Publ. Math. Debrecen
48 / 1-2 (1996), 103–115

Functional equations, DEs and distributions

By J. A. BAKER (Waterloo, Canada)

Abstract. The notion of pullback (which generalizes that of composition) from the theory of Schwartz distributions is used to find the locally integrable solutions of certain functional equations involving complex valued functions of several real variables. In the process the solutions of natural distributional analogues of these equations are also determined.

Hilbert remarked in connection with his fifth problem that, despite its power and elegance, the method of solving functional equations by reduction to differential equations is unsatisfying in that it requires unnatural smoothness assumptions on the unknown functions; see [1]. In certain cases this deficiency can be overcome by employing regularity theorems (for example, of the type "measurability implies differentiability" to be found in JÁRAI [6]) or by appealing to the theory of Schwartz distributions (see, e.g., [2] and the references contained therein).

The aim of this paper is to illustrate how distribution theory can be useful in the study of certain functional equations of the form

$$\sum_{k=0}^{N} c_k(\xi) f_k(F_k(\xi)) = 0, \quad \xi \in \Omega. \tag{1}$$

Here Ω is an open (nonempty) subset of $\mathbb{R}^n$, $c_k : \Omega \to \mathbb{C}$ and $F_k : \Omega \to \mathbb{R}^{m_k}$, $m_k < n$, are given C^∞ maps for $0 \le k \le N$ and $f_k : F_k(\Omega) \to \mathbb{C}$, $0 \le k \le N$, are the unknowns. Note that (1) may be written more concisely as

$$\sum_{k=0}^{N} c_k (f_k \circ F_k) = 0. \tag*{(1)$'$}$$

Supported by NSERC (Canada) Grant #7153

As we will see, if each F_k has surjective derivatives then $(1)'$ can be interpreted in a distributional sense in such a way that reduction to distributional differential equations is at least plausible. In such cases, regularity theorems for distributional differential equations (e.g. the elliptic regularity theorem) sometimes allow one to determine the continuous (or even locally integrable) solutions of (1) and to solve natural distributional analogues thereof.

Páles, in [8], has studied (1) in the case $n = 2$ and $m_k = 1$ for $0 \leq k \leq N$. Assuming some smoothness of the given data and assuming a kind of independence and spanning property of $\{F_0, \dots, F_N\}$ he showed that each $f_k \circ F_k$ satisfies a certain linear partial differential equation. This has provided motivation for the present paper.

Some distributional background

We will, for the most part, use the notation of Rudin [9] but the operator $\frac{\partial}{\partial x_i}$ will be written ∂_i when applied to C^1 functions and written D_i when applied to distributions.

Suppose that Ω is an open subset of $\mathbb{R}^n$ and $f : \Omega \to \mathbb{C}$. Then we say that f is *locally integrable* on Ω provided it is Lebesgue measurable on Ω and $\int_K |f(x)|dx < +\infty$ for every compact $K \subseteq \Omega$; the set of all such f is denoted by $L^1_{\text{loc}}(\Omega)$. If $f \in L^1_{\text{loc}}(\Omega)$ and we define $\Lambda_f(\varphi) = \int_\Omega f(x)\varphi(x)dx$ for $\varphi \in \mathcal{D}(\Omega)$ then $\Lambda_f \in \mathcal{D}'(\Omega)$ and we call Λ_f the *regular* distribution corresponding to f. For $f, g \in L^1_{\text{loc}}(\Omega)$, $\Lambda_f = \Lambda_g$ if and only if $f(x) = g(x)$ for a.e. $x \in \Omega$.

By "$u_j \to u$ in $\mathcal{D}'(\Omega)$" we mean $u \in \mathcal{D}'(\Omega)$, $u_j \in \mathcal{D}'(\Omega)$ for $j = 1, 2, \dots$ and $\lim_{j\to\infty} u_j(\varphi) = u(\varphi)$ for every $\varphi \in \mathcal{D}(\Omega)$. We will need the fact that $C_c(\Omega)$ is "dense" in $\mathcal{D}'(\Omega)$ in the sense that, given $u \in \mathcal{D}(\Omega)$, there exists a sequence $\{f_j\}_{j=1}^\infty$ in $C_c(\Omega)$ such that $\Lambda_{f_j} \to u$ in $\mathcal{D}'(\Omega)$; see [5], page 89.

