PARTIAL REGULARITY OF MAPPINGS BETWEEN EUCLIDEAN SPACES

BY

JAN BOMAN

The University of Stockholm, Stockholm, Sweden

1. Introduction

Let f be a locally bounded function from a p-dimensional Euclidean space E_p to a q-dimensional Euclidean space F_q. For a given subset Λ of $E_p \times F_q$ we will consider conditions on f of the following type: for each $(\xi, \eta) \in \Lambda$, $\xi \in E_p$, $\eta \in F_q$, the function $x \to \langle \eta, f(x) \rangle$ has a certain regularity property in the direction ξ. Here $\langle \cdot, \cdot \rangle$ denotes the inner product in F_q. The problem is to determine the condition on Λ in order that these conditions on f imply a corresponding (unrestricted) regularity property for the function f.

The answer to these problems is formulated in terms of the following two algebraic conditions on Λ. Let $\mathbf{R}$ denote the real numbers.

(A) *if Φ is a bilinear form $(E_p, F_q) \to \mathbf{R}$ and $\Phi(\Lambda) = 0$, then $\Phi = 0$.*

$(\hat{A})$ *if Φ is a bilinear form $(E_p, F_q) \to \mathbf{R}$ of rank 1 and $\Phi(\Lambda) = 0$, then $\Phi = 0$.*

As examples of our results we mention the following. If the regularity property is continuity or infinite differentiability, the condition $(\hat{A})$ is necessary and sufficient for an assertion of the above-mentioned type to hold. If we consider continuity of the first derivatives, the condition (A) plays the same role. If f is locally bounded and $\langle \eta, f \rangle$ is constant in the direction ξ for each $(\xi, \eta) \in \Lambda$, then it follows that f is constant if and only if (A) holds. The same assumption implies that f is a polynomial, if and only if $(\hat{A})$ holds.

If (A) holds, Λ contains at least pq elements. On the other hand, there exist subsets Λ of $E_p \times F_q$, which satisfy $(\hat{A})$ and contain only $p+q-1$ elements. If $q=1$, then (A) and $(\hat{A})$ are equivalent and mean simply that the linear hull of $\{\xi; (\xi, \eta) \in \Lambda, \eta \neq 0\}$ is equal to E_p. An analogous statement holds of course if $p=1$. Our results are trivial in case p or q is equal to one.

The above-mentioned problem becomes particularly interesting if the regularity in question is described by the modulus of continuity. Then both of the conditions (A) and

$(\hat{A})$ come into consideration. Assume that f is a locally bounded function from E_p to F_q and that $\langle \eta, f\rangle$ has modulus of continuity $\leqslant \sigma(\varepsilon)$ in the direction ξ for each $(\xi, \eta) \in \Lambda$ (see section 1 for exact definition). If Λ satisfies $(\hat{A})$, these assumptions imply that f is continuous, and that the modulus of continuity of f is $\leqslant C\hat{\sigma}(\varepsilon)$, where $\hat{\sigma}(\varepsilon)$ is related to $\sigma(\varepsilon)$ by formula (1.2), roughly $\hat{\sigma}(\varepsilon) \approx \varepsilon \int_\varepsilon^1 t^{-2}\sigma(t)\,dt$. The function $\hat{\sigma}(\varepsilon)$ is always larger than $\sigma(\varepsilon)$ and sometimes of a strictly larger order of magnitude than $\sigma(\varepsilon)$ when $\varepsilon \to 0$. If Λ satisfies (A), the same assumptions imply that f has modulus of continuity $\leqslant C\sigma(\varepsilon)$. Theorem 1 gives complete information on these questions.

