió al problema original (1.9). Definió \underline{u} como el operador de grado 0 correspondiente a la función igual a u para $t \geq 0$, y nula para $t < 0$ y probó que si u

satisface (1.9) y (1.10), entonces se debería cumplir

$$L(\underline{u}) = \rho \quad \underline{u}|_{t < 0} = 0 \quad (1.11)$$

en Z_s , donde ρ es el funcional

$$(\rho, \varphi) = \iiint_{t > 0} F\varphi \, dx_1 \dots dt + \iint_{t > 0} \left(u^{(0)} \frac{\partial \varphi}{\partial t} - u^{(1)} \varphi \right) dx_1 \dots dx_{2k+1}$$

y $\underline{u}|_{t < 0} = 0$, significa que (\underline{u}, φ) no depende de los valores de φ para $t < 0$ (en lenguaje moderno, soporte de $\underline{u} \subset \{x_1, \dots, x_{2k+1}, t\} : t \geq 0\}$). Recíprocamente si \underline{u} es un funcional solución de (1.11), que corresponde a una función u con derivadas de primer y segundo órdenes integrables, entonces Sobolev probó fácilmente que u era una solución ordinaria del problema original. En consecuencia, Sobolev abordó la resolución de (1.11). En lugar de intentar hallar un inverso de L en todo Z_s Sobolev introdujo ciertos espacios intermedios de funciones, de modo que

$$\Phi_s \subset \Omega_s \subset \Psi_s, \quad Y_s = (\Omega_s)', \quad W_s = (\Psi_s)'$$

y, por dualidad, $Z_s \supset Y_s \supset W_s$. Tomando después $Y = \bigcup_{s=0}^{\infty} Y_s$, $W = \bigcup_{s=0}^{\infty} W_s$ y usando la expresión explícita de la solución u de clase $k+1$ que había obtenido al principio, Sobolev probó que el operador L definido en Y y en W tenía inverso por la izquierda y por la derecha. Como tanto ρ como la solución deseada de (1.11) estaban en Y , esto permitía probar la existencia y unicidad de una solución de (1.11). En consecuencia, si la solución en Y venía dada por una función u de clase \mathcal{C}^2 , esta función era solución del problema original (1.9) y (1.10), y si la solución no era de este tipo, el problema original no tenía solución clásica. Por ejemplo, este método permitió a Sobolev probar que si $F \in \mathcal{C}^{k+1}$, $u^{(0)} \in \mathcal{C}^{k+3}$ y $u^{(1)} \in \mathcal{C}^{k+2}$, el problema tiene solución clásica.

Comparemos los resultados de Sobolev con la teoría de distribuciones de Schwartz desarrollada posteriormente. Los puntos de semejanza son muchos: Sobolev, como Schwartz, quería generalizar el concepto de función y algunas operaciones clásicas con funciones a un conjunto mayor, donde cierto problema (el problema de Cauchy para (1.9), (1.10) en el caso de Sobolev) pudiera resolverse más fácilmente. Los métodos de generalización de los conceptos de función y las operaciones son análogos: funcionales y transposición. Incluso

la noción de convergencia en los espacios de funciones test coinciden. En este sentido, puede decirse que Sobolev inventó las distribuciones. Sin embargo, hay también grandes diferencias entre ambas teorías. Quizá las más profundas están en los conceptos que Schwartz introduce y no Sobolev: los espacios de distribuciones temperadas y con soporte compacto, las transformaciones de Fourier y Laplace de distribuciones, los productos tensoriales y la convolución de distribuciones, la concepción de la δ y las «partes finitas» de Hadamard como distribuciones, etc. Pero la diferencia esencial es que Sobolev creó estos métodos como herramienta para resolver un problema concreto, mientras que Schwartz desarrolló una teoría completa, versátil y muy potente, aplicable y aplicada por él mismo a la solución de muchos problemas diferentes. Así puede decirse que, si bien Sobolev inventó las distribuciones, Schwartz creó la teoría de distribuciones como cuerpo de doctrina.

Sobolev no continuó el estudio de los espacios de funcionales. En sus trabajos posteriores se mantuvo siempre dentro de los conceptos tradicionales de función, aunque sí usó sistemáticamente la noción generalizada de diferenciación. En un trabajo sobre difracción de ondas, utilizó también el método de las sucesiones para definir soluciones generalizadas de una ecuación $P(u) = 0$: $u \in L^1$ es solución generalizada si existe una sucesión (u_n) de soluciones clásicas que convergen a u en L^1 . En el mismo trabajo, Sobolev probó que esta noción es equivalente a la definición por medio de funciones test.

En sus trabajos posteriores, Sobolev introdujo los espacios funcionales que hoy conocemos con su nombre

$$W_p^m \text{ (o } H_p^m) = \{f \in L^p : D^\alpha f \in L^p \text{ para } |\alpha| \leq m\}$$

(donde las derivadas en la definición anterior están tomadas en sentido generalizado). En 1938 anunció sus dos teoremas de inmersión, que fueron (y son) ampliamente utilizados, pues en muchas ocasiones permiten concluir que una solución generalizada de una E.D.P. en un espacio de Sobolev adecuado, es de hecho una solución clásica.

Como ya hemos dicho en otro lugar, Sobolev no fue el primero en definir o usar los espacios que llevan su nombre. De hecho, habían aparecido en mu-

chos problemas sobre E.D.P., especialmente en el tratamiento variacional del problema de Dirichlet [Beppo Levi (1906), Tonelli (1926), Nikodym (1933), Evans (1920), Calkin (1940) y Morrey (1940)], con diferentes nociones de derivadas generalizadas y diferentes valores de m y p . Sobolev, que conocía bien el trabajo de algunos de sus predecesores, escribió (1936): *Nos théorèmes donnent une précision des évaluations connues, due à l'école de Göttingen, qu'on rencontre souvent dans différents problèmes de la théorie des E.D.P.*

En resumen, la diferenciación generalizada influyó decisivamente en el desarrollo y utilización de los espacios de Sobolev, pero los funcionales de Sobolev (distribuciones) no volvieron a ser usados en este contexto hasta su redescubrimiento por L. Schwartz, quien según sus propias declaraciones, no conoció los trabajos de Sobolev al respecto antes de 1945.