Composing a distribution with a smooth map

In order to treat $(1)'$ in a distributional way we must generalize the notion of composition appropriately. Suppose that U is an open subset of $\mathbb{R}^m$, V is an open subset of $\mathbb{R}^n$, $F : V \to U$, F is C^∞ on V and rank $\nabla F(x) = m$ for all $x \in V$. Then F is called a *submersion* of V into U (see [4] page 84); this implies that $m \leq n$ and F is open. Our development rests heavily upon the following

Composition Theorem (see [5], pages 133–135 and [4] pages 80–85.). *Suppose that F is a submersion of an open $V \subseteq \mathbb{R}^n$ into an open $U \subseteq \mathbb{R}^m$. Then there exists a unique continuous linear map $F^* : \mathcal{D}'(U) \to \mathcal{D}'(V)$ such that $F^*\Lambda_f = \Lambda_{f\circ F}$ for all $f \in C_c(U)$; we will usually write $u \circ F$ instead of F^*u for $u \in \mathcal{D}'(U)$. Moreover, the following are true.*

(i) *The chain rule holds: if $u \in \mathcal{D}'(U)$ and $1 \leq j \leq n$ then*

$$D_j(u \circ F) = \sum_{\nu=1}^{m} (\partial_j F_\nu)((D_\nu u) \circ F)$$

where $F(x) = (F_1(x), \dots, F_m(x))$ for $x \in V$.

(ii) *If $c \in C^\infty(U)$ and $u \in \mathcal{D}'(U)$ then $(cu) \circ F = (c \circ F)(u \circ F)$.*

(iii) *If W is an open subset of $\mathbb{R}^d$ and G is a submersion of W into V then $(u \circ F) \circ G = u \circ (F \circ G)$ for all $u \in \mathcal{D}'(U)$.*

The map F^* is often called the *pullback* by F. To say that F^* is continuous means that if $u_j \to u$ in $\mathcal{D}'(U)$ then $u_j \circ F \to u \circ F$ in $\mathcal{D}'(V)$.

We will show that if F, U and V are as in the Composition Theorem and $f \in L^1_{\text{loc}}(U)$ then $f \circ F \in L^1_{\text{loc}}(V)$ and $\Lambda_f \circ F = \Lambda_{f\circ F}$. Assuming this for the moment, suppose that Ω is an open subset of $\mathbb{R}^n$, for $0 \leq k \leq N$, $c_k \in C^\infty(\Omega)$, F_k is a submersion of Ω into an open subset U_k of $\mathbb{R}^{m_k}$, $f_k \in L^1_{\text{loc}}(U_k)$ and (1) holds almost everywhere, i.e.

$$\sum_{k=0}^{N} c_k(\xi) f_k(F_k(\xi)) = 0 \quad \text{for a.e. } \xi \in \Omega.$$

If we let $u_k = \Lambda_{f_k}$ for $0 \leq k \leq N$ it follows that

$$(1)'' \qquad \sum_{k=0}^{N} c_k(u_k \circ F) = 0.$$

The advantage of $(1)''$ is that we can apply distributional derivatives.

There are two kinds of submersions that are of particular importance; diffeomorphisms and projections. As we will see, every submersion is locally a composition of at most three submersions, each of one of these two types. In proving the following two propositions we will essentially rely on ideas from the proof of the Composition Theorem given on page 135 of [5].

Suppose that F is a diffeomorphisms of an open subset V of $\mathbb{R}^n$ *onto* an open subset U of $\mathbb{R}^n$, i.e. F is a C^∞ bijection of V onto U whose inverse is C^∞. For $x \in V$ let

$$JF(x) = |\det \nabla F(x)| > 0 \quad \text{for } x \in V.$$

Suppose that $f \in C_c(U)$ and $\varphi \in \mathcal{D}(V)$. Then, by the change of variable theorem, $\int_V f(F(x))\varphi(x)dx = \int_U f(y)\ JF^{-1}(y)\ \varphi(F^{-1}(y))dy$,
i.e. $(\Lambda_f \circ F)(\varphi) = \Lambda_{f\circ F}(\varphi) = \Lambda_f(F^\# \varphi)$ where $F^\#\varphi = (JF^{-1})(\varphi \circ F^{-1})$ for $\varphi \in \mathcal{D}(V)$. It is not difficult to check that $F^\#$ is a continuous linear bijection of $\mathcal{D}(V)$ onto $\mathcal{D}(U)$ and $(F^\#)^{-1} = (F^{-1})^\#$. Now if $u \in \mathcal{D}'(U)$ and $\{f_j\}_{j=1}^\infty$ is a sequence in $C_c(U)$ such that $\Lambda_{f_j} \to u$ it follows that

$$(U \circ F)(\varphi) = \lim_{j\to\infty} (\Lambda_{f_j} \circ F)(\varphi) = \lim_{j\to\infty} \Lambda_{f_j}(F^\#\varphi) = u(F^\#\varphi)$$

for all $\varphi \in \mathcal{D}(V)$. In summary we have

Proposition 1. *If F is a diffeomorphism of an open subset V of $\mathbb{R}^n$ onto an open subset U of $\mathbb{R}^n$ then, for all $u \in \mathcal{D}'(U)$, $F^*u = u \circ F^\#$, i.e. $(u \circ F)(\varphi) = u(F^\#\varphi) = u((JF^{-1})(\varphi \circ F^{-1}))$ for all $\varphi \in \mathcal{D}(V)$. Moreover, for $u \in \mathcal{D}'(U)$, $u \circ F = 0$ if and only if $u = 0$. In fact $F^\#$ is a bijection and $(F^\#)^{-1} = (F^{-1})^\#$.*

