GENERALIZED FOURIER TRANSFORMATIONS: THE WORK OF BOCHNER AND CARLEMAN VIEWED IN THE LIGHT OF THE THEORIES OF SCHWARTZ AND SATO

CHRISTER O. KISELMAN

Uppsala University, P. O. Box 480, SE-751 06 Uppsala, Sweden

E-mail: kiselman@math.uu.se

Salomon Bochner (1899–1982) and Torsten Carleman (1892–1949) presented generalizations of the Fourier transform of functions defined on the real axis. While Bochner's idea was to define the Fourier transform as a (formal) derivative of high order of a function, Carleman, in his lectures in 1935, defined his Fourier transform as a pair of holomorphic functions and thus foreshadowed the definition of hyperfunctions. Jesper Lützen, in his book on the prehistory of the theory of distributions, stated two problems in connection with Carleman's generalization of the Fourier transform. In the article these problems are discussed and solved.

Contents:

1 Introduction

In order to define in an elementary way the Fourier transform of a function we need to assume that it decays at infinity at a certain rate. Already long ago mathematicians felt a need to extend the definition to more general functions. In this paper I shall review some of the attempts in that direction: I shall explain the generalizations presented by Salomon Bochner (1899–1982) and Torsten Carleman (1892–1949) and try to put their ideas into the framework of the later theories developed by Laurent Schwartz and Mikio Sato.

In his book on the prehistory of the theory of distributions, Jesper Lützen [1982] gives an account of various methods to extend the definition of the Fourier transformation. This paper has its origin in a conversation with

Anders Öberg, who pointed out to me that Lützen had left open two questions. I shall try to answer these here. Thus this paper is in a way an historical survey, but not exclusively.

2 Bochner

In his book *Vorlesungen über Fouriersche Integrale* [1932], translated as *Lectures on Fourier Integrals* [1959], Salomon Bochner extended the definition of the Fourier transform to functions such that $f(x)/(1+|x|)^k$ is integrable for some number k. The usual Fourier transform of f is defined as

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbf{R}} f(x)e^{-ix\xi}dx, \qquad \xi \in \mathbf{R},$$

provided the integral has a sense; e.g., if f is integrable in the sense of Lebesgue. If f is sufficiently small near the origin we may form

$$g_k(\xi) = \int_{\mathbf{R}} \frac{f(x)e^{-ix\xi}}{(-ix)^k}dx, \qquad \xi \in \mathbf{R},$$

and this integral now has a sense if $f(x)/x^k$ is integrable. If both f and f/x^k are integrable, then the k^{th} derivative of g_k is equal to $\hat{f}$. So g_k is a k^{th} primitive function of $\hat{f}$ in the classical sense. This is the starting point of Bochner's investigation.

To overcome the somewhat arbitrary assumption that f is small near the origin, Bochner [1932:112, 1959:140] adjusted the integrand by using the Taylor expansion of the exponential function and defined

$$E(\alpha, k) \overset{k}{\asymp} \frac{1}{2\pi}\int_{\mathbf{R}} f(x)\frac{e^{-i\alpha x} - L_k(\alpha, x)}{(-ix)^k}dx, \qquad \alpha \in \mathbf{R}, \quad k \in \mathbf{N},$$

where

$$L_k(\alpha, x) = \begin{cases} \displaystyle\sum_{j=0}^{k-1} \frac{(-i\alpha x)^j}{j!}, & |x| \leqslant 1, \\ 0, & |x| > 1. \end{cases}$$

The symbol $\overset{k}{\asymp}$ means that the difference between the two sides is a polynomial of degree less than k. Thus E is undetermined, but its k^{th} derivative is not influenced by this ambiguity. Bochner's Fourier transform is this k^{th} derivative, a formal object. Calculations are done on E, not on its derivative $E^{(k)}$.

In his review of Schwartz [1950, 1951], Bochner [1952:79–80] remarks that any distribution in $\mathcal{D}'(\mathbf{R}^n)$ agrees in a given bounded domain with the k^{th} derivative of a continuous function for some sufficiently large k. Thus the Fourier transforms that Bochner constructs are locally not less general than distributions. Bochner's review portrays the theory of distributions as not going much beyond what he himself has presented in his book [1932]; "it would not be easy to decide what the general innovations in the present work are, analytical or even conceptual" [1952:85]. Later generations of mathematicians have been more appreciative.

In several papers, starting in 1954, Sebastio e Silva developed the idea of defining distributions as derivatives of functions. He used an axiomatic approach; see, e.g., [1964].

3 Streamlining Bochner's definition

In particular, if f vanishes for $|x| \leqslant 1$, then the definition of $E(\alpha, k)$ simplifes to

$$E(\alpha, k) \overset{k}{\asymp} \frac{1}{2\pi} \int_{\mathbf{R}} f(x) \frac{e^{-i\alpha x}}{(-ix)^k} dx.$$

We may therefore split any function f into two, $f = f_0 + f_1$, where $f_0(x) = 0$ for $|x| > 1$, $f_1(x) = 0$ for $|x| \leqslant 1$. For f_1 we then define the function $E(\alpha, k)$ as above without the need to use Taylor expansions, while f_0, a function of compact support, has a Fourier transform in the classical sense; the latter is an entire function of exponential type.

Another way to avoid the division by $(-ix)^k$ is to divide instead by some power of $1 + x^2$. This can easily be done in any number of variables, defining x^2 as an inner product, $x^2 = x \cdot x = \sum x_j^2$. Since the function $1 + x^2$ has no zeros, the mapping $f \mapsto (1 + x^2)f$ is a bijection. We may define

$$\mathcal{F}_{s,\varepsilon}(f)(\xi) = \int_{\mathbf{R}^n} f(x) e^{-ix\xi} (1 + \varepsilon^2 x^2)^{-s} dx, \qquad \xi \in \mathbf{R}^n, \quad s, \varepsilon \in \mathbf{R}.$$

Then the usual Fourier transform is obtained when s or ε vanishes:

$$\mathcal{F}_{0,\varepsilon} = \mathcal{F}_{s,0} = \mathcal{F}_{0,0} = \mathcal{F}.$$

If $f(x)(1+x^2)^{-t} \in L^1(\mathbf{R})$, then $\mathcal{F}_{s,\varepsilon}(f)$ is a bounded continuous function for all $s \geqslant t$ and all $\varepsilon \neq 0$. By applying a differential operator of order $2m$, $m \in \mathbf{N}$, we can lower the index s by m units:

$$(1 - \varepsilon^2 \Delta)^m \mathcal{F}_{s,\varepsilon}(f) = \mathcal{F}_{s-m,\varepsilon}(f),$$

where Δ is the Laplacian, $\Delta = \sum \partial^2/\partial x_j^2$. In particular, if $(1+x^2)^{-m} f$ is integrable, then $\mathcal{F}_{m,\varepsilon}(f)$ has a sense, and $(1-\varepsilon^2\Delta)^m \mathcal{F}_{m,\varepsilon}(f)$ is a generalized Fourier transform of f. This is a somewhat streamlined version of Bochner's idea of defining a primitive function of the Fourier transform of f. The word *primitive* must now be understood in terms of the differential operator $1-\varepsilon^2\Delta$. We shall come back to this idea in section 9.