Quizá sea éste un buen momento para resumir los distintos métodos de definición de soluciones generalizadas de E.D.P. que hemos visto hasta ahora.

A) *Sustitución de una E.D. por otro modelo de sistema físico*

Método dirigido al tratamiento de singularidades que el modelo original no puede abordar dentro de un marco riguroso: ecuaciones de movimiento: Lagrange (1759), Riemann (1859), Christoffel (1876), etc.

B) *Sustitución de varios límites por uno*

Se definen las derivadas de orden superior tomando un sólo límite en lugar de límites sucesivos en las derivadas de orden inferior: Lagrange (1760-1761), Riemann (1867, en relación con las series trigonométricas), etc. Las expresiones diferenciales compuestas se consideran como un todo, en lugar de una suma de distintos operadores, cada uno con un proceso de paso al límite [Petrini (1905), Wiener (1927), etc.]. Dini, en 1878, sustituyó el límite ordinario en la definición de derivada por límite superior o límite inferior («derivadas de Dini»).

C) *Diferenciación en casi todo punto*

En dimensión uno, en relación con la teoría de la medida e integral, fue usado por Lebesgue (1902) y Vitali (1905), entre otros. En dos

dimensiones, fue aplicado a la teoría de la medida e integración por Tonelli (1926) y Morrey, y en el cálculo de variaciones por Beppo Levi (1906), Fubini (1907), Tonelli (1929) y Nikodym (1933). En este último contexto, esta noción sirvió para la introducción de los espacios de Sobolev.

D) *Método de las curvas y superficies de prueba*

Partiendo del teorema de Stokes, en cualquiera de sus versiones clásicas, por integración de la E.D. objeto de estudio, ésta se puede transformar en una ecuación integro-diferencial, que debe cumplirse para todos los dominios de integración suficientemente regulares (curvas o superficies). En general, la nueva ecuación tiene sentido para un conjunto de funciones mayor que la original y, por tanto, la generaliza. El método fue empleado esencialmente en la teoría del potencial [Bôcher (1905-1906), Evans (1914-1933), Calkin (1940), etc.].

E) *Método de las funciones test*

La E.D. a estudiar se multiplica por una función test suficientemente regular con soporte compacto en un cierto dominio, y el resultado se integra por partes. De esta forma, el operador diferencial se transfiere a la función test, y la ecuación integro-diferencial resultante, que ha de verificarse para todas las funciones test, no presupone ninguna regularidad de la solución. Es un método muy relacionado con el anterior, y su origen está en el estudio de las ecuaciones hiperbólicas y parabólicas. Fue anticipado por Lagrange (1761) y formulado explícitamente por Wiener (1926), y después por Leray, Sobolev y Courant-Hilbert. Es el método básico de la teoría de distribuciones.

F) *Método de las sucesiones*

Se define una solución generalizada de una E.D. como una función tal que es límite (en algún sentido) de una sucesión de soluciones clásicas. Anticipado por Euler (1765) y Laplace (1772), y explícitamente formulado por Sobolev (1935; convergencia en L^1), Friedrichs (1939; convergencia en $\|f\| + \|Df\|$), Schwartz (1944; convergencia uniforme sobre compactos), Bochner (1945; convergencia débil en L^1 , en el entorno de cada punto), Wiener (1926; convergencia en

L^2) enunció el método como consecuencia de su definición, así como Leray (1934; convergencia débil o en norma en L^2).

Como hemos visto a lo largo de la exposición, estos métodos son interdependientes, e incluso un mismo autor ha usado distintos métodos en distintos trabajos.

2. Las funciones singulares

1. La función δ

La más importante de todas las funciones singulares que fueron surgiendo en matemáticas, especialmente en relación con la mecánica cuántica es, sin duda, la llamada función δ (notación introducida por Dirac), y será a ella principalmente a quien dedicaremos gran parte de nuestro análisis.

El significado físico de la función δ es tan grande, que la idea subyacente en esta primera «función singular» puede remontarse en el tiempo tan lejos como se desee. Pero nos restringiremos a los casos en los que este concepto se aborda en un contexto matemático.

Es de destacar la habilidad de físicos y matemáticos para evitar el uso de la paradójica y matemáticamente problemática función δ en lugares en los que parece jugar un papel fundamental. Por ejemplo, en el estudio de las fuerzas eléctricas o gravitacionales, el método tradicional consiste en estudiar primero las cargas (o masas) puntuales y después pasar a estudiar el caso de distribuciones continuas de cargas o masas en el espacio. Sin embargo, esta última teoría no puede tratar los problemas originales de cargas o masas puntuales en un sentido riguroso.

En cualquier caso, la función δ tuvo una infancia desgraciada, pues ni los matemáticos ni los físicos la reconocieron como un objeto propio de sus respectivos dominios. Para los matemáticos se trataba de una noción intuitiva, sin realidad matemática; para los físicos, se trataba de una idealización puramente matemática, que no existe en la naturaleza. Así, por ejemplo, Maxwell, en su tratado sobre electricidad y magnetismo (1873), considera

las cargas puntuales como «pequeños cuerpos cuyas dimensiones son despreciables comparadas con las distancia consideradas».