Now suppose that V is an open subset of $\mathbb{R}^n$, $m < n$, $P(x) = (x_1, \ldots, x_m)$ for all $x = (x_1, \ldots, x_n) \in V$, U is an open subset of $\mathbb{R}^m$ and $P(V) \subseteq U$. Then we call P a *projection* of V into U; it is clearly a submersion. Define $1_{n-m} \in \mathcal{D}'(\mathbb{R}^{n-m})$ by $1_{n-m}(\chi) = \int_{\mathbb{R}^{n-m}} \chi$ for all $\chi \in \mathcal{D}(\mathbb{R}^{n-m})$. If $f \in C_c(U)$, $\varphi \in \mathcal{D}(P(V))$, $\chi \in \mathcal{D}(\mathbb{R}^{n-m})$ and $\varphi \otimes \chi \in \mathcal{D}(V)$ then

$$\begin{aligned}
&(\Lambda_f \circ P)(\varphi \otimes \chi) \\
&= \int_V f(x_1, \ldots, x_m)\varphi(x_1, \ldots, x_m)\chi(x_{m+1}, \ldots, x_n)d(x_1, \ldots, x_n) \\
&= \Lambda_f(\varphi)\ 1_{n-m}(\chi) \ = \ (\Lambda_f \otimes 1_{n-m})(\varphi \otimes \chi).
\end{aligned}$$

It follows that $\Lambda_f \circ P = (\Lambda_f \otimes 1_{n-m})|_{\mathcal{D}(V)}$; see [5], page 127. Using the denseness of $C_c(U)$ in $\mathcal{D}'(U)$ and the continuity of P^* we conclude that $u \circ P = (u \otimes 1_{n-m})|_{\mathcal{D}(V)}$ for all $u \in \mathcal{D}'(U)$. We therefore have proved most of

Proposition 2. *If P is a projection of an open subset V of $\mathbb{R}^n$ into an open subset U of $\mathbb{R}^m$ (where $m < n$), then $u \circ P = (u \otimes 1_{n-m})|_{\mathcal{D}(V)}$ for all $u \in \mathcal{D}'(U)$. Moreover, for $u \in \mathcal{D}'(U)$, $u \circ F = 0$ if and only if u vanishes on $P(V)$.*

Proof. It suffices to prove the last assertion. To this end, suppose that $u \in \mathcal{D}'(U)$ and u does not vanish on $\mathcal{D}(P(V))$. Choose $\varphi \in \mathcal{D}(P(V))$ and $\chi \in \mathcal{D}(\mathbb{R}^{n-m})$ such that $\varphi \otimes \chi \in \mathcal{D}(V)$, $u(\varphi) \neq 0$ and $\int_{\mathbb{R}^{n-m}} \chi = 1$. Then $(u \circ F)(\varphi \otimes \chi) = u(\varphi) \neq 0$. □

According to the Inverse Function Theorem ([3], page 42) a submersion of an open subset V of $\mathbb{R}^n$ into an open subset U of $\mathbb{R}^m$ is locally a diffeomorphism if $m = n$. When $m < n$ the local nature of a submersion is revealed by the following (special case of the)

Rank Theorem ([3], page 47)**.** *Suppose that F is a submersion of an open subset V of $\mathbb{R}^n$ into an open subset U of $\mathbb{R}^m$, $m < n$, $a \in V$ and $b = F(a) \in U$. Then there exists an open subset V_0 of $\mathbb{R}^n$ such that $a \in V_0 \subseteq V$, an open subset U_0 of $\mathbb{R}^m$ such that $b \in U_0 \subseteq U$ and $F(V_0) = U_0$, a diffeomorphism G of V_0 onto an open subset V' of $\mathbb{R}^n$ and a diffeomorphism H of U_0 onto an open subset U' of $\mathbb{R}^m$ such that $H \circ F \circ G^{-1}(x) = (x_1, \ldots, x_m) =: P(x)$ for all $x = (x_1, \ldots, x_n) \in V'$ and $H \circ F \circ G^{-1}(V') = U'$ or, equivalently, $F|_{V_0} = H^{-1} \circ P \circ G$.*

The following proposition highlights the local nature of composition.

Proposition 3. *Suppose that U, V and F are as in the Composition Theorem, V_0 is an open subset of V and $u \in \mathcal{D}'(U)$. Then*

$$(*) \qquad (u \circ F)|_{\mathcal{D}(V_0)} = u \circ (F|_{V_0}).$$

Proof. If $f \in C_c(U)$ then, for every $\varphi \in \mathcal{D}(V_0)$, $(\Lambda_f \circ F)(\varphi) = \int_V f(F(x))\varphi(x)dx = \int_{V_0} (f \circ (F|_{V_0}))(x)\varphi(x)dx = (\Lambda_f \circ (F|_{V_0}))(\varphi)$. That is, $(*)$ holds if $u = \Lambda_f$ for some $f \in C_c(U)$.