The transform $\mathcal{F}_{s,\varepsilon}(f)$ depends continuously on (s,ε): for $\varphi \in \mathcal{S}(\mathbf{R}^n)$,

$$\mathcal{F}_{s,\varepsilon}(\varphi) \to \mathcal{F}_{s_0,\varepsilon_0}(\varphi) \qquad \text{as} \quad (s,\varepsilon) \to (s_0,\varepsilon_0) \in \mathbf{R}^2.$$

This is easy to prove using norms which define the topology of $\mathcal{S}(\mathbf{R}^n)$, either the norms defined in (5.4) below or those given in Proposition 9.3. By duality we get the same statement for $u \in \mathcal{S}'(\mathbf{R}^n)$. In particular, if one of s_0 and ε_0 is zero, then for all $\varphi \in \mathcal{S}(\mathbf{R}^n)$,

$$\mathcal{F}_{s,\varepsilon}(\varphi) \to \widehat{\varphi} \qquad \text{as} \quad (s,\varepsilon) \to (s_0,\varepsilon_0).$$

4 Carleman

In 1935, Torsten Carleman lectured on a generalization of the Fourier transformation at the Mittag-Leffler Institute near Stockholm, Sweden. His notes, however, were not published until nine years later. In his book [1944] he quotes Bochner [1932] and the work of Norbert Wiener. In June, 1947, Carleman participated in a CNRS meeting in Nancy organized by Szolem Mandelbrojt and presented his theory there; see Carleman [1949].

Carleman's approach is quite different from Bochner's and foreshadows the definition of hyperfunctions. In fact, in modern terminology, he defines the Fourier transform for a large class of hyperfunctions of one variable.

He remarks in the beginning that he will cover the case of functions which are integrable in Lebesgue's sense on each bounded interval and which satisfy the condition

$$\int_0^x |f(x)|dx = O(|x|^\kappa), \qquad x \to \pm\infty, \tag{4.1}$$

for some positive number κ. This condition is equivalent to the one imposed by Bochner, i.e., that $f(x)/(1+|x|)^k$ be integrable for some k. He then remarks that the usual Fourier transform of an integrable function can be written

$$g(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-izy} f(y)dy = g_1(z) - g_2(z),$$

where

$$g_1(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} e^{-izy} f(y)dy \quad \text{and} \quad g_2(z) = -\frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-izy} f(y)dy;$$

the function g_1 is well-defined and continuous for $\operatorname{Im} z \geqslant 0$ and is holomorphic in the open upper half plane; similarly with g_2 in the lower half plane. So the Fourier transform of f appears as the difference between the boundary values of two holomorphic functions g_1 and g_2, each defined and holomorphic in a half plane. In the case we are now considering, i.e., when f is integrable, the holomorphic function g_j has boundary values in a very elementary sense: it admits a continuous extension to the closed half plane, given by the same integral. We shall write $B(g_1, g_2)(x)$ for the difference $\lim_{y \to 0+}(g_1(x+iy) - g_2(x-iy))$.

Carleman then asks [1944:37] whether it is always possible to decompose a function defined on the real axis in this way, and whether this decomposition, if it exists, is unique. In the sequel he answers in the affirmative these two questions. Thus he shows that any measurable function of one variable satisfying (4.1) can be represented as a hyperfunction, and that the representation is unique in a reasonable sense, i.e., as could be expected, unique modulo an entire function—in fact, in view of the growth conditions he imposes, modulo a polynomial.

Carleman then proceeds to define the Fourier transform of a pair of functions. He considers functions f_1, f_2 defined respectively for $\operatorname{Im} z > 0$ and $\operatorname{Im} z < 0$ such that there exist nonnegative numbers α and β and, for all θ_0 in the interval $]0, \pi/2[$, a number $A(\theta_0)$ such that

$$|f_1(re^{i\theta})| < A(\theta_0)(r^\alpha + r^{-\beta}), \qquad r > 0, \quad \theta_0 < \theta < \pi - \theta_0, \tag{4.2}$$

and

$$|f_2(re^{i\theta})| < A(\theta_0)(r^\alpha + r^{-\beta}), \qquad r > 0, \quad \pi + \theta_0 < \theta < -\theta_0. \tag{4.3}$$

Let us call such a pair (f_1, f_2) a *Carleman pair of class* (α, β). He then defines [1944:48] another pair of holomorphic functions G, H by

$$G(z) = \frac{1}{\sqrt{2\pi}} \int_L e^{-izy} f_1(y)dy \quad \text{and} \quad H(z) = \frac{1}{\sqrt{2\pi}} \int_{L'} e^{-izy} f_2(y)dy. \tag{4.4}$$

Here L is a half line in the upper half plane issuing from the origin, and similarly with L' in the lower half plane. Thus, for a particular choice of L, the function G will be defined in a half plane $\{z; \operatorname{Im}(zy) < 0\}$; by letting L vary in the upper half plane, we will get a function defined in the complement of the positive real half axis; similarly H will be defined in the complement of the negative real half axis. In particular the difference $H - G$ is defined in $\mathbf{C} \smallsetminus \mathbf{R}$.

The integrals are well-defined if $\beta < 1$; if not, Carleman has to resort to the kind of trick that Bochner used: he defines the m^{th} derivatives as

$$G^{(m)}(z) = \frac{1}{\sqrt{2\pi}} \int_L e^{-izy}(-iy)^m f_1(y)dy \tag{4.5}$$

and

$$H^{(m)}(z) = \frac{1}{\sqrt{2\pi}} \int_{L'} e^{-izy}(-iy)^m f_2(y)dy, \tag{4.6}$$

so that G and H are determined only up to a polynomial of degree at most $m-1$. (This ambiguity will not affect the definition of the Fourier transform as we shall see.) The factor y^m attenuates the singularity at the origin. He chooses m such that $0 \leqslant \beta - m < 1$; in fact, any $m > \beta - 1$ will do. Next he defines

$$g_1(z) = H(z) - G(z) \text{ for } \operatorname{Im} z > 0 \text{ and } g_2(z) = H(z) - G(z) \text{ for } \operatorname{Im} z < 0,$$

and remarks that it is easily proved that g_1 and g_2 satisfy inequalities similar to those for f_1 and f_2,

$$|g_1(re^{i\theta})| < A_1(\theta_0)(r^{\alpha'} + r^{-\beta'}), \qquad \theta_0 < \theta < \pi - \theta_0,$$