Veamos ahora alguna de las formas en las que la ubicua δ ha aparecido en matemáticas:

En su renombrada *Théorie analytique de la chaleur* (1822), Fourier «prueba» la convergencia de la serie de Fourier de una función arbitraria de período 2π , simplemente considerando la expresión de la suma N -sima:

$$S_N(f)(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} \left(\sum_{i=-N}^N \cos(i(x-t)) \right) f(t) dt$$

y transformando el núcleo de Dirichlet que aparece en la expresión anterior

$$\sum_{i=-N}^N \cos(i\theta) = \cos(N\theta) + \operatorname{sen}(N\theta) \frac{\operatorname{sen}(\theta)}{1 - \cos(\theta)} \quad (\theta = t - x)$$

donde N puede ser infinito. Por una serie de razonamientos, Fourier intenta convencer al lector que la contribución del coseno para N infinito se anula, mientras que

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \operatorname{sen}(N(t-x)) \frac{\operatorname{sen}(t-x)}{1 - \cos(t-x)} dt = f(x)$$

para N infinito. En lenguaje moderno, esta fórmula puede expresarse como

$$\lim_{N \rightarrow \infty} \frac{1}{\pi} \operatorname{sen}(N\Theta) \frac{\operatorname{sen}(\Theta)}{1 - \cos(\Theta)} = \delta(\Theta)$$

fórmula que es correcta en la teoría de distribuciones.

Aunque Fourier escribe literalmente que la expresión

$$\frac{1}{2} + \sum_{i=1}^{\infty} \cos(i(x-t)) \quad (\text{otra expresión del núcleo anterior}) \quad (2.12)$$

representa una función con la propiedad de la δ , los rigoristas del siglo XIX interpretaron la igualdad

$$f(x) = \int_{-\pi}^{\pi} \delta(x-t) f(t) dt \quad (2.13)$$

donde $\delta(x - t)$ es la expresión (2.12), como un límite, ya que «como la serie (2.12) tiene una suma indeterminada, no se le puede asignar un sentido a la expresión considerada por Fourier» (Darboux).

Fourier emplea también razonamientos análogos, usando discrecionalmente los infinitos, para «demostrar» el teorema de inversión de la integral de Fourier, estableciendo la fórmula (válida en la teoría de distribuciones)

$$\lim_{p \rightarrow \infty} \frac{1}{\pi} \frac{\text{sen}(p(t-x))}{t-x} = \delta(t-x)$$

La idea de demostrar el teorema de inversión de Fourier o la convergencia de la serie de Fourier probando que el núcleo de Dirichlet u otro análogo tiende a la función [caracterizada por la propiedad (2.13)], aparece con frecuencia (obviamente, no con esta formulación), en muchos trabajos sobre análisis de Fourier de la primera mitad del siglo XIX. Las pruebas rigurosas de los teoremas de convergencia, a partir de la dada por Dirichlet, utilizan las llamadas integrales singulares o bien los métodos generalizados de convergencia (Abel, Cesaro, Gauss), que pueden interpretarse (y de hecho son) como aproximaciones de la δ ; pero como los límites se toman fuera del signo integral, no aparece explícitamente el uso de «funciones singulares».

La primera definición matemática explícita de la función δ se debe a Gustav R. Kirchoff, en su tratamiento del principio de Huygens para la ecuación de ondas

$$\frac{\partial^2 u}{\partial t^2} = a^2 \Delta u \quad (2.14)$$

en 2 y 3 coordenadas espaciales (1882). Siguiendo un proceso ya habitual, Kirchoff comenzó usando el teorema de Green para dos funciones $u(\bar{x}, t)$ y $v(\bar{x}, t)$, con t fijo y \bar{x} en un dominio G :

$$\int_{\partial G} \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) d\sigma = \int_G (v \Delta u - u \Delta v) d\bar{x}$$

Si u y v satisfacen (2.14), entonces el segundo miembro de la expresión anterior se convierte en

$$\frac{1}{a^2} \int_G \left(v \frac{\partial^2 u}{\partial t^2} - u \frac{\partial^2 v}{\partial t^2} \right) d\bar{x} = \frac{1}{a^2} \frac{\partial}{\partial t} \int_G \left(v \frac{\partial u}{\partial t} - u \frac{\partial v}{\partial t} \right) d\bar{x}$$

que después de integrar entre $-t'$ y t'' se transforma en

$$\int_{-t'}^{t''} dt \int_{\partial G} \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) d\sigma = \frac{1}{a^2} \left[\int_G \left(v \frac{\partial u}{\partial t} - u \frac{\partial v}{\partial t} \right) d\bar{x} \right]_{-t'}^{t''}$$

Entonces Kirchoff consideró la función auxiliar v , tal que

$$v(\bar{x}, t) = \frac{F(|\bar{x} - \bar{x}'| + at)}{|\bar{x} - \bar{x}'|}$$

donde F es una «función», tal que

$$\int_I F(s) ds = 1$$

«para todo intervalo I con extremos de distinto signo». Kirchoff pone como ejemplo la función $F(r) = \frac{r}{\sqrt{\pi}} \exp(-r^2 t^2)$, para r constante positiva muy grande [notemos de paso que $\lim_{r \rightarrow \infty} \frac{r}{\sqrt{\pi}} \exp(-r^2 t^2) = \delta$ en D']. Es decir, Kirchoff toma como F la δ . Posteriormente, mediante una serie de manipulaciones realizadas como si F fuera una función para la que valen todas las reglas formales del cálculo, Kirchoff es capaz de obtener una expresión concreta para u (¡en la que no aparece F !) que es una transcripción matemática del principio de Huygens: la amplitud de la onda en un punto \bar{x}' de G se obtiene como superposición de ondas propagadas desde la frontera, con velocidad a .