Let $u \in \mathcal{D}'(U)$. Choose a sequence $\{f_j\}_{j=1}^{\infty}$ in $C_c(U)$ such that $\Lambda_{f_j} \to u$ in $\mathcal{D}'(U)$. Then $u \circ F|_{\mathcal{D}(V_0)} = \left(\lim_{j\to\infty}(\Lambda_{f_j} \circ F)\right)|_{\mathcal{D}(V_0)}$
$= \lim_{j\to\infty}(\Lambda_{f_j} \circ F)|_{(\mathcal{D}V_0))} = \lim_{j\to\infty} \Lambda_{f_j} \circ (F|_{V_0}) = u \circ (F|_{V_0})$. □

Combining this assertion with Theorem 2.2.1 of [5] we have the

Corollary. *With U, V and F as in the Composition Theorem and $u, v \in \mathcal{D}'(U)$, $u = v$ if and only if for each $a \in V$, there exists an open $V_0 \subseteq V$ such that $a \in V_0$ and $u \circ (F|_{V_0}) = v \circ (F|_{V_0})$.*

The next result implies that if F is a submersion of V *onto* U then F^* is injective.

Proposition 4. *Suppose that F is a submersion of an open subset $V \subseteq \mathbb{R}^n$ into an open subset $U \subseteq \mathbb{R}^m$. If $u \in \mathcal{D}'(U)$ and $u \circ F = 0$ then u vanished on $F(V)$, i.e. $u(\varphi) = 0$ for all $\varphi \in \mathcal{D}(F(V))$.*

Proof. We know from Propositions 1 and 2 that the assertion is true for surjective diffeomorphisms and projections. The general assertion follows by localization using the Corollary, the Rank Theorem and (iii) of the Composition Theorem. □

Pullbacks of regular distributions

The aim of this section is to prove

Proposition 5. *Suppose that F is a submersion of an open subset V of $\mathbb{R}^n$ into an open subset U of $\mathbb{R}^m$ and $f \in L^1_{\mathrm{loc}}(U)$. Then $f \circ F \in L^1_{\mathrm{loc}}(V)$ and $\Lambda_{f\circ F} = \Lambda_f \circ F$.*

Proof. Suppose first that F is a diffeomorphism of V onto U. By the change of variable theorem, $f \circ F \in L^1_{\mathrm{loc}}(V)$ and $\Lambda_{f\circ F}(\varphi) = \int_V f(F(x))\ \varphi(x)dx = \int_U f(y)F^{\#}\varphi(y)dy = \Lambda_f(F^{\#}\varphi)$ for all $\varphi \in \mathcal{D}(V)$. Choose a sequence $\{f_j\}_{j=1}^{\infty}$ in $C_c(U)$ such that $\Lambda_{f_j} \to \Lambda_f$ in $\mathcal{D}'(U)$. Then

$$\begin{aligned}\Lambda_{f\circ F}(\varphi) &= \Lambda_f(F^{\#}\varphi) = \lim_{j\to\infty} \Lambda_{f_j}(F^{\#}\varphi) = \lim_{j\to\infty} \Lambda_{f_j\circ F}(\varphi) \\ &= \lim_{j\to\infty} (\Lambda_{f_j} \circ F)(\varphi) = ((\lim_{j\to\infty} \Lambda_{f_j}) \circ F)(\varphi) \\ &= (\Lambda_f \circ F)(\varphi) \quad \text{for all } \varphi \in \mathcal{D}(V)\end{aligned}$$

by the continuity of $F^{\#}$, i.e. $\Lambda_{f\circ F} = \Lambda_f \circ F$.

Next suppose that F is a projection of V onto U. To see that $f \circ F$ is Lebesgue measurable note that, for any open $\mathcal{O} \subseteq \mathbb{C}$, $f^{-1}(\mathcal{O})$ is Lebesgue measurable and so $(f \circ F)^{-1}(\mathcal{O}) = F^{-1}(f^{-1}(\mathcal{O})) = \{(x_1, \dots, x_n) \in V : (x_1, \dots, x_m) \in f^{-1}(\mathcal{O})\} = V \cap (f^{-1}(\mathcal{O}) \times \mathbb{R}^{n-m})$ is Lebesgue measurable. Moreover, for each $\varphi \in \mathcal{D}(V)$,

$$\Lambda_{f\circ F}(\varphi) = \int_V f(x_1, \dots, x_m)\varphi(x_1, \dots, x_n)d(x_1, \dots, x_n).$$

It follows that $\Lambda_{f\circ F} = (\Lambda_f \otimes 1_{n-m})|_{\mathcal{D}(V)}$. Choose a sequence $\{f_j\}_{j=1}^{\infty}$ in $C_c(U)$ such that $\Lambda_{f_j} \to \Lambda_f$. Then, by the prelude to Proposition 2,

$$\Lambda_{f\circ F} = \lim_{j\to\infty} (\Lambda_{f_j} \otimes 1_{n-m})|_{\mathcal{D}(V)} = \lim_{j\to\infty} (\Lambda_{f_j} \circ F) = \Lambda_f \circ F$$

where the continuity of F^* has been used once again.