and

$$|g_2(re^{i\theta})| < A_1(\theta_0)(r^{\alpha'} + r^{-\beta'}), \qquad -\pi + \theta_0 < \theta < -\theta_0,$$

where we may choose $\alpha' = \beta - 1 \geqslant -1$ and $\beta' = \alpha + 1 \geqslant 1$ if we assume that $\beta \neq 1, 2, 3, \ldots$. If $\beta = 1, 2, 3, \ldots$, there appears a logarithmic term in the estimate at infinity, and we may take α' as any number strictly larger than $\beta - 1$ while $\beta' = \alpha + 1$ as before. The interchange between α and β means that the growth of the f_j near the origin is reflected in the growth of the g_j at infinity and conversely. A convenient comparison function is $r^{\gamma-1/2} + r^{-\gamma-1/2}$, i.e., with $\alpha = \gamma - \frac{1}{2}$, $\beta = \gamma + \frac{1}{2}$. Then we achieve symmetry for $\gamma \neq \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$.

Thus Carleman's Fourier transform $\mathrm{CF}(f_1, f_2)$ of the pair $f = (f_1, f_2)$ is the pair (g_1, g_2); let us denote it by $g = S(f)$. He needs to interchange the g_j, so he defines a new operation T by $T(g) = (h_1, h_2)$, where $h_1(z) = \overline{g_2(\overline{z})}$ and $h_2(z) = \overline{g_1(\overline{z})}$. Carleman's version of Fourier's inversion formula [1944:49] then reads $(T \circ S \circ T \circ S)(f_1, f_2) = (f_1 + P, f_2 + P)$, where P is a polynomial; the latter does not influence the difference between the two functions. Since the calculation has to be done on the derivatives, the proof [1944:50–52] is a bit involved.

5 Schwartz

To extend the Fourier transformation Laurent Schwartz took the formula

$$\int_{\mathbf{R}^n} \hat{f}(\xi) g(\xi) d\xi = \int_{\mathbf{R}^n} f(x)\hat{g}(x) dx \tag{5.1}$$

as his starting point. The formula holds under quite general conditions and for most definitions of the Fourier transformation; no constant is needed. In particular it is true if both f and g are integrable on $\mathbf{R}^n$. To be precise, Schwartz [1966:231] defined

$$\mathcal{F}(f)(\xi) = \int_{\mathbf{R}^n} f(x) e^{-2i\pi x\cdot\xi} dx, \qquad \xi \in \mathbf{R}^n, \tag{5.2}$$

so that the inversion formula reads

$$f(x) = \int_{\mathbf{R}^n} \mathcal{F}(f)(\xi) e^{2i\pi x\cdot\xi} d\xi, \qquad x \in \mathbf{R}^n.$$

Formula (5.1) makes it natural to define Schwartz's Fourier transform $\mathrm{SF}(u)$ of a functional u by

$$\mathrm{SF}(u)(\varphi) = u(\widehat{\varphi}), \qquad \varphi \in \Phi, \tag{5.3}$$

Schwartz [1966:250]. In this way $\mathrm{SF}(u)$ is defined as a functional on a space of test functions Φ provided u itself is defined on the space $\widehat{\Phi}$ of all transforms of functions in Φ. Schwartz made this situation completely symmetric by defining Φ so that $\widehat{\Phi} = \Phi$. Since he wished Φ to contain $\mathcal{D}(\mathbf{R}^n)$, it must also contain $\mathcal{D}(\mathbf{R}^n) \cup \widehat{\mathcal{D}}(\mathbf{R}^n)$, and this is indeed the case for the Schwartz space $\mathcal{S}(\mathbf{R}^n)$. It is defined as the space of all smooth functions on $\mathbf{R}^n$ such that the norms

$$\varphi \mapsto \sup_{x\in\mathbf{R}^n} |x^\alpha \partial^\beta \varphi / \partial x^\beta|, \qquad \alpha, \beta \in \mathbf{N}^n, \tag{5.4}$$

are finite, and is equipped with the weakest topology making all these norms continuous. This makes the dual space smaller than the dual of $\mathcal{D}(\mathbf{R}^n)$; it is the well-known space $\mathcal{S}'(\mathbf{R}^n)$ of *temperate distributions,* strictly contained in $\mathcal{D}'(\mathbf{R}^n)$. These distributions, which are also known as *tempered distributions,* were called *distributions sphériques* in the beginning (see Schwartz [1949:3]), since they are the restrictions of the distributions defined on the n-dimensional sphere, which is identified with the one-point compactification $\mathbf{R}^n \cup \{\infty\}$.

Let us denote by $[f] \in \mathcal{D}'(\mathbf{R}^n)$ the distribution defined by a function $f \in L^1_{\mathrm{loc}}(\mathbf{R}^n)$, thus

$$[f](\varphi) = \int_{\mathbf{R}^n} f(x)\varphi(x) dx, \qquad f \in L^1_{\mathrm{loc}}(\mathbf{R}^n), \quad \varphi \in \mathcal{D}(\mathbf{R}^n).$$

Then, in view of (5.1), Schwartz's Fourier transform of $[f]$, defined by (5.3), is the distribution defined by the function $\hat{f}$:

$$\mathrm{SF}([f])(\varphi) = [f](\widehat{\varphi}) = [\hat{f}](\varphi), \qquad f \in L^1(\mathbf{R}^n), \quad \varphi \in \mathcal{D}(\mathbf{R}^n),$$

which means that SF extends the classical Fourier transformation to a larger class.

For any $c_1, c_2 \geqslant 0$, the mappings

$$\varphi \mapsto (1 + c_1 x^2)\varphi, \qquad \varphi \mapsto (1 - c_2\Delta)\varphi \tag{5.5}$$

are topological isomorphisms of the space $\mathcal{S}(\mathbf{R}^n)$ of Schwartz test functions. Here, again, we write x^2 for the inner product $x \cdot x = \sum x_j^2$ and Δ for the Laplacian $\sum \partial^2/\partial x_j^2$. They correspond to each other under the Fourier transformation in the sense that, for $c_1 = 4\pi^2 c_2$,

$$\mathcal{F}\big((1 + c_1 x^2)\varphi\big) = (1 - c_2\Delta)\widehat{\varphi} \quad \text{and} \quad \mathcal{F}\big((1 - c_2\Delta)\varphi\big) = (1 + c_1\xi^2)\widehat{\varphi}.$$