A pesar de que Kirchoff es el primero en introducir explícitamente una función con las propiedades de la δ , hay que hacer notar que esta función singular aparece más o menos maquillada desde el mismo inicio de la teoría de lo que hoy se conoce como «funciones de Green» o «soluciones fundamentales de un operador diferencial en un punto»: en su revolucionario *Ensayo sobre la aplicación del Análisis Matemático a la teoría de la Electricidad y el Magnetismo* (1828), Green obtiene una fórmula para determinar el potencial eléctrico V en un cuerpo en el vacío acotado por conductores con potenciales dados, mostrando que el correspondiente problema matemático consiste en resolver la ecuación de Laplace $\Delta V = 0$ en el dominio G considerado, con ciertas condiciones de contorno en ∂G . Para ello utiliza de forma esencial, amén del teorema de su nombre, la existencia de una función de dos variables en G (que luego Riemann llamó función de Green asociada al operador

y al dominio G) que, en lenguaje moderno, satisface la ecuación

$$\Delta_x U(\bar{x}, \bar{x}') = -4\pi\delta(\bar{x} - \bar{x}')$$

y que como función de \bar{x} es de clase 2 en $G - \{\bar{x}'\}$ y se anula en la frontera de G . Green dio por supuesto la existencia de tal función, pues sabía que físicamente describía el potencial de una carga en \bar{x}' cuando G es una superficie conductora metálica. Más tarde, uno de los grandes problemas de la teoría fue probar la existencia de la función de Green para diferentes ecuaciones y diferentes dominios. En el Vol. I, pág. 351 y sigs. del libro *Methods of Mathematical Physics*, de Courant-Hilbert (Interscience, 1953), puede verse una excelente introducción heurística a la función de Green en el contexto del problema de Sturm-Liouville general.

El ingeniero eléctrico Oliver Heaviside desarrolló, a finales del siglo XIX, un cálculo operacional, de difícil justificación matemática, basado en razonamientos experimentales, y que fue ampliamente utilizado en el primer tercio de este siglo. La introducción que hace de la δ es típica: aparece como una «función impulso», derivada de la función $H(t)$, que vale 0 si $t < 0$ y 1 si $t > 0$. El razonamiento de Heaviside es el siguiente:

Si Q es cualquier función del tiempo, Q' es su velocidad de crecimiento. Por tanto, si, como en nuestro caso, H es 0 antes de $t = 0$ y constante después, H' es cero, excepto en $t = 0$, donde es infinita. Pero su suma total es 1. Esto es, H' es una función de t enteramente concentrada en $t = 0$, de suma total 1. Por así decirlo, es una función impulsiva.

La ingeniería eléctrica fue, junto con la mecánica cuántica, el más importante campo de aplicación de la función δ antes de la teoría de distribuciones. La amplia difusión del cálculo operacional de Heaviside condujo inevitablemente a la aparición de manipulaciones formales con la δ . Sin embargo, en muchos casos la δ aparecía simplemente como H' y desaparecía en el resultado. En los intentos posteriores de rigORIZAR el cálculo operacional, la actitud hacia la δ varió de unos autores a otros. Así, entre los dos mayores responsables del uso generalizado de la transformación de Laplace como justificación del cálculo operacional, la actitud es radicalmente distinta. El

rigorista Doetsch la consideró «ilegítima» y la ignoró en sus primeros trabajos. El ingeniero eléctrico van der Pol, la usó limitadamente y dio fórmulas de transformación de δ y sus derivadas. Aunque no dio una definición explícita, describió la δ por las siguientes propiedades:

$$\int_{-\infty}^{\infty} \delta(x) dx = 1 \quad \text{y} \quad \delta(x) = 0, \quad \text{excepto en } 0, \text{ donde es } \infty$$

Además, puso de manifiesto explícitamente la propiedad

$$\int_{-\infty}^{\infty} f(x) \delta^{(n)}(x) dx = (-1)^n f^{(n)}(0)$$

La función δ se llama «función de Dirac» por el amplio uso que de la misma hizo el físico P. A. M. Dirac en su intento de unificar la mecánica matricial y la ondulatoria. Su teoría apareció en 1926 en un artículo publicado en los Proceedings de la Royal Society de Londres, y fue desarrollada posteriormente en su libro *Los Principios de la Mecánica Cuántica* (1930). Dirac representó los estados de un sistema mecánico por vectores ψ y los observables por operadores lineales. En sus razonamientos, Dirac suponía que en un espacio vectorial era posible elegir siempre una «base» finita o infinita numerable (ψ_p) , de modo que los vectores ψ podían representarse por coordenadas, y los operadores α por matrices (α_{pq}) . De esta manera, escribía

$$\psi = \sum_p a_p \psi_p \quad \alpha(\psi_q) = \sum_p \alpha_{pq} \psi_p$$

Las representaciones físicamente interesantes son aquellas en las que la «base» está formada por autovalores de un operador correspondiente a algún observable. Sin embargo, en este caso el requerimiento de numerabilidad de la «base» no se cumple en general, ya que los operadores más importantes desde el punto de vista físico tienen un espectro continuo y, por lo tanto,

... el número total de estados independientes es infinito, e igual al número de puntos de una línea. Así pues, la condición que expresa que cualquier ψ es función lineal de los ψ' s fundamentales debe escribirse como una integral, en lugar de una suma:

$$\psi = \int a_p \psi_p \quad (2.15)$$

Pero entonces, cuando, por ejemplo, se quiere expresar en la forma (2.15) uno de los autovectores surge un problema, pues se tiene que escribir

$$\psi_q = \int \delta(p - q) \psi_p dp$$

«donde la función impropia δ está definida por $\int_{-\infty}^{\infty} \delta(x) dx = 1$ y $\delta(x) = 0$ para $x \neq 0$ »

La derivada la representa Dirac por

$$-\frac{\partial \psi_q}{\partial q} = \int_{-\infty}^{\infty} \psi_p \delta'(p - q) dp$$

También la representación de operadores en el caso de un espectro continuo requería para Dirac el uso de la δ , escribiendo

$$\alpha(\psi_q) = \int \psi_p \alpha_{pq} dp \quad (2.16)$$

por lo que, obviamente, el operador de multiplicación por una constante c debía representarse por el núcleo $\alpha_{pq} = c\delta(p - q)$, que Dirac interpretaba como el análogo continuo de la matriz identidad $I = (\delta_{pq})$. Esta fue seguramente la razón por la que Dirac dio a su función impropia el nombre de δ , no por Dirac, sino por Kronecker.