The general case follows from the Rank Theorem and the Corollary. □

Applications to functional equations

1. Consider the example presented by Páles [8];

$$f_1(x+y)+f_2(x-y)+f_3(xy)=0. \tag{2}$$

Let $V=\{(x,y)\in\mathbb{R}^2 : 0<y<x\}$ and $U=(0,+\infty)$.

Suppose that $f_j\in C^2(U)$ for $1\le j\le 3$ and (2) holds for all $(x,y)\in V$. By replacing x by $\frac{s+t}{2}$ and y by $\frac{s-t}{2}$ in (2) we conclude that

$$f_1(s)+f_2(t)+f_3\left(\frac{s^2-t^2}{4}\right)=0 \quad \text{for all } (s,t)\in V. \tag{3}$$

Indeed (3) is equivalent to (2) because the map $(s,t)\to\left(\frac{s+t}{2},\frac{s-t}{2}\right)$ is a diffeomorphism of V onto itself; for future reference call this map G. Now differentiate (3) with respect to s to conclude that

$$f_1'(s)+\frac{s}{2}\,f_3'\left(\frac{s^2-t^2}{4}\right)=0 \quad \text{for } (s,t)\in V. \tag{4}$$

By differentiating (3) with respect to t we find that

$$f_2'(t)-\frac{t}{2}\,f_3'\left(\frac{s^2-t^2}{4}\right)=0 \quad \text{for } (s,t)\in V. \tag{5}$$

By differentiating (4) with respect to t we find that

$$f_3''\left(\frac{s^2-t^2}{4}\right)=0 \quad \text{for all } (s,t)\in V \tag{6}$$

or, equivalently,

$$f_3''(x)=0 \quad \text{for all } x>0. \tag{7}$$

Hence there exist $\beta,\gamma\in\mathbb{C}$ such that

$$f_3(x)=\beta+4\gamma x \quad \text{for all } x>0. \tag{8}$$

Now (4) and (8) imply that

$$f_1'(s)+2\gamma s=0 \quad \text{for all } s>0 \tag{9}$$

so that, for some $\alpha_1\in\mathbb{C}$,

$$f_1(s)=\alpha_1-\gamma s^2 \quad \text{for all } s>0. \tag{10}$$

Similarly, from (5) and (8) we surmise that

$$f_2'(t)-2\gamma t=0 \quad \text{for all } t>0 \tag{11}$$

and hence, for some $\alpha_2 \in \mathbb{C}$,

$$(12) \qquad f_2(t) = \alpha_2 + \gamma t^2 \quad \text{for all } t > 0.$$

Now (2), (8), (10) and (12) imply that

$$(13) \qquad \alpha_1 + \alpha_2 + \beta = 0.$$

Conversely, if (8), (10), (12) and (13) holds then it is trivial to check that (2) holds for all $(x, y) \in V$. We thus have proved

Proposition 6. *If $f_j \in C^2(0, +\infty)$ for $1 \le j \le 3$ then* (2) *holds for all $(x, y) \in V$ if and only if there exists $\alpha_1, \alpha_2, \beta, \gamma \in \mathbb{C}$ such that* (8), (10), (12) *and* (13) *hold.*

We aim to solve a distributional analogue of (2) and thereby determine those locally integrable f_1, f_2 and f_3 for which (2) holds almost everywhere.

To this end, let $F_1(x, y) = x+y,\ F_2(x, y) = x-y$ and $F_3(x, y) = xy$ for $(x, y) \in V$. Note that F_1, F_2 and F_3 are submersions of V onto U. Suppose that $f_1, f_2, f_3 \in L^1_{\mathrm{loc}}(U)$, (2) holds for a.e. $(x, y) \in V$ and let $u_j = \Lambda_{f_j}$ for $1 \le j \le 3$. Then it follows from Proposition 5 that

$$(2)' \qquad \sum_{j=1}^{3} u_j \circ F_j = 0.$$

Now suppose that $u_j \in \mathcal{D}'(U)$ for $1 \le j \le 3$ and $(2)'$ holds.
Then $0 = \sum_{j=1}^3 (u_j \circ F_j) \circ G = \sum_{j=1}^3 u_j \circ (F_j \circ G)$. But