By duality the mappings (5.5) give rise to isomorphisms of $\mathcal{S}'(\mathbf{R}^n)$,

$$u \mapsto (1 + c_1 x^2)u, \qquad u \mapsto (1 - c_2\Delta)u. \tag{5.6}$$

To define not only the Fourier transform $\hat{u}(\xi)$ for all $\xi \in \mathbf{R}^n$ but more generally the Fourier–Laplace transform $\hat{u}(\zeta)$ for all $\zeta \in \mathbf{C}^n$ (at least as a functional), it would be desirable to find a space Φ such that

$$\mathcal{D} \subset \Phi \subset \mathcal{S}, \tag{5.7}$$

and such that

$$\text{for all } \varphi \in \Phi \text{ and all } \zeta \in \mathbf{C}^n, \int_{\mathbf{R}^n} e^{-i\zeta\cdot x}\varphi(x)dx \text{ is well defined.} \tag{5.8}$$

In 1961 I attempted to define the Fourier–Laplace transform in $\mathbf{C}^n$, inspired by Schwartz's definition of $\mathcal{S}(\mathbf{R}^n)$. I realized then that it is not possible to require (5.7) and (5.8) and keep the symmetry in the sense that $\widehat{\Phi} = \Phi$. Indeed, the function defined as $\psi(\xi) = \exp\big(-1/(1 - \|\xi\|^2)\big)$ for $\|\xi\| < 1$ and $\psi(\xi) = 0$ for $\|\xi\| \geqslant 1$ is in $\mathcal{D}(\mathbf{R}^n)$ but its Fourier transform $\varphi = \widehat{\psi}$ does not satisfy $\int_{\mathbf{R}^n} |e^{-i\zeta\cdot x}\varphi(x)|dx < \infty$ for any $\zeta \in \mathbf{C}^n \smallsetminus \mathbf{R}^n$.

By abandoning the requirement that $\widehat{\Phi}$ be equal to Φ, Gel′fand & Shilov [1953] found other interesting spaces of test functions. In particular they defined the Fourier transform of a distribution as a functional on $\widehat{\mathcal{D}}$. See also Ehrenpreis [1954, 1956]. Hörmander [1955] announced a very general theory of this nature.

In my work of 1961, I kept the symmetry $\widehat{\Phi} = \Phi$ and instead relaxed the condition (5.7) that Φ contain $\mathcal{D}$. In this work I defined a space $\mathcal{W}$ of test functions consisting of all entire functions φ on $\mathbf{C}^n$ such that the norms

$$\|\varphi\|_m = \sup_{\|\operatorname{Im} z\| \leqslant m} |\varphi(z)| e^{m\|\operatorname{Re} z\|}, \qquad m \in \mathbf{N},$$

are all finite; $\mathcal{W}$ is equipped with the topology defined by these norms. The Fourier transformation is an isomorphism of $\mathcal{W}$ onto itself, and the same is true of the dual space $\mathcal{W}'$. I studied the Fourier transformation and convolution in these spaces and developed several of their properties but my work was not published. Kelly McKennon independently discovered the same space and published his results in [1976]; he was kind enough to mention my work (McKennon [1976:178]).

Hörmander [1998] gives a full account of the ideas he presented in his short note [1955].

6 Sato

Mikio Sato presented his theory of hyperfunctions in [1958a,b,c, 1959, 1960]. Boundary values of holomorphic functions (without any growth condition) are the basic objects of his theory; in particular, all distributions in one variable are represented as the difference of such boundary values from the upper and lower half planes. The Fourier transform in one variable is defined for pairs of functions with infra-exponential growth, generalizing Carleman's conditions.

The theory of Fourier hyperfunctions in several variables is a theme outside the scope of this article. Let us only mention that it was developed by Kawai [1970a,b] and further developed by Morimoto [1973, 1978] and Saburi [1985].

7 On Carleman's Fourier transformation

In this section we shall comment on Carleman's theory and also show how Carleman pairs can be constructed.

Carleman's theory does not lend itself easily to calculations. For the pair of functions representing the Dirac measure placed at the origin one has to take $\beta = 1$ in (4.2), (4.3) and so has to use $m \geqslant 1$ in (4.5), (4.6). It is easy to calculate explicitly the functions G' and H' in (4.5), (4.6), and the jump in $H - G$ is found to be the constant $\frac{1}{\sqrt{2\pi}}$ as expected. For the Dirac measure placed at a point $a \neq 0$ we may take $\beta = 0$; it is, however, difficult to calculate G and H from (4.4), although their difference $H - G$ can be easily found. For $a > 0$, $H - G$ is 0 in the upper half plane and $-\frac{1}{\sqrt{2\pi}}e^{-iza}$ in the lower, so that the jump is $\frac{1}{\sqrt{2\pi}}e^{-iza}$ as we should expect. One even receives the impression that Carleman avoids examples and applications of his theory to simple generalized functions.

We note that if (f_1, f_2) is a Carleman pair of class (α, β), then the pair (zf_1, zf_2), which is of class $(\alpha + 1, \beta - 1)$, has a Carleman transform which is just i times the derivative of the transform of (f_1, f_2). Similarly, the derivative of (f_1, f_2), which is a pair of class $(\alpha - 1, \beta + 1)$, has a transform which is iz times the transform of (f_1, f_2). Thus the usual rules hold. However, Carleman does not mention these simple rules.

Bremermann & Durand [1961:241] write that Carleman's work is limited to L^2 and L^p functions. As we have seen, this is not so: the Carleman pairs are much more general. Along the rays through the origin Carleman assumes that the f_j have a temperate behavior (see (4.2) and (4.3)), but there is no restriction in the growth of $A_0(\theta_0)$ or $A_1(\theta_0)$ when θ_0 tends to zero. If we impose a temperate growth also on $A_0(\theta_0)$, then the condition can be written as $|f_j(z)| \leqslant C|\mathrm{Im}\, z|^{-\gamma}(|z|^\alpha + |z|^{-\beta})$, which means temperate growth both at infinity and at the real axis, and we get exactly the temperate distributions. Thus Carleman's classes are more general than the temperate distributions. On the other hand, the hyperfunctions are even more general, because for them we do not impose temperate growth at infinity or the origin.

To make the last remark clearer we may map the upper half plane onto the unit disk by a Möbius mapping, with the origin going to the point 1 and infinity going to -1, say. Then the temperate distributions correspond to pairs of holomorphic functions of temperate growth at the boundary of the disk, which means that $|f(z)| \leqslant C(1 - |z|)^{-\alpha}$, $|z| < 1$, for some constants α and C, while the hyperfunctions impose no restriction on the growth at all. The intermediate Carleman pairs have a temperate behavior along all circles through 1 and -1.

To define (g_1, g_2) it would actually be enough to assume that f_1 and f_2 grow slower than $e^{\varepsilon|z|}$ for every positive ε along every ray (infra-exponential growth). This, however, would allow for a faster growth of (g_1, g_2) at the origin, and it would then not be possible to attenuate the singularity simply by multiplying with a power of y as in (4.5), (4.6); another definition of the transform would be needed. Although Carleman does not offer any comment on this problem, I would surmise that this is the reason why he limited the admissible growth to powers of $|z|$ along the rays.