Our main results deal with bounds for differences instead of local regularity properties. Assume that f is a continuous function from E_p to F_q and that any difference in the direction ξ of the function $\langle \eta, f\rangle$, i.e. any expression of the form $|\langle \eta, f(x+t\xi) - f(x)\rangle|$, is bounded by a constant C for all $t \in \mathbf{R}$, $x \in E_p$ and $(\xi, \eta) \in \Lambda$. If (A) holds we can then estimate an arbitrary difference $|f(x) - f(y)|$ by $C_1 C$, where C_1 depends only on Λ (Theorem 3 and corollaries). If Λ satisfies only the weaker condition $(\hat{A})$, we can not estimate the differences $f(x) - f(y)$, but we obtain a similar estimate for the qth order differences, where q is the dimension of F_q (Theorem 5 and corollaries). These results are used to prove Theorem 1.

1. The directional modulus of continuity

Let σ be a function defined on the non-negative reals which tends to zero at the origin.

Definition. $K(\sigma)$ denotes the set of functions $f: E_p \to F_q$ such that to every compact subset $K \subset E_p$ there exists a constant C such that for x and $x+y$ in K

$$|f(x+y) - f(x)| \leqslant C\sigma(\varepsilon), \quad \text{if} \quad |y| \leqslant \varepsilon.$$

If $0 \neq \xi \in E_p$, we denote by $K(\xi, \sigma)$ the set of locally bounded functions $f: E_p \to F_q$ such that to every compact subset $K \subset E_p$ there exists a constant C such that for x and $x + t\xi \in K$, t real

$$|f(x+t\xi) - f(x)| \leqslant C\sigma(\varepsilon), \quad \text{if} \quad |t| \leqslant \varepsilon.$$

Denote by Σ the set of all real-valued continuous subadditive and increasing functions defined on the non-negative reals and vanishing at the origin. It is easy to see that any class $K(\sigma)$ is equal to some $K(\sigma_1)$ where $\sigma_1 \in \Sigma$. In fact we can take

$$\sigma_1(\varepsilon) = \inf\{\textstyle\sum_1^n \sigma(\varepsilon_i);\ \sum_1^n \varepsilon_i \geqslant \varepsilon,\ \varepsilon_i \geqslant 0\},$$

which is the largest subadditive and increasing minorant of σ. If $\sigma, \tau \in \Sigma$, the expression $\sigma \prec \tau$ will mean that there is a constant C such that $\sigma(\varepsilon) \leqslant C\tau(\varepsilon)$ when $\varepsilon < 1$. Then $\sigma \prec \tau$, if and only if $K(\sigma) \subset K(\tau)$. If $\sigma \prec \tau$ and $\tau \prec \sigma$, we write $\sigma \approx \tau$ and say that σ is equivalent to τ.

We will often use the following simple inequality. If $a>0$, $t>0$, and $[a]$ denotes the integral part of a, then

$$\sigma(at) \leqslant \sigma(([a]+1)t) \leqslant (a+1)\sigma(t). \tag{1.1}$$

If $\sigma \in \Sigma$, we set

$$\hat{\sigma}(\varepsilon) = \varepsilon\left(1+\sigma(1)+\int_{\min(\varepsilon,1)}^{1} t^{-2}\sigma(t)\,dt\right), \quad \varepsilon>0. \tag{1.2}$$

LEMMA 1. *If $\sigma \in \Sigma$, then $\hat{\sigma} \in \Sigma$ and $\sigma \prec \hat{\sigma}$.*

Proof. We will only prove that $\lim_{\varepsilon\to 0}\hat{\sigma}(\varepsilon)=0$ and that $\sigma\prec\hat{\sigma}$. These statements follow from the inequalities

$$\sigma(\varepsilon)\,\varepsilon\int_{\varepsilon}^{1} t^{-2}dt \leqslant \varepsilon\int_{\varepsilon}^{1} t^{-2}\sigma(t)\,dt \leqslant \sigma(\delta)\,\varepsilon\int_{\varepsilon}^{\delta} t^{-2}dt + \varepsilon\int_{\delta}^{1} t^{-2}\sigma(t)\,dt, \quad \text{if} \quad 0<\varepsilon<\delta<1.$$

For example, if $\sigma(\varepsilon)=\varepsilon^b$, $0<b<1$, we get $\hat{\sigma}(\varepsilon)=2\varepsilon+(\varepsilon^b-\varepsilon)/(1-b)$ when $\varepsilon<1$, hence $\hat{\sigma}\approx\sigma$ in this case. If $\sigma(\varepsilon)=\varepsilon$ we get $\hat{\sigma}(\varepsilon)=2\varepsilon+\varepsilon\log(1/\varepsilon)$, which shows that $K(\sigma)$ is sometimes a proper subset of $K(\hat{\sigma})$.