$$\begin{aligned} F_1 \circ G\ (s, t) &= s =: P_1(s, t), \\ F_2 \circ G\ (s, t) &= t =: P_2(s, t) \quad \text{and} \\ F_3 \circ G\ (s, t) &= \frac{s^2 - t^2}{4} =: F(s, t) \quad \text{for all } (x, t) \in V \end{aligned}$$

so that

$$(3)' \qquad u_1 \circ P_1 + u_2 \circ P_2 + u_3 \circ F = 0.$$

Note that $(3)'$ was deduced from $(2)'$ in essentially the same way that (3) was deduced from (2). By applying D_1 to $(3)'$ we find, with the help of the chain rule, that

$$(4)' \qquad (Du_1) \circ P_1 + (\partial_1 F)((Du_3) \circ F) = 0$$

— a distributional analogue of (4). Similarly,

$$(5)' \qquad (Du_2) \circ P_2 + (\partial_2 F)((Du_3) \circ F) = 0$$

Now apply D_2 to $(4)'$ and use the chain rule again to conclude that

$$(\partial_2 P_1)((D^2u) \circ P_1) + (\partial_2\partial_1 F)((Du_3) \circ F) + (\partial_1 F)(\partial_2 F)((D^2u_3) \circ F) = 0$$

or

$$(D^2u_3) \circ F = 0 \tag{6'}$$

since $\partial_2 P_1 \equiv 0, \partial_2\partial_1 F \equiv 0$ and $(\partial_1 F)(\partial_2 F)(s,t) = \frac{-st}{4} \neq 0$ for all $(s,t) \in V$. Since F maps V onto U, from $(6)'$ and last part of Proposition 4 we conclude that

$$D^2u_3 = 0. \tag{7'}$$

Now the distributional solutions of a linear ordinary differential equation with C^∞ coefficients (and nowhere vanishing leading coefficient) are regular distributions corresponding to the classical solutions (all of which are C^∞ functions); see [5], page 58. Thus it follows from $(7)'$ that there exists $\beta, \gamma \in \mathbb{C}$ such that

$$u_3 = \Lambda_{f_3} \tag{8'}$$

where f_3 is defined by (8). Similarly, by using an argument like that employed to deduce (10) and (12) we find that there exist $\alpha_1, \alpha_2 \in \mathbb{C}$ such that $u_1 = \Lambda_{f_1}$ and $u_2 = \Lambda_{f_2}$ where f_1 and f_2 are defined by (10) and (12). Now $0 = \sum_{j=1}^{3} \Lambda_{f_j} \circ F_j = \sum_{j=1}^{3} \Lambda_{f_j \circ F_j} = \Lambda_f$, where $f = \sum_{j=1}^{3} f_j \circ F_j$, and f is continuous, so $f \equiv 0$ on V, i.e. (2) holds. By Proposition 6, $\alpha_1 + \alpha_2 + \beta = 0$. We have proved most of

Proposition 7. *Suppose that U, V, F_1, F_2 are F_3 are as above.*

(i) *If $u_1, u_2, u_3 \in \mathcal{D}'(U)$ then $(2)'$ holds if and only if there exist $\alpha_1, \alpha_2, \beta, \gamma \in \mathbb{C}$ such that $u_j = \Lambda_{f_j}$ for $1 \leq j \leq 3$ and* (8), (10), (12) *and* (13) *hold.*

(ii) *If $g_i \in L^1_{\text{loc}}(U)$ for $1 \leq j \leq 3$ and*

$$g_1(x+y) + g_2(x-y) + g_3(xy) = 0 \quad \text{for a.e. } (x,y) \in V \tag{2''}$$

then there exist $f_1, f_2, f_3 \in C^\infty(U)$ such that (2) *holds and, for each $j = 1, 2, 3$, $g_j(x) = f_j(x)$ for a.e. $x > 0$.*

Proof. It remains only to prove (ii). So assume that $(2)''$ holds with $g_1, g_2, g_3 \in L^1_{\mathrm{loc}}(0,+\infty)$. Let $u_j = \Lambda_{g_j}$ for $1 \le j \le 3$ and let $h = \sum_{j=1}^3 g_j \circ F_j$. Then $(2)''$ says that $h(x,y) = 0$ for a.e. $(x,y) \in V$ so that, by Proposition 5, $0 = \Lambda_h = \sum_{j=1}^3 \Lambda_{g_j \circ F_j} = \sum_{j=1}^3 \Lambda_{g_j} \circ F_j = \sum_{j=1}^3 u_j \circ F_j$. Hence, by (i) there exist $f_j \in C^\infty(U)$ such that (2) holds and, for $1 \le j \le 3$, $\Lambda_{g_j} = u_j = \Lambda_{f_j}$. Hence

$$g_j(x) = f_j(x) \quad \text{for a.e. } x > 0. \qquad \square$$

In the interest of clarity we have included more details than many readers may find necessary. We will be more concise in the remaining two examples.