Given an integrable function we have seen how its Fourier transform is the difference between the boundary values of two holomorphic functions, each defined in a half plane. But how do we represent the function itself as such a difference? The answer is: by forming its convolution with $1/z$. It follows from Plemelj's formulas

$$\lim_{\substack{y\to 0\\ \pm y>0}} \frac{1}{x+iy} = \mathsf{vp}\left(\frac{1}{x}\right) \mp \pi i\delta,$$

that the difference between the limits from the upper and lower half planes of $1/z$ is $-2\pi i\delta$. So, apart from a factor, $1/z$ represents the Dirac measure, the most fundamental distribution.

Let us define a function $E(z) = i/(2\pi z)$ for $z \in \mathbf{C} \setminus \{0\}$ and convolution with E by

$$(E*f)(z) = \int_{\mathbf{R}} E(z-t)f(t)dt = \int_{\mathbf{R}} E(t+iy)f(x-t)dt, \qquad z = x+iy \in \mathbf{C}\setminus\mathbf{R},$$

whenever the integral has a sense, e.g., if $f(x)/(1+|x|)$ is integrable. We may also form the convolution $E*u$ for any distribution u with compact support; it is holomorphic in $\mathbf{C}\setminus \operatorname{supp} u$, where we consider the support of u as a subset of the complex plane.

Let us now see when the two holomorphic functions have a limit at the real axis in the classical sense.

Proposition 7.1. *If $f \in C^1(\mathbf{R})$ and $f(x)/(1+|x|)$ is integrable, then $h(z) = (E*f)(z)$, $\operatorname{Im} z > 0$, is the restriction of a continuous function defined in the closed upper half plane.*

Proof. Any function of class C^1 can be written as

$$f(x+t) = f(x) + t\int_0^1 f'(x+ts)ds = f(x) + tg(x,t), \qquad x,t \in \mathbf{R},$$

where $g(x,t) = \int_0^1 f'(x+ts)ds$ is a continuous function of $(x,t) \in \mathbf{R}^2$. We shall study the behavior of $(E*f)(z)$ when $\operatorname{Re} z$ belongs to a bounded interval $[-a,a]$.

We assume first that f has compact support. We choose a positive number b which is so large that $f(x-t)$ vanishes when $x \in [-a,a]$ and $t \notin [-b,b]$. Then for $x = \operatorname{Re} z \in [-a,a]$ and $y = \operatorname{Im} z > 0$,

$$h(z) = \frac{i}{2\pi}\int_{-b}^{b} \frac{f(x-t)}{t+iy}dt = \frac{if(x)}{2\pi}\int_{-b}^{b}\frac{dt}{t+iy} - \frac{i}{2\pi}\int_{-b}^{b}\frac{t}{t+iy}g(x,-t)dt.$$

The first integral in the last expression can be evaluated, and it is easily seen that it tends to $\frac{1}{2}f(x_0)$ as $x+iy \to x_0$ with $y>0$. In the second integral we note that $t/(t+iy)$ tends to 1 almost everywhere as $y \to 0$ and that $|t/(t+iy)| \leqslant 1$. Lebesgue's theorem on dominated convergence can be applied and we see that the second integral tends to a limit too. The extension to $\mathbf{R}$ is therefore given by

$$\begin{aligned}\lim_{\substack{x+iy\to x_0\\ y>0}} h(x+iy) &= \frac{1}{2}f(x_0) - \frac{i}{2\pi}\int_{-b}^{b} g(x_0,-t)dt\\ &= \frac{1}{2}f(x_0) - \frac{i}{2\pi}\int_{-b}^{b} dt \int_0^1 f'(x_0-ts)ds, \qquad x_0 \in [-a,a].\end{aligned}$$

Next we consider the general case and write $f = \sum_{j\in\mathbf{Z}} f_j$ using a partition of unity, where f_j has its support in the interval $[j-1, j+1]$, say. The argument just presented applies to any finite sum of the $E * f_j$. For indices $j > a+1$ and points z satisfying $\operatorname{Re} z \leqslant a < j-1$ we have the estimate

$$|(E*f_j)(z)| = \left|\int_{j-1}^{j+1} E(z-t)f_j(t)dt\right| \leqslant \frac{1}{2\pi(j-1-a)}\int_{j-1}^{j+1}|f_j(t)|dt.$$

Therefore the sum $\sum_{j>a+1} E * f_j$ converges uniformly for $\operatorname{Re} z \leqslant a$ in view of our hypothesis that $f(t)/(1+|t|)$ is integrable. The terms with $j < -a-1$ can be estimated in the same way, and we are done.

Tillmann [1961a,b] and Martineau [1964] studied systematically the boundary values in the sense of distributions of holomorphic functions.

In the framework of Proposition 7.1 we can form the difference of the extensions from the upper and lower half planes. We see that $h(x+iy) - h(x-iy)$ tends to $f(x_0)$ as $x+iy \to x_0$ while $y > 0$. However, this conclusion holds even if we assume only that f is continuous as the next result shows.

Proposition 7.2. *If $f \in C^0(\mathbf{R})$ and $f(x)/(1+|x|)$ is integrable, then*

$$(E*f)(x+iy) - (E*f)(x-iy) \to f(x_0)$$

locally uniformly as $x+iy \to x_0$ while y is positive.

Proof. We have

$$h(z) = (E*f)(z) = \frac{i}{2\pi}\int_{\mathbf{R}} \frac{f(x-t)}{t+iy}dt, \qquad z = x+iy \in \mathbf{C} \smallsetminus \mathbf{R},$$

so that

$$h(z) - h(\overline{z}) = \frac{i}{2\pi}\int_{\mathbf{R}} f(x-t)\frac{t-iy-(t+iy)}{t^2+y^2}dt = \frac{1}{\pi}\int_{\mathbf{R}} f(x-t)\frac{y}{t^2+y^2}dt$$

for positive y. This is the Poisson integral of f; $\frac{1}{\pi}\frac{y}{x^2+y^2}$ is a well-known approximate identity, so the integral tends to $f(x_0)$ as $x+iy \to x_0$ while $y > 0$, even locally uniformly.

Thus the difference $H(z) = h(z) - h(\overline{z})$, a harmonic function, is much easier to work with than each of the terms when it comes to passage to the limit. (However, there are of course other difficulties connected with the harmonic functions: for instance, they do not form an algebra as do the holomorphic functions.) This transform H is mentioned by Arne Beurling in his note [1949:10]; he called it *la transformée harmonique* and used it to define the spectrum of f. In his book [1983], Hörmander chose this as the main definition of hyperfunctions. A systematic development of the theory of hyperfunctions as boundary values of harmonic functions was undertaken by Komatsu [1991, 1992].