The term $\varepsilon\sigma(1)$ in formula (1.2) is needed to make $\hat{\sigma}(\varepsilon)$ increasing, and the term ε is needed to make $\hat{\sigma}(\varepsilon)$ positive in case $\sigma(\varepsilon)$ is identically zero. The following theorem gives a complete solution to our problem in the case when the regularity is described by the modulus of continuity.

THEOREM 1. *Let $\sigma, \tau \in \Sigma$ and $\Lambda \subset E_p \times F_q$. Assume that at least one of the following two conditions holds*

$$\sigma \prec \tau \text{ and } \Lambda \text{ satisfies } (A), \tag{1.3}$$

$$\hat{\sigma} \prec \tau \text{ and } \Lambda \text{ satisfies } (\hat{A}). \tag{1.4}$$

Assume moreover that

$$\langle \eta, f\rangle \in K(\xi,\sigma) \text{ for each } (\xi,\eta)\in\Lambda. \tag{1.5}$$

Then $f\in K(\tau)$. Conversely, if (1.5) *implies that $f \in K(\tau)$, then at least one of the conditions* (1.3) *and* (1.4) *holds.*

The theorem is trivial if p or q is equal to one. In this case (1.4) implies (1.3), and hence (1.4) drops out.

A few corollaries of this result are given in Section 4. Let us here consider only one very simple but nevertheless quite illuminating example. Let $E_p=F_q=\mathbf{R}^2$ and let $\sigma(\varepsilon)=\varepsilon$. Then $\hat{\sigma}(\varepsilon)\approx\varepsilon\log(1/\varepsilon)$. The theorem implies the following. In order to prove that a function $f:\mathbf{R}^2\to\mathbf{R}^2$ belongs to $K(\sigma)$ we need at least four conditions of the type $\langle\eta, f\rangle\in K(\xi,\sigma)$. To prove that f is continuous we need three conditions. Assume that Λ consists of the following three elements: $\xi=\eta=(1,0)$, $\xi=\eta=(0,1)$, $\xi=\eta=(1,1)$. Then it is easy to see that $(\hat{A})$ holds, so that the modulus of continuity of f must be $O(\varepsilon \log (1/\varepsilon))$. Conversely, since (A) does not hold, Theorem 1 implies that no stronger assertion about the modulus of continuity of f follows from the assumptions. This fact can be seen directly as follows. Take

$$f(x) = (x_2 \log |x|,\ -x_1 \log |x|), \quad \text{where} \quad |x| = (x_1^2 + x_2^2)^{\frac{1}{2}}, \quad f(0, 0) = (0, 0).$$

Then $\langle \eta, f \rangle \in K(\xi, \sigma)$ for each $(\xi, \eta) \in \Lambda$ (in fact for every (ξ, η) such that $\xi = \eta$). For reasons of symmetry it is enough to verify this when $\xi = \eta = (1, 0)$. And this follows from the fact that $|(d/dx_1) x_2 \log |x|| = |x_1 x_2| / |x|^2 \leqslant 1$ when $|x| \neq 0$. But it is clear that the statement $f \in K(\hat{\sigma})$ can not be strengthened.