2. As an example of (1) in which conconstant coefficients appear, consider the functional equation

$$f(x) + (1-x)^\alpha f\left(\frac{y}{1-x}\right) = f(y) + (1-y)^\alpha f\left(\frac{x}{1-y}\right) \tag{14}$$

on $\Omega = \{(x,y) \in \mathbb{R}^2 : x, y > 0 \text{ and } x + y < 1\}$ and where α is a given real constant. A function $f : (0,1) \to \mathbb{R}$ which satisfies (14) is usually called an *information function of degree* α. Such functions have been investigated by numerous authors; see e.g. [7] and its references. We aim to consider (14) from a distributional point of view. We will also assume that $\alpha \notin \{0,1,2\}$; the other three cases can be treated in a similar manner. As in the first example, we begin by exhibiting a method for finding the smooth solutions of (14) and then observe that an analogous strategy bears fruit in a distributional setting.

So suppose that $f \in C^2(0,1)$ and (14) holds. Differentiate with respect to x and then with respect to y to conclude that

$$\begin{aligned}(1-\alpha)(1-x)^{\alpha-2} f'\left(\frac{y}{1-x}\right) \;+\; y(1-x)^{\alpha-3} f''\left(\frac{y}{1-x}\right)\\ = (1-\alpha)(1-y)^{\alpha-2} f'\left(\frac{x}{1-y}\right) \;+\; x(1-y)^{\alpha-3} f''\left(\frac{x}{1-y}\right)\end{aligned} \tag{15}$$

for all $(x,y) \in \Omega$.

Let $H(s,t) = \left(\frac{t-st}{1-st}\,,\ \frac{s-st}{1-st}\right)$ for $(s,t) \in \Omega$. Then H is a diffeomorphism of Ω onto itself and $H^{-1}(x,y) = \left(\frac{y}{1-x}\,,\ \frac{x}{1-y}\right)$ for all $(x,y) \in \Omega$. It follows that if in (15) we replace $\frac{y}{1-x}$ by s and $\frac{x}{1-y}$ by t and simplify we obtain the equivalent equation

$$(1-s)^{2-\alpha}[(1-\alpha)f'(s)+sf''(s)] \\ =(1-t)^{2-\alpha}[(1-\alpha)f'(t)+tf''(t)],\ (s,t)\in\Omega. \tag{16}$$

By differentiating (16) with respect to s we conclude that there exists $\lambda\in\mathbb{C}$ such that

$$sf''(s)+(1-\alpha)f'(s)=\lambda(1-s)^{\alpha-2} \quad \text{for } 0<s<1. \tag{17}$$

If $f_0(s)=\frac{\lambda}{\alpha(\alpha-1)}\,[s^\alpha+(1-s)^\alpha]$ for $0<s<1$ then f_0 is a particular solution of (17). The associated homogenous equation is easy to solve and it follows that, for some $a,b,c\in\mathbb{C}$,

$$f(s)=as^\alpha+c(1-s)^\alpha+b \quad \text{for } 0<s<1. \tag{18}$$

But (18) and (14) imply that $b+c=0$. Thus

$$f(s)=as^\alpha+b(1-(1-s)^\alpha) \quad \text{for } 0<s<1. \tag{19}$$

Conversely, any such f satisfies (14).

Let $P_1(x,y)=x$, $P_2(x,y)=y$, $F_1(x,y)=\frac{y}{1-x}$, $F_2(x,y)=\frac{x}{1-y}$, $c_1(x,y)=(1-x)^\alpha$ and $c_2(x,y)=(1-y)^\alpha$ for $(x,y)\in\Omega$. Notice that P_1,P_2,F_1 and F_2 are submersions of Ω onto $(0,1)$ and $c_1,c_2\in C^\infty(\Omega)$.

Arguing as we did in the first example we can prove the following proposition concerning a distributional analogue of (14).

Proposition 8. *If $u\in\mathcal{D}'(0,1)$ then*

$$u\circ P_1+c_1(u\circ F_1)=u\circ P_2+c_2\ (u\circ F_2) \tag{14$'$}$$

if and only if there exist $a,b\in\mathbb{C}$ such that $u=\Lambda_f$ where f is defined by (19). If $f\in L^1_{\text{loc}}(0,1)$ and (14) holds for a.e. $(x,y)\in\Omega$ then there exist $a,b\in\mathbb{C}$ such that

$$f(x)=ax^\alpha+b(1-(1-x)^\alpha) \quad \textit{for a.e. } x\in(0,1).$$

3. With the aid of tensor products, see Chapter V of [5], it is possible to fruitfully apply our method to certain nonlinear equations. For example, consider the famous "cosine equation":