8 Lützen's first question

In his book Jesper Lützen wrote [1982:192]:

> I do not know whether Carleman's function pairs under the conditions (42) always represent distributions. Tillmann's growth condition in [1961b] suggests that this is not the case.

The conditions (42) that Lützen refers to are the conditions (4.2), (4.3) of the present paper.

We shall confirm Lützen's conjecture. In doing so we shall allow ourselves to use freely the language of the later theories of distributions and hyperfunctions.

Fix a point $a \in \mathbf{R}$, $a \neq 0$, and define

$$f(z) = \exp\left(\frac{1}{z-a}\right), \qquad z \in \mathbf{C} \smallsetminus \{a\}.$$

Since f is bounded in a neighborhood of 0 as well as in a neighborhood of ∞, the pair of functions obtained by taking the restriction of f to the upper and lower half planes is a Carleman pair of class $(\alpha, \beta) = (0, 0)$, but it does not represent a Schwartz distribution. Indeed, if it did, then this distribution would have its support contained in the singleton set $\{a\}$, and so would be a finite linear combination of the Dirac measure at a and its derivatives. Since $(x-a)\delta_a = 0$, and similarly $(x-a)^m u = 0$ when u is a derivative of δ_a if only m is large enough, the pair representing u would be entire after multiplication by $(z-a)^m$ for some m. Now this is obviously not the case with $\exp(1/(z-a))$; the singularity at a is essential and cannot be removed just by multiplying with some power of $z-a$.

We use here the fact, well known since the work of Sato, that all distributions, in particular all distributions with compact support, can be represented by pairs of functions holomorphic in the upper and lower half plane, and that this representation is unique up to adding an entire function to both functions in the pair. So the zero distribution is only represented by a pair (f_1, f_2) where the f_j are restrictions of the same entire function.

9 Lützen's second question

Lützen writes in his book [1982:192]

> I have not been able to rigorously prove that Carleman's and Schwartz's Fourier transforms of a tempered distribution are equal; but formal calculations strongly suggest that this is the case.

We can confirm Lützen's suggestion:

Theorem 9.1. *For any temperate distribution $u \in \mathcal{S}'(\mathbf{R})$, Carleman's Fourier transform agrees with Schwartz's Fourier transform; more precisely, u is represented by a Carleman pair (f_1, f_2) and the difference between the boundary values, taken in the sense of distributions, from the upper and lower half planes of Carleman's Fourier transform $\mathrm{CF}(f_1, f_2)$ is equal to Schwartz's Fourier transform $\mathrm{SF}(u)$ of u.*

Writing as before $E(z) = i/(2\pi z)$, defined for $z \in \mathbf{C} \smallsetminus \{0\}$, we know that a distribution $u \in \mathcal{E}'(\mathbf{R})$ is the difference between the boundary values of the holomorphic function $(E * u)(z)$, $z \in \mathbf{C} \smallsetminus \mathbf{R}$. In fact, it is not necessary that u have compact support; it is enough that u is so small at infinity that the convolution has a good sense. In particular we may assume that u is a continuous function which satisfies $|u(x)| \leqslant C(1 + |x|)^{-\alpha}$, $x \in \mathbf{R}$, for some positive α.

If f is a function in $C^1(\mathbf{R}) \cap L^1(\mathbf{R})$, we may form by convolution a Carleman pair (f_1, f_2) to represent it (Proposition 7.1). We then know (Proposition 7.2) that the Carleman Fourier transform of this pair is a Carleman pair representing the classical Fourier transform $\hat{f}$ of f.

We also know that the Schwartz Fourier transform $\mathrm{SF}(f)$ of a function $f \in L^1(\mathbf{R})$ agrees with the classical Fourier transform. When comparing the definitions, we must agree on a definition of the classical Fourier transform. Let us use in the sequel Carleman's definition

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} f(x) e^{-ix\xi} dx, \qquad \xi \in \mathbf{R}.$$

Modifying Schwartz's definition accordingly, we can say that $B(\mathrm{CF}(f_1, f_2)) = \hat{f}$ and $\mathrm{SF}(f) = [\hat{f}]$ for $f \in C^1(\mathbf{R}) \cap L^1(\mathbf{R})$. We express this fact by saying that CF and SF agree on these functions.

To go from these special functions to distributions we shall use the rules $\mathcal{F}(xf) = i\mathcal{F}(f)'$ and $\mathcal{F}(f') = i\xi\mathcal{F}(f)$, which hold for both Carleman's and Schwartz's definitions. By applying them twice we see that

$$\mathcal{F}\big((1-\Delta)f\big) = (1+\xi^2)\mathcal{F}(f) \text{ and } \mathcal{F}\big((1+x^2)f\big) = (1-\Delta)\mathcal{F}(f).$$

This yields, for functions $f \in C^1(\mathbf{R}) \cap L^1(\mathbf{R})$,

$$\mathrm{SF}\big((1+x^2)f\big) = (1-\Delta)\mathrm{SF}(f) = (1-\Delta)\mathrm{CF}(f) = \mathrm{CF}\big((1+x^2)f\big)$$

and

$$\mathrm{SF}\big((1-\Delta)f\big) = (1+\xi^2)\mathrm{SF}(f) = (1+\xi^2)\mathrm{CF}(f) = \mathrm{CF}\big((1-\Delta)f\big).$$

Repeated use of these rules proves that the Schwartz and Carleman transformations agree on all generalized functions of the form $(P_s \circ P_{s-1} \circ \cdots \circ P_1)f$ for some $f \in C^1(\mathbf{R}) \cap L^1(\mathbf{R})$, where each P_j is one of the operators $1 + x^2$, $1 - \Delta$. But this class of generalized functions is equal to all of $\mathcal{S}'(\mathbf{R})$ as shown by Theorem 9.4 below. (We will actually need that result only for $k = 1$ and $m = 0$.)

The mapping $1 - \Delta$ has an inverse, which in one variable is convolution with the function $w(x) = \frac{1}{2}e^{-|x|}$, $x \in \mathbf{R}$. This is an integrable function, and its derivative in the sense of distributions is $w'(x) = w(x)$ for $x < 0$ and $w'(x) = -w(x)$ for $x > 0$, which is also an integrable function, and $\|w\|_1 = \|w'\|_1 = 1$. Its second derivative in the sense of distributions is the measure $w'' = w - \delta$, whose total mass is 2. We thus have three well-defined convolution operators $\varphi \mapsto w*\varphi$, $w'*\varphi$, $w''*\varphi$, satisfying inequalities $\|w*\varphi\|_\infty$, $\|w' * \varphi\|_\infty \leqslant \|\varphi\|_\infty$ and $\|w'' * \varphi\|_\infty \leqslant 2\|\varphi\|_\infty$.