Dirac dio también «ciertas propiedades elementales de la función δ que se deducen de la definición, o al menos no son inconsistentes». Entre ellas, estaban:

$$\delta(-x) = \delta(x) \quad x\delta(x) = 0$$

$$-\delta'(x) = \delta'(-x) \quad x\delta'(x) = -\delta(x)$$

$$\int_{-\infty}^{\infty} f(x)\delta(x) dx = f(0) \quad \int_{-\infty}^{\infty} f(x)\delta(x - a) dx = f(a)$$

$$\int_{-\infty}^{\infty} \delta(a - x)\delta(x - b) dx = \delta(a - b) \quad \int_{-\infty}^{\infty} f(x)\delta'(x) dx = -f'(a)$$

De la derivada, dijo que, «desde luego, es una función aún más discontinua e impropia que la misma δ ».

En posteriores ediciones de su obra fue añadiendo nuevas propiedades, como

$$\delta(ax) = a^{-1}\delta(x) \qquad \frac{d}{dx} \log(x) = \frac{1}{x} - i\pi\delta(x)$$

que prueban que Dirac era un hábil manipulador de la δ . Notemos de pasada que la fórmula (2.16), según la cual todo operador puede representarse por un núcleo, indica de forma imprecisa lo que hoy se conoce como el *Teorema de los Núcleos* de L. Schwartz. En la tercera edición de su libro, Dirac menciona la «forma alternativa de definir la δ » como la derivada de la función de Heaviside.

El libro de Dirac se convirtió rápidamente en un clásico y con él se generalizó el uso de la función δ entre los físicos. En 1926, cuando apareció el primer artículo de Dirac, un grupo de matemáticos y físicos (entre ellos, Hilbert, Von Neumann y Nordheim) estaban explorando ideas similares. Supusieron, como Dirac, que los operadores asociados a observables físicos eran operadores integrales

$$Tf(x) = \int \varphi(x, y)f(y) dy \qquad (2.17)$$

y de esta manera su desarrollo de la teoría cuántica llevó la función δ a manos de los matemáticos. Las contradicciones con las nociones habituales de función, derivada, etc., hicieron dificultosa la obtención de una teoría rigurosa. Sin embargo, uno de los miembros del grupo, J. Von Neumann, formuló en 1927 una nueva y revolucionaria base matemática de la mecánica cuántica. En la introducción de su libro *Fundamentos Matemáticos de la Mecánica Cuántica* (1932), Von Neumann probó que la representación de la identidad en la forma (2.17) requeriría que φ tuviera las propiedades de la δ , que eran incompatibles con la integración Riemann y Lebesgue, por lo que había que buscar otra vía alternativa. La idea de Von Neumann fue describir la Mecánica Cuántica en un espacio de Hilbert (que él axiomatizó) separable, y los observables como operadores no acotados sobre él. Von Neumann desarrolló la teoría espectral para operadores no acotados, reemplazando las representaciones de Dirac en términos de integrales de autofunciones, por integrales del tipo de Stieltjes, como

$$\psi = \int_{\sigma} dE_{\lambda}$$

donde E_λ es una familia de proyecciones. Así pudo Von Neumann evitar el uso de la función δ ; pero muchos físicos continuaron usando los métodos de Dirac.

El uso de la función δ y otras funciones singulares se fue incrementando, a pesar de la falta de rigor, lo que originó diversos intentos de rigorización. Recordemos las diversas definiciones de δ :

a) $\delta(x) = \frac{d}{dx}H(x)$.

b) $\delta(x) = \lim_{n \rightarrow \infty} f_n(x)$ para una sucesión adecuada (f_n) .

c) $\delta(x) = 0$ para $x \neq 0$ y $\int_{-\infty}^{\infty} \delta = 1$

d) $\int_{-\infty}^{\infty} \delta(x-a)f(x) dx = f(a)$ o $\int_{-\infty}^{\infty} \delta(x)f(x) = f(0)$

La propiedad d), la más próxima a la definición funcional, fue utilizada en todas las aplicaciones de la δ por la mayoría de los autores. Dirac, en la tercera edición de sus *Principios...*, se acerca mucho a la definición funcional, cuando dice:

... aunque una función impropia no tiene un valor bien definido, cuando aparece como factor en un integrando, la integral sí que tiene un valor bien definido. En la teoría cuántica, siempre que aparece una función impropia, en último término será usada dentro de una integral. Por tanto, sería posible desarrollar la teoría de modo que las funciones impropias aparecieran únicamente como integrandos, y así se podrían eliminar. El uso de funciones impropias, por tanto, no supone ninguna pérdida de rigor en la teoría, sino que simplemente es una notación conveniente...

y antes, en la primera edición de 1930, había escrito:

Cualquier ecuación en la que aparece la función δ puede convertirse en otra equivalente, pero generalmente más complicada, en la que δ no aparece.

No nos dice Dirac la forma de hacer la transcripción, pero la presentación intuitiva que hace de la δ como límite impropio de funciones ordinarias, sugiere que el método riguroso consistiría en efectuar los cálculos con las aproximaciones de δ y luego pasar al límite en el resultado final. Esta justificación de la δ y otras funciones singulares, indicando un posible argumento riguroso alternativo (¡pero sin darlo!) se encuentra también en otros autores, pero el método usado frecuentemente por los matemáticos para tratar fenómenos que, por su descripción física intuitiva, comportaban el uso de la δ , fue dar sólo el método riguroso, o sin hacer referencia al argumento intuitivo que había detrás. Parece haber evidencias claras de que muchos matemáticos desarrollaron o utilizaron distintas teorías como una forma de rigORIZAR los argumentos que empleaban la δ , pero no para fundamentar rigurosamente la función δ misma; probablemente porque estaban tan convencidos de su ilegitimidad que ni siquiera lo intentaban. Sin embargo, hubo algunos intentos de rigORIZACIÓN de la δ antes de la teoría de distribuciones.