Several of our statements can be generalized if one considers, instead of the conditions (A) and $(\hat{A})$, certain hull operations P and $\hat{P}$ defined on subsets of $E_p \times F_q$, which we now define. $P\Lambda$ is the set of all $(\xi^0, \eta^0) \in E_p \times F_q$ such that every bilinear form $\Phi : (E_p, F_q) \to \mathbf{R}$ which vanishes on Λ also vanishes on (ξ^0, η^0). $\hat{P}\Lambda$ is the set of all $(\xi^0, \eta^0) \in E_p \times F_q$ such that every bilinear form Φ of rank one which vanishes on Λ also vanishes on (ξ^0, η^0). Then $P\Lambda = E_p \times F_q$ if and only if (A) holds, and $\hat{P}\Lambda = E_p \times F_q$ if and only if $(\hat{A})$ holds. The following statement is proved in the same way as Theorem 1. The conditions (1.6) and (1.7) below are equivalent:

$$\left.\begin{array}{l} f \textit{ is continuous and } \langle \eta, f \rangle \in K(\xi, \sigma) \textit{ for each } (\xi, \eta) \in \Lambda \\ \textit{implies } \langle \eta^0, f \rangle \in K(\xi^0, \tau) \end{array}\right\} \tag{1.6}$$

$$\left.\begin{array}{l} \sigma \prec \tau \textit{ and } (\xi^0, \eta^0) \in P\Lambda \quad \textit{or} \\ \hat{\sigma} \prec \tau \textit{ and } (\xi^0, \eta^0) \in \hat{P}\Lambda. \end{array}\right\} \tag{1.7}$$

2. Consequences of the stronger algebraic condition (A)

The most important consequences of (A) are deduced from the theorem on decomposition of vector valued measures (Theorem 3). In order to illustrate the condition (A) we give first a few simple consequences of (A).

The condition (A) is equivalent to the following: the set of tensor products $\xi \otimes \eta$, where $(\xi, \eta) \in \Lambda$, spans the linear space $E_p \otimes F_q$. Or equivalently, using a pair of bases, the set of $p \times q$-matrices $(\xi_i \eta_j)$, where $(\xi, \eta) \in \Lambda$, spans the linear space of all $p \times q$-matrices. To see that these conditions are equivalent to (A) we need of course only represent bilinear forms by $\Phi(\xi, \eta) = \Sigma a_{ij} \xi_i \eta_j$.

Definition. Let $0 \neq \xi \in E_p$ and $k \geqslant 1$. We denote by $C^k(\xi)$ the class of continuous functions defined in E_p such that the derivatives $D_\xi^j f(x) = [(d/dt)^j f(x + t\xi)]_{t=0}$ exist and are continuous when $j \leqslant k$. C^k is the class of k times continuously differentiable functions. $C^\infty(\xi)$ denotes $\bigcap_{k=1}^{\infty} C^k(\xi)$ and C^∞ denotes $\bigcap_{k=0}^{\infty} C^k$.

Using a standard approximation technique we can prove the following theorem.

THEOREM 2. *Assume that $\Lambda \subset E_p \times F_q$ satisfies (A) and that $f: E_p \to F_q$ satisfies*

$$\langle \eta, f \rangle \in C^1(\xi) \quad \text{for each} \quad (\xi, \eta) \in \Lambda. \tag{2.1}$$

Then $f \in C^1$. Moreover, there exists a constant B which depends only on Λ, such that

$$\|Df(x)\| \leqslant B \sup_{(\xi,\eta)\in\Lambda} |D_\xi \langle \eta, f(x) \rangle| \quad \text{for every } x \in E_p.$$

Here $\|Df(x)\|$ denotes the norm of the differential of f at x considered as an operator from E_p to F_q. Conversely, if (2.1) implies that $f \in C^1$, then (A) holds.

COROLLARY. *Assume that Λ satisfies (A), that f is continuous and that $\langle \eta, f \rangle$ is constant on all lines parallel with ξ for each $(\xi, \eta) \in \Lambda$. Then f is constant.*

In section 3 we will be able to prove the same assertion without assuming that f is continuous.