Lemma 9.2. *When u is any of the distributions w, w', w'' defined above, then for all $p \geqslant 0$,*

$$\|(1+x^2)^p(u*\varphi)\|_\infty \leqslant C_p\|(1+x^2)^p\varphi\|_\infty, \qquad \varphi \in \mathcal{S}(\mathbf{R}).$$

Proof. Writing $\psi(x) = (1+x^2)^p\varphi(x)$ we see that we have to prove that

$$\left|u * \frac{\psi}{(1+x^2)^p}\right| \leqslant \frac{C_p}{(1+x^2)^p}$$

when $\|\psi\|_\infty \leqslant 1$. For $u = \delta$ this is clear; for $u = w, w'$ it suffices by symmetry to prove that

$$\int_0^\infty \frac{e^{-y}}{(1+(x-y)^2)^p}dy \leqslant \frac{C_p}{(1+x^2)^p}, \qquad x \in \mathbf{R}.$$

This is easy when $x \leqslant 0$, for then $1/(1+(x-y)^2)^p \leqslant 1/(1+x^2)^p$. When $x > 0$ we consider two integrals. First

$$\int_{x/2}^{\infty} \frac{e^{-y}}{(1+(x-y)^2)^p} dy \leqslant \int_{x/2}^{\infty} e^{-y} dy = e^{-x/2} \leqslant \frac{C_p}{(1+x^2)^p}, \qquad x > 0.$$

Over the interval $[0, x/2]$ we may estimate as follows:

$$\begin{aligned} \int_0^{x/2} \frac{e^{-y}}{(1+(x-y)^2)^p} dy &\leqslant \int_0^{x/2} \frac{e^{-y}}{(1+(x/2)^2)^p} dy \\ &\leqslant \frac{1}{(1+x^2/4)^p} \int_0^{\infty} e^{-y} dy \leqslant \frac{C_p}{(1+x^2)^p}. \end{aligned}$$

Proposition 9.3. *The topology of the space $\mathcal{S}(\mathbf{R})$ is the weakest topology such that all norms*

$$\varphi \mapsto \|\varphi\|_{p,q} = \sup_{x \in \mathbf{R}^n} |(1+x^2)^p (1-\Delta)^q \varphi|, \qquad p, q \in \mathbf{N},$$

are continuous. More explicitly,

$$\|x^j D^k \varphi\|_\infty \leqslant C \|(1+x^2)^p (1-\Delta)^q \varphi\|_\infty, \qquad \varphi \in \mathcal{S}(\mathbf{R}),$$

where $p = j/2$ and $q = k/2$ when $k \in 2\mathbf{N}$; $q = (k+1)/2$ when $k \in 2\mathbf{N}+1$. The norms are essentially increasing in their indices: $\|\varphi\|_{p,q} \leqslant C\|\varphi\|_{p',q'}$ if $p \leqslant p'$, $q \leqslant q'$. (Hence it suffices to use the norms $\|\varphi\|_{p,p}$.)

This implies that the continuity of a temperate distribution is conveniently expressed by an estimate

$$|u(\varphi)| \leqslant C \|\varphi\|_{p,q}, \qquad \varphi \in \mathcal{S}(\mathbf{R}^n). \tag{9.1}$$

Proof. When k is even we have to prove that

$$\|(1+x^2)^p D^{2q} \varphi\|_\infty \leqslant C \|(1+x^2)^p (1-\Delta)^q \varphi\|_\infty,$$

which may be written as

$$\|(1+x^2)^p (w'')^{*q} * \psi\|_\infty \leqslant C \|(1+x^2)^p \psi\|_\infty.$$

To prove this we use the lemma q times with $u = w''$.

When k is odd we have to prove that

$$\|(1+x^2)^p D^{2q-1} \varphi\|_\infty \leqslant C \|(1+x^2)^p (1-\Delta)^q \varphi\|_\infty,$$

which may be written as

$$\|(1+x^2)^p (w'')^{*(q-1)} * w' * \psi\|_\infty \leqslant C \|(1+x^2)^p \psi\|_\infty.$$

Here we use the lemma $q-1$ times with $u=w''$ and once with $u=w'$.

Finally, the inequality $\|\varphi\|_{p,q} \leqslant C\|\varphi\|_{p,q'}$, where $q \leqslant q'$, follows from $q'-q$ applications of the lemma with $u=w$.

Theorem 9.4. *Given any temperate distribution $u \in \mathcal{S}'(\mathbf{R})$ and any numbers $k, m \in \mathbf{N}$, there exist a number $s \in \mathbf{N}$ and a function $f \in C^k(\mathbf{R})$ satisfying $(1+x^2)^m f \in L^1(\mathbf{R})$ such that $u = (P_s \circ P_{s-1} \circ \cdots \circ P_1)f$, where each P_j is equal either to $1-\Delta$ or to multiplication by $1+x^2$.*

This theorem is similar to that of Schwartz [1966:239]; however, it is adapted to the operator $1-\Delta$ and its proof is more direct.

Proof. If u is a temperate distribution, we know that

$$|u(\varphi)| \leqslant C\|\varphi\|_{p,q} = C \sup_{x\in\mathbf{R}} \left|(1+x^2)^p(1-\Delta)^q\varphi(x)\right| \qquad \varphi \in \mathcal{S}(\mathbf{R}),$$

for some constants C, p, and q; see (9.1). There is a distribution v such that

$$(1-\Delta)^q(1+x^2)^p v = u,$$

for the mappings $(1-\Delta)^q$ and $(1+x^2)^p$ are isomorphisms. We see that $|v(\psi)| \leqslant C\|\psi\|_{0,0}$, if ψ is of the form $(1+x^2)^p(1-\Delta)^q\varphi$ for some $\varphi \in \mathcal{S}(\mathbf{R})$. But this means that the estimate holds for all $\psi \in \mathcal{S}(\mathbf{R})$, and thus v is a measure of finite total mass. The convolution product $g = (1-\Delta)^{-1}v = w * v$ is a bounded continuous function. We can then form $h = (1-\Delta)^{-r}g$, which is a bounded function of class C^k if $2r \geqslant k$. Indeed, $(1-\Delta)^{-1}$ maps $C^j \cap L^\infty$ into $C^{j+2} \cap L^\infty$, $j \in \mathbf{N}$, so $(1-\Delta)^{-r}$ maps $C^0 \cap L^\infty$ into $C^{2r} \cap L^\infty$. Finally $f = (1+x^2)^{-m-1}h$ is such that $(1+x^2)^m f \in L^1(\mathbf{R})$.