Uno de ellos fue debido a J. J. Smith, de la General Electric Co., que en un largo artículo publicado en 1926 trató de rigORIZAR no sólo la δ , sino todo el cálculo operacional de Heaviside por medio de una cierta teoría de H -funciones. A pesar de sus afirmaciones, realmente su trabajo no reúne los mínimos requisitos de rigor, pero tiene algunos puntos interesantes. Así, opta por un desarrollo experimental de la teoría, que le lleva a definir las ideas más complejas de la misma (como H -diferenciación, etc.) antes de los conceptos más simples, como el de H -límite, por medio de ejemplos. Así, la H -derivada de una H -función f en $[a, b]$ (que puede ser multiforme) se define como el H -límite de

$$\frac{f(x + \Delta x_1) - f(x - \Delta x_2)}{\Delta x_1 + \Delta x_2}$$

cuando Δx_1 y Δx_2 tienden a 0 y $\Delta x_1/\Delta x_2 = p$, con p una constante arbitraria, pero fija, $0 \leq p \leq \infty$. Todos los valores obtenidos para diferentes valores de p son valores de la H -derivada. Como ejemplo más importante, calcula las derivadas de la H -función f definida por $f(x) = (1 - a)x$ para

$x \leq a$, $f(x) = (1-x)a$ para $x \geq a$. La primera derivada es

$$f'(x) = \begin{cases} 1-a & \text{si } x < a \\ -a & \text{si } x > a \\ \frac{1}{p+1} - a & \text{si } x = a \quad (0 \leq p \leq \infty) \end{cases}$$

y las derivadas sucesivas son

$$f''(x) = \begin{cases} 0 & \text{si } x < a \\ 0 & \text{si } x > a \\ -k & \text{si } x = a \quad (0 \leq k \leq \infty) \end{cases}$$

$$f^{(n)}(x) = \begin{cases} 0 & \text{si } x < a \text{ ó } x > a \\ -k & \text{si } x = a \quad (-\infty \leq k \leq \infty) \end{cases} \quad (n \geq 3)$$

En general, sus definiciones son tan oscuras que es difícil tener una idea de lo que es una H -función, un H -límite, etc., y sólo los ejemplos que da permiten intuir alguna de sus nociones. En todo caso, es muy dudosa la utilidad de la teoría de Smith, pues los ingenieros eléctricos sabían perfectamente bien cómo operar intuitivamente con el cálculo de Heaviside, y los matemáticos no podían aceptar sus razonamientos como una teoría rigurosa y coherente.

Un intento mucho más serio fue realizado por el matemático inglés E. W. Sumpner en 1931. Su teoría se basa en el uso de infinitesimales e infinitos, cuidadosamente elegidos. Así, dice:

Sea c un infinitesimal positivo, que puede ser tan pequeño como queramos, pero que, una vez elegido, debe tratarse como una cantidad fija.

Sea n un entero positivo, elegido después de c , tan grande que nc^2 es infinito.

Sea w un infinito positivo, pero de orden inferior a $1/c$, y tal que cw es infinitesimal

Con estos números no standard, define la función de Heaviside H o

unidad como

$$H(t) = \frac{t^c}{c!} \left[\frac{1}{2} + \frac{1}{\pi} Si(nct) \right]$$

donde

$$Si(x) = \int_0^x \frac{\text{sen}(u)}{u} du$$

y después la función impulso

$$pH(t) = \frac{d}{dt} H(t)$$

Sumpner prueba que $H(t)$, para todo t de $[-w, w]$ real, tiene las propiedades usuales de la función de Heaviside, por ejemplo: «... para todo a de $+c$ a $+w$, H será la unidad, excepto una cantidad muy infinitesimal comparada con c » y (por el término $t^c/c!$), «para todo t de $-w$ a 0 , $H(t) = 0$, excepto por una cantidad altamente infinitesimal comparada con c .»

Examinada al microscopio, $H(t)$ es una «función de onda de longitud de onda constante (infinitamente pequeña), que fluctúa entre dos valores: la unidad para $t > 0$ y 0 para $t < 0$ ». Además, mirada de esta forma, $H(t)$ es, según Sumpner, tan regular que se puede derivar, y su derivada es continua: «Su derivada pH es una función de onda de la misma longitud de onda, que fluctúa alrededor de 0 para todos los valores de t .»

Sumpner tiene que postular que $t^c/c!$ es 0 para $t < 0$, pero excepto este punto, Sumpner trabaja muy cuidadosamente para obtener las anteriores propiedades de H y pH a partir de los diferentes infinitesimales e infinitos elegidos. Bien es verdad, que Sumpner no da ninguna prueba de que puedan definirse infinitesimales e infinitos de modo que, junto a los reales, puedan usarse con ellos las reglas ordinarias del álgebra tal como él hace. Sin embargo, sus ideas señalan una rigorización de la δ en términos de infinitesimales e infinitos que fueron desarrolladas mucho más tarde en el marco del análisis no standard (Langwitz, 1958; Robinson, 1961)

2. La teoría de distribuciones de L. Schwartz

Laurent Schwartz (nacido en 1915) se graduó en Matemáticas en la Ecole Normale Supérieure en 1937. Después del servicio militar, comenzó sus trabajos de investigación en 1940, en la Facultad de Ciencias de Estrasburgo,

que se había trasladado a Clermont-Ferrand durante la ocupación alemana. Allí se doctoró en 1943. En 1944 se trasladó a Grenoble y en 1945 fue profesor en la Facultad de Ciencias de Nancy. En 1953 se incorporó al núcleo universitario de París, primero como profesor en la facultad de ciencias (1953-1959) y después en la Ecole Polytechnique. Ha sido miembro del grupo Bourbaki y ha recibido numerosos premios y distinciones, entre los que destaca la Medalla Field, que recibió en 1950 por su creación de la teoría de distribuciones.