Proof of Theorem 2. We have to prove that an arbitrary first partial derivative of f exists and is continuous. Choose bases in E_p and F_q such that this derivative is $D_1 f_1$. (We use the notation $D_1 = \partial/\partial x_1$.) Since Λ satisfies (A), there exist b_ν and $(\xi^\nu, \eta^\nu) \in \Lambda$ such that

$$\sum_{\nu=1}^{n} b_\nu \xi_i^\nu \eta_j^\nu = \begin{cases} 1 \text{ when } & (i,j) = (1,1), \\ 0 \text{ when } & (i,j) \neq (1,1). \end{cases} \tag{2.2}$$

Take ψ of class C^1 with compact support such that $\int \psi \, dx = 1$ and for any $\varepsilon > 0$ set

$$f_\varepsilon(x) = \int f(x + \varepsilon y) \, \psi(y) \, dy.$$

Then for each $(\xi, \eta) \in \Lambda$, $D_\xi \langle \eta, f_\varepsilon \rangle$ converges uniformly on compact sets to $D_\xi \langle \eta, f \rangle$ when $\varepsilon \to 0$. Denoting the first component of f_ε by $(f_\varepsilon)_1$ we have by (2.2)

$$D_1 (f_\varepsilon)_1 = \sum_{\nu=1}^{n} b_\nu D_{\xi^\nu} \langle \eta^\nu, f_\varepsilon \rangle. \tag{2.3}$$

Thus $D_1 (f_\varepsilon)_1$ is continuous and converges uniformly on compact sets to some continuous function g when $\varepsilon \to 0$. Also, since f_1 is continuous, $(f_\varepsilon)_1$ converges uniformly on compact sets to f_1. Using a suitable result from elementary calculus we conclude that f_1 is differentiable with respect to x_1 and that $D_1 f_1 = g$.

We also obtain formula (2.3) with $f_\varepsilon = f$. Since the constants b_ν depend only on Λ, this proves the estimate of $\|Df(x)\|$.

We now prove the necessity of the condition (A). If (A) does not hold there exists a non-trivial bilinear form Φ such that $\Phi(\xi, \eta) = 0$ for each $(\xi, \eta) \in \Lambda$. We can represent Φ by $\Phi(\xi, \eta) = \langle B\xi, \eta\rangle$, where B is a linear operator from E_p to F_q. Take a function $\theta: \mathbf{R} \to \mathbf{R}$ of class C^1 such that $0 \leqslant \theta \leqslant 1$, $\theta(t) = 1$ when $t < \frac{1}{4}$ and $\theta(t) = 0$ when $t > \frac{1}{2}$, take δ and a such that $0 < \delta < 1$, $0 < a < 1$ and set

$$f_{\delta, a}(x) = \theta(|x/a|^\delta)\, Bx, \quad \text{for} \quad x \in E_p.$$

Then $f_{\delta, a} \in C^1$. If $\langle B\xi, \eta\rangle = 0$ we have

$$D_\xi \langle \eta, f_{\delta, a}(x)\rangle = \left[\frac{d}{dt}\,\theta(|x + t\xi|^\delta a^{-\delta})\,\langle Bx, \eta\rangle\right]_{t=0}$$

$$= \langle Bx, \eta\rangle \cdot \langle x, \xi\rangle \cdot a^{-\delta}\delta |x|^{\delta-2}\theta'(|x/a|^\delta)$$

$$\leqslant \|B\|\, |\xi|\, |\eta| \delta |x|^\delta a^{-\delta} \max |\theta'|.$$

Hence
$$|D_\xi \langle \eta, f_{\delta, a}(x)\rangle| \leqslant C\delta a^{-1},$$

where C is independent of x, δ and a. On the other hand, the differential of $f_{\delta, a}$ at $x = 0$ is equal to the operator B. Finally we note that $|f_{\delta, a}(x)| \leqslant \|B\| \cdot a$ for every x. Now choose a_ν and δ_ν, $\nu = 1, 2, \ldots$ such that $\Sigma a_\nu < \infty$ and $\Sigma \delta_\nu / a_\nu < \infty$, and set

$$f(x) = \sum_{\nu=1}^\infty f_{\delta_\nu, a_\nu}(x).$$

Then $\langle \eta, f\rangle \in C^1(\xi)$ for each (ξ, η) such that $\langle B\xi, \eta\rangle = 0$, but f is not differentiable at the origin.(1) This completes the proof of Theorem 2.