Collecting what we have done we see that

$$u = (1-\Delta)^q(1+x^2)^p(1-\Delta)^{r+1}(1+x^2)^{m+1}f.$$

10 Conclusion

In his lectures in 1935, Torsten Carleman represented the Fourier transform of a function of temperate growth as a pair of functions defined in the upper and lower half planes, respectively. He also extended the Fourier transformation to be defined on such pairs. In modern parlance, he defined the Fourier transformation of a class of hyperfunctions, but only in one variable. Although very different in nature, his definition agrees with the one given later by Laurent Schwartz for temperate distributions. His calculus, however, is valid for a class of hyperfunctions strictly larger than the temperate distributions.

Carleman's monograph [1944] was probably not well-known; the same goes for the proceedings article [1949]. However, a pirate edition of Carleman's book [1944] was published in Japan after the war (Professor Hikosaburo Komatsu, personal communication, November 28, 2001). Nevertheless, it seems that the work of Carleman did not play a role in the early development of the theory of hyperfunctions in Japan. Lützen [1982:191] writes that the connection was pointed out only by Bremermann & Durand in [1961]. Indeed they quote both Carleman [1944] and Sato [1958a].

References

Beurling, Arne

1949 Sur les spectres des fonctions. *Analyse harmonique; Nancy, 15–22 Juin 1947,* pp. 9–29. Colloques internationaux du Centre national de la Recherche scientifique, XV. Paris: CNRS. 133 pp.

Bochner, Salomon

1932 *Vorlesungen über Fouriersche Integrale.* 229 pp. (Reprinted 1948 by Chelsea Publishing Company, New York.)

1952 Review of *Théorie des distributions* by Laurent Schwartz [1950, 1951]. *Bull. Amer. Math. Soc.* **58**, 78–85.

1959 *Lectures on Fourier Integrals.* Annals of Mathematics Studies, Number 42. Princeton: Princeton University Press. x + 333 pp. Translated by Morris Tenenbaum and Harry Pollard.

Bremermann, H. J.; Durand, L., III

1961 On analytic continuation, multiplication, and Fourier transformations of Schwartz distributions. *J. Mathematical Physics* **2**, 240–258.

Carleman, T[orsten]

1944 *L'intégrale de Fourier et questions qui s'y rattachent.* Publications Scientifiques de l'Institut Mittag-Leffler, 1. 119 pp. (Reprinted 1967 with an additional note on page 107.)

1949 Sur l'application de la théorie des fonctions analytiques dans la théorie des transformées de Fourier. *Analyse harmonique; Nancy, 15–22 Juin 1947,* pp. 45–53. Colloques internationaux du Centre national de la Recherche scientifique, XV. Paris: CNRS. 133 pp.

Ehrenpreis, Leon

1954 Solution of some problems of division. I. Division by a polynomial of derivation. *Amer. J. Math.* **76**, 883–903.

1956 Analytic functions and the Fourier transform of distributions. I. *Ann. of Math.* (2) **63**, 129–159.

Gel′fand, I. M.; Šilov, G. E.

1953 Fourier transforms of rapidly increasing functions and the questions of uniqueness of the solution of Cauchy's problem. [Russian.] *Uspehi Matem. Nauk (N.S.)* **8**, no. 6(58), 3–54.

Hörmander, Lars

1955 La transformation de Legendre et le théorème de Paley–Wiener. *C. R. Acad. Sci. Paris* **240**, 392–395.

1983 *The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis.* Berlin: Springer-Verlag. ix + 391 pp.

1998 On the Legendre and Laplace transformations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **25,** 517–568.

Kawai, T.

1970a *The theory of Fourier transforms in the theory of hyperfunctions and its applications.* [Japanese.] Master's Thesis, University of Tokyo.

1970b On the theory of Fourier hyperfunctions and its applications to partial differential equations with constant coefficients. *J. Fac. Sci. Univ. Tokyo Sect. IA,* **17**, 467–517.

Komatsu, Hikosaburo

1991 Microlocal analysis in Gevrey classes and in complex domains. *Microlocal analysis and applications (Montecatini Terme, 1989),* 161–236. Lecture Notes in Mathematics, 1495. Berlin: Springer-Verlag.

1992 An elementary theory of hyperfunctions and microfunctions. *Partial differential equations, Part 1, 2 (Warsaw, 1990),* 233–256. Banach Center Publ., 27, Part 1, 2. Warsaw: Polish Academy of Sciences.

Lüzten, Jesper

1982 *The Prehistory of the Theory of Distributions.* New York: Springer-Verlag. viii + 232 pp.

McKennon, Kelly

1976 Analytic distributions. *J. reine angew. Math.* **281,** 164–178.

Martineau, André

1964 Distributions et valeurs au bord des fonctions holomorphes. *Theory of distributions.* Proceedings of an International Summer Institute held in Lisbon, September 1964, pp. 195–326. Lisbon: Instituto Gulbenkian de Cincia.

Morimoto, Mitsuo
1973 On the Fourier ultra-hyperfunctions I. [Japanese.] *Res. Inst. Math. Sci. Kôkyûroku* **192**, 10–34.
1978 Analytic functionals with noncompact carrier. *Tokyo J. Math.* **1**, 72–103.

Saburi, Yutaka
1985 Fundamental properties of modified Fourier hyperfunctions. *Tokyo J. Math.* **8**, 231–273.

Sato, Mikio
1958a On a generalization of the concept of functions. *Proc. Japan Acad.* **34**, 126–130.
1958b On a generalization of the concept of functions. II. *Proc. Japan Acad.* **34**, 604–608.
1958c Theory of hyperfunctions. [Japanese.] *Sukagu* **10**, 1–27.
1959 Theory of hyperfunctions I. *J. Fac. Sci. Tokyo* Sect. I, **8**, 139–193.
1960 Theory of hyperfunctions II. *J. Fac. Sci. Tokyo* Sect. I, **8**, 387–437.

Schwartz, Laurent
1949 Théorie des distributions et transformation de Fourier. *Analyse harmonique; Nancy, 15–22 Juin 1947,* pp. 1–8. Colloques internationaux du Centre national de la Recherche scientifique, XV. Paris: CNRS. 133 pp.
1950/51 *Théorie des distributions.* Vol. I, II. Publications de l'Institut de Mathématique de l'Université de Strasbourg, nos. 9 and 10; Actualités scientifiques et industrielles, nos. 1091 and 1122.
1966 *Théorie des distributions.* Paris: Hermann. XIII + 420 pp.

Sebastio e Silva, J.
1964 Integrals and orders of growth of distributions. *Theory of distributions.* Proceedings of an International Summer Institute held in Lisbon, September 1964, pp. 327–390. Lisbon: Instituto Gulbenkian de Cincia.

Tillmann, Heinz Günther
1961a Distributionen als Randverteilungen analytischer Funktionen II. *Math. Z.* **76,** 5–21.
1961b Darstellungen der Schwartzen Distributionen durch analytische Funktionen. *Math. Z.* **77,** 106–124.