En la Ecole Normale, Schwartz tuvo como profesores, entre otros, a Leray, P. Levy y J. Hadamard que, como sabemos, habían trabajado en el problema de soluciones generalizadas de ecuaciones diferenciales. En su tesis, consagrada al estudio de sumas de exponenciales, Schwartz utiliza métodos de análisis funcional para resolver problemas de aproximación de tipo clásico. Esta idea de aplicar el análisis funcional a los problemas clásicos es una constante en el trabajo de Schwartz e influyó decisivamente en la creación de la teoría de distribuciones. Parece que la influencia del grupo Bourbaki, representado en Clermont-Ferrand por J. Dieudonné, H. Cartan, A. Weil y otros, no fue ajena a las preferencias de Schwartz por esta línea de investigación.

En Grenoble, Schwartz estaba aislado y, como mero ejercicio, generalizó la teoría de dualidad de espacios de Banach para espacios de Frechet, definiendo en estos espacios el dual fuerte y caracterizando la reflexividad. Los ejemplos más importantes que tenía a su disposición Schwartz de este tipo de espacios eran los espacios de funciones holomorfas y el espacio E de las funciones de clase infinito, con la convergencia uniforme de las funciones y sus derivadas sobre cada compacto. Este trabajo de Schwartz en análisis funcional abstracto no se publicó nunca, pero tuvo una influencia destacada en la creación de la teoría de distribuciones:

«Este trabajo abstracto previo sería la clave de la teoría de distribuciones. Es esta formación anterior la que hizo que el “descubrimiento” de las distribuciones fuera casi instantáneo al comienzo de 1945» (Autobiografía de L. Schwartz en la Academia de Ciencias).

El punto de partida para la creación por Schwartz de la teoría de distribuciones fue su trabajo sobre soluciones generalizadas de ecuaciones en

derivadas parciales, motivado por un artículo de 1944, *Sur certaines familles non fondamentales de fonctions continues*, en el que generalizaba un teorema de Cartan sobre caracterización de funciones que engendraban conjuntos fundamentales (*e.d.*, tales que las combinaciones lineales fueran densas) en el espacio de funciones continuas sobre \mathbb{R}^n con la topología de la convergencia uniforme sobre compactos. El teorema probado por Schwartz es el siguiente:

Para que el sistema de funciones continuas

$$U(\lambda x_1 + \xi_1, \lambda x_2 + \xi_2, \dots, \lambda x_n + \xi_n)$$

de las variables x_1, \dots, x_n sea no fundamental sobre un compacto K del espacio de n dimensiones, cuando λ y ξ_1, \dots, ξ_n toman todos los valores reales posibles, es necesario (y suficiente si el interior de K no es vacío) que U sea solución generalizada de al menos una ecuación en derivadas parciales del tipo

$$0 = \sum_{p_1 + \dots + p_n = m} a_{p_1 p_2 \dots p_n} \frac{\partial^m U}{(\partial x_1)^{p_1} \dots (\partial x_n)^{p_n}} \quad (2.18)$$

donde $a_{p_1 \dots p_n}$ son constantes.

Schwartz definió U como solución generalizada de la ecuación (2.18) si era límite uniforme de soluciones ordinarias sobre cada compacto de \mathbb{R}^n . Así pues, Schwartz, como Sobolev, partió de una definición secuencial de solución generalizada, aunque el tipo de convergencia difería de la convergencia en L^1 usada por Sobolev. También observó Schwartz que la ecuación de ondas

$$\frac{\partial^2 U}{\partial x^2} - \frac{\partial^2 U}{\partial y^2} = 0$$

tenía las soluciones generalizadas $f(x+y) + g(x-y)$ para todo par de funciones continuas f y g . Posteriormente probó que todas las soluciones generalizadas de la ecuación de Laplace son soluciones ordinarias.

Pocos días después de escribir este artículo, Schwartz atacó el problema de obtener una definición mejor de solución generalizada. Fijó su atención en el hecho de que la solución generalizada obtenida en la demostración del teorema actuaba como un operador de convolución que transformaba una

función infinitamente diferenciable con soporte compacto en una función \mathcal{C}^∞ . Esto le llevó a introducir un nuevo objeto, que llamó «operador de convolución»: lo definió como un operador lineal continuo de D (espacio de las funciones infinitamente diferenciables con soporte compacto, con la noción de convergencia secuencial idéntica a la de Sobolev) en E (espacio de las funciones de clase \mathcal{C}^∞ , con la convergencia uniforme sobre compactos de todas las derivadas), con la propiedad de que

$$\forall \varphi, \psi \in D \quad T \cdot (\varphi * \psi) = (T \cdot \varphi) * \psi$$

(la continuidad tomada en sentido secuencial).

Schwartz observó que toda función continua f podía identificarse con el operador

$$\varphi \rightarrow f * \delta \quad (\varphi \in D)$$

y que la función δ podía interpretarse como el operador

$$\varphi \rightarrow \delta \cdot \varphi = \varphi \quad (\varphi \in D)$$

Como el punto de partida para la definición de operadores de convolución había sido la teoría de soluciones generalizadas de ecuaciones diferenciales, Schwartz abordó, naturalmente, la cuestión de definir la derivada de un operador, lo que hizo por la fórmula

$$\left(\frac{\partial}{\partial x_i} T \right) \cdot \varphi = T \cdot \frac{\partial \varphi}{\partial x_i}$$

que generaliza la fórmula usual de derivación de un producto de convolución. Definió también la convolución de dos operadores, de modo obvio, y el producto de un operador por una función de D , aunque aquí se encontró con dificultades. Para resolverlas, tuvo que probar que todo operador de convolución era localmente la derivada u -sima de (el operador asociado a) una función continua, para lo que tuvo que echar mano de sus conocimientos del espacio $D(K)$ y la teoría por él desarrollada en espacios de Frechet.