$$f(x+y)+f(x-y)=2f(x)f(y). \tag{20}$$

If we let $A(x,y) = x+y$ and $S(x,y) = x-y$ for $(x,y) \in \mathbb{R}^2$ then A and S are submersions of $\mathbb{R}^2$ onto $\mathbb{R}$ and the following is a distributional analogue of (20):

$$(20)' \qquad u \circ A + u \circ S = 2u \otimes u.$$

Suppose that $u \in \mathcal{D}'(\mathbb{R})$ and $(20)'$ holds. By applying D_1 to $(20)'$ and using the chain rule we find that

$$(Du) \circ A + (Du) \circ S = 2(Du) \otimes u.$$

Similarly

$$(Du) \circ A - (Du) \circ S = 2u \otimes (Du),$$

$$(D^2u) \circ A + (D^2u) \circ S = 2(D^2u) \otimes u$$

and

$$(D^2u) \circ A + (D^2u) \circ S = 2u \otimes (D^2u)$$

so that

$$(D^2u) \otimes u = u \otimes (D^2u).$$

Assuming that $u \neq 0$, choose $\chi_0 \in \mathcal{D}(\mathbb{R})$ such that $u(\chi_0) = 1$. Then, for all $\varphi \in \mathcal{D}(\mathbb{R})$ we have

$$D^2u(\varphi) = (D^2u \otimes u)(\varphi \otimes \chi_0) = (u \otimes D^2u)(\varphi \otimes \chi_0) = u(\varphi)D^2u(\chi_0).$$

That is, $D^2u = -\lambda^2 u$ where $\lambda^2 = -D^2u(\chi_0)$. Hence there exist $a, b \in \mathbb{C}$ such that $u = \Lambda_f$ where f is defined by $f(x) = a\cos\lambda x + b\sin\lambda x$ for all $x \in \mathbb{R}$. Since $u \circ A = \Lambda_f \circ A = \Lambda_{f\circ A}$, $u \circ S = \Lambda_{f\circ S}$ and $u \otimes u = \Lambda_{f\otimes f}$, $(20)'$ implies that (20) holds (for almost every $(x,y) \in \mathbb{R}^2$, and hence by continuity) for all $(x,y) \in \mathbb{R}^2$. It follows from well known properties of the cosine equation that $a = 1$ and $b = 0$.

Using arguments that have been illustrated in the first two examples one can deduce the following Proposition which may have some novelty.

Proposition 9. *For $f \in L^1_{\mathrm{loc}}(\mathbb{R})$, (20) holds for almost every $(x,y) \in \mathbb{R}^2$ if and only if either $f(x) = 0$ for a.e. $x \in \mathbb{R}$ or, for some $\lambda \in \mathbb{C}$, $f(x) = \cos\lambda x$ for a.e. $x \in \mathbb{R}$.*

Remarks. In all three examples, assuming that the functional equation is satisfied everywhere, the local integrability assumptions can be replaced simply by measurability according to Járai [6]; our method may be viewed as a (partial) alternative to such regularity theory. The main point is that, in certain instances, the method of reduction to differential equations may be useful, even in case the unknowns are assumed only to be locally integrable, provided differentiation is interpreted in the sense of distributions. As we have seen, by using distributions one can essentially arrive at the same conclusion by assuming only that the functional equation holds almost everywhere instead of everywhere. In many cases, as we have illustrated, functional equations have natural distributional analogues which can be solved by reduction to (distributional) differential equations. Like [8], which has provided motivation, our discussion has been somewhat algorithmic. Our main concern has been with methodology and not with substantial new theorems concerning functional equations.

References

[1] J. Aczél, The state of the second part of Hilbert's fifth problem, *Bull. Amer. Math. Soc. (N.S.)* **20** (1989), 153–163.
[2] J. A. Baker, Difference Operators, Distributions and Functional Equations, *Periodica Mathematica Hungarica* **23** no. 3 (1991), 171–183.
[3] William M. Boothby, An Introduction to Differentiable Manifolds and Riemannian Geometry, *Academic Press*, *New York*, 1975.
[4] F. G. Friedlander, Introduction to the theory of distributions, *Cambridge*, 1982.
[5] Lars Hörmander, The Analysis of Linear Partial Differential Operators I, Second edition, *Springer-Verlag*, *Berlin*, 1990.
[6] A. Járai, On regular solutions of functional equations, *Aequationes Math.* **30** (1986), 21–54.
[7] Gy. Maksa, The general solution of a functional equation related to the mixed theory of information, *Aequationes Math.* **22** (1981), 90–96.
[8] Zsolt Páles, On reduction of linear two variable functional equations to differential equations without substitutions, *Aequationes Math.* **43** (1992), 236–247.
[9] Walter Rudin, Functional Analysis, *McGraw-Hill*, *New York*, 1973.

JOHN A. BAKER
DEPARTMENT OF PURE MATHEMATICS
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
CANADA N2L 3G1

(Received October 14, 1994; revised March 1, 1995)