Schwartz desarrolló estos conceptos y varios teoremas sobre sus operadores de convolución durante una noche de octubre de 1944. En los próximos seis meses continuó trabajando sobre el mismo tema, obteniendo nuevos

teoremas. Alrededor de febrero de 1945 comenzó a desarrollar una teoría de transformadas de Fourier para estos operadores, pero se encontró con grandes dificultades. Intentó resolverlas durante varios meses. Por fin, un día, probablemente en abril de 1945, de repente se dio cuenta que estos problemas podían resolverse fácilmente si definía las funciones generalizadas no como operadores, sino como funcionales sobre D , que llamó distribuciones. Según sus propias afirmaciones, hubo dos hechos que le pudieron sugerir la definición «real» de distribución:

- 1) Su trabajo anterior sobre la dualidad de espacios de Frechet le había suministrado una teoría abstracta de los operadores lineales continuos sobre E , sin una representación «concreta» con la que calcular.
- 2) Sabía, por sus contactos con Cartan y el grupo Bourbaki, que las medidas, y en especial la medida δ , podían representarse como funcionales.

Durante la primavera de 1945, Schwartz desarrolló su nueva teoría de distribuciones. Si en la teoría de operadores de convolución su trabajo sobre análisis funcional abstracto había jugado un papel importante, en la nueva teoría ocupaba una posición mucho más importante aún, y ayudó considerablemente a su desarrollo. Sin embargo, la influencia era recíproca. Por ejemplo, el concepto de límite inductivo de espacios de Frechet tiene sus raíces en la teoría de distribuciones. Schwartz sabía muy bien que la convergencia que había definido en D no podía obtenerse a partir de una topología de Frechet en ese espacio. Por tanto, él no trató de definir un sistema de entornos en D , sino sólo un sistema de conjuntos acotados que le permitieran (por dualidad) definir una topología en D' . Cuando Dieudonné conoció la descripción topológica de D hecha por Schwartz, la relacionó con la teoría abstracta de límite inductivo de espacios topológicos. En 1949, Dieudonné y Schwartz escribieron un artículo conjunto sobre el tema, en el que probaban los principales teoremas que Schwartz había usado en el caso concreto de los espacios D y D' en un marco general, y a partir de este trabajo se produjo un gran impulso en la teoría general de espacios localmente convexos.

Gracias a su trabajo previo en análisis funcional. Schwartz desarrolló su nueva teoría de distribuciones tan rápidamente que pudo dar un curso sobre

ella en el invierno 1945-1946 en el Collège de France de Paris.

Aparte de sus conocimientos sobre la teoría general de soluciones generalizadas de ecuaciones en derivadas parciales y la teoría de la función δ , Schwartz no conocía en 1944 el resto de lo que hemos visto como prehistoria de la teoría de distribuciones: los trabajos de Sobolev, el cálculo de Heaviside, las funciones impropias usadas por los físicos en la mecánica cuántica, los funcionales analíticos de Fantappié o los trabajos de Bochner y Carleman sobre la transformación de Fourier generalizada.

A partir de su curso en el Collège de France, tuvo noticia de alguno de estos desarrollos. En particular, los ingenieros le hablaron del cálculo operacional y le animaron a continuar su trabajo, especialmente con el tratamiento de las transformaciones de Laplace y Fourier y la convolución. Además, le pidieron que hiciera una monografía sobre el tema, tan elemental que todos ellos pudieran entenderla.

Más tarde, Schwartz tuvo conocimiento a través de sus colegas de los trabajos previos sobre el tema. Sin embargo, para entonces había desarrollado tanto su teoría, que no encontró en ellos nada nuevo que le inspirara nuevos conceptos o técnicas.

Schwartz escribió cuatro artículos sobre la teoría de distribuciones antes de la publicación de su monografía *Theorie des Distributions* en 1950-1951. En el primero y cuarto se encuentra un resumen general de la teoría, con énfasis en los aspectos matemáticos en el primero y más en las aplicaciones en el cuarto. El segundo y tercer artículos tratan fundamentalmente de la transformada de Fourier, definida para funciones de crecimiento lento por la fórmula

$$F(f)(x) = \lim_{A, B \rightarrow \infty} \int_{-A}^B f(t) \exp(2\pi itx) dt$$

donde el límite se toma en D' . Sólo más tarde apareció la idea de la transformación de Fourier usual en el espacio S . Ciertamente, en junio de 1947 Schwartz ya había descubierto las distribuciones temperadas, pues en esa fecha habló sobre *Teoría de distribuciones y transformación de Fourier* en un congreso de análisis armónico celebrado en Nancy.

La monografía de Schwartz, publicada en dos volúmenes en 1950-1951, se convirtió inmediatamente en la obra de referencia standard sobre la teoría

de distribuciones. En las sucesivas ediciones la obra se ha ido manteniendo al día, de modo que aún hoy es una de las obras más completas sobre el tema.

Schwartz continuó trabajando en la teoría de distribuciones después de la publicación de su monografía. El primer gran éxito de la teoría no contenido en su texto es la formulación y demostración rigurosa del teorema de los núcleos.

Bibliografía

- (1) J. DIEUDONNE: *History of functional Analysis*, North Holland, 1981.
- (2) J. HORVATH: *Topological vector spaces and distributions*, Addison Wesley, New York, 1966.
- (3) J. HORVATH: *An introduction to distributions*, Amer. Math. Monthly (1970), 227-240.
- (4) J. LÜTZEN: *The prehistory of the Theory of Distributions*, Springer, 1982.
- (5) A. F. MONNA: *Functional Analysis in Historical Perspectiva*, Uthecht, 1973.
- (6) L. SCHWARTZ: *Theorie des distributions*, Vols. I y II, Hormann, París, 1950/51.
- (7) F. TREVES: *Topological vector spaces, distributions and kernels*, Academic Press, 1967.
- (8) F. TREVES: *Applications of distributions to PDE theory*, Amer. Math. Monthly (1970),