If we replace $C^1(\xi)$ by $C^k(\xi)$, $k > 1$, the assertion of Theorem 2 becomes false, even in the trivial case $q = 1$ and $p = 2$. To see this it is sufficient to observe that there exists a function $f: \mathbf{R}^2 \to \mathbf{R}$ such that $D_1^2 f$ and $D_2^2 f$ exist and are continuous and $D_1 D_2 f$ does not exist everywhere. (Such a function is $f(x) = x_1 x_2 \log|\log|x||$, $0 < |x| < 1/2$, $f(0) = 0$.)

Let $M(E_p)$ be the space of all complex-valued measures in E_p with compact support. The elements of $M(E_p)$ can be considered as linear functionals from the space $C^0(E_p, \mathbf{C})$ of complex-valued continuous functions on E_p to the complex numbers $\mathbf{C}$. Denote by $M_\xi(E_p)$ the subset of $M(E_p)$ consisting of those $\mu \in M(E_p)$ which satisfy $\mu(\varphi) = 0$ if φ is constant on all lines parallel with ξ, that is if $\varphi(x + t\xi)$ is independent of t for every x. $M_\xi(E_p)$ is clearly a linear subspace of $M(E_p)$. The total mass of μ is denoted $\|\mu\|$. We

(1) A simpler example is $f(x) = (Bx) \log|\log|x||$, $0 < |x| < 1/2$, $f(0) = 0$.

have $\|\mu\| = \sup\{|\mu(\varphi)|; \sup|\varphi| \leq 1\}$. The support of μ is denoted $\operatorname{supp}\mu$. Finally we form the tensor product $F_q \otimes M(E_p)$ of the spaces F_q and $M(E_p)$ considered as vector spaces over the field of real numbers. Let $C^0(E_p, F_q)$ denote the real linear space of continuous functions from E_p to F_q. To any element $\eta \otimes \mu \in F_q \otimes M(E_p)$ is associated in a canonical way a linear functional from $C^0(E_p, F_q)$ to $\mathbf{C}$ as follows: $\eta \otimes \mu(f) = \mu(\langle \eta, f \rangle)$.

We can now state the main result of this section.

THEOREM 3. *Assume that* $\Lambda \subset E_p \times F_q$ *and that* $(\xi^0, \eta^0) \in P\Lambda$. *Then there exists a constant* C, *which depends only on* Λ, *with the following properties. For each* $\mu_0 \in M_{\xi^0}(E_p)$ *such that* $\operatorname{supp}\mu_0 \subset B_r = \{x; |x| \leq r\}$, *there exist* $(\xi^\nu, \eta^\nu) \in \Lambda$ *and* $\mu_\nu \in M_{\xi^\nu}(E_p)$, $\nu = 1, ..., n$, *such that* $\operatorname{supp}\mu_\nu \subset B_{Cr}$ *for each* ν, $\sum_{\nu=1}^n \|\mu_\nu\| \leq C\|\mu_0\|\,|\eta^0|$ *and*

$$\eta^0 \otimes \mu_0 = \sum_{\nu=1}^{n} \eta^\nu \otimes \mu_\nu. \tag{2.4}$$

Conversely, if for each $\mu_0 \in M_{\xi^0}(E_p)$ *there exist* $(\xi^\nu, \eta^\nu) \in \Lambda$ *and* $\mu_\nu \in M_{\xi^\nu}(E_p)$ *such that* (2.4) *holds, then* $(\xi^0, \eta^0) \in P\Lambda$.

By introducing some additional notation we can give Theorem 3 a condensed formulation. To any subset Γ of $F_q \otimes E_p$ we associate a subset $\mathfrak{M}_\Gamma$ of $F_q \otimes M(E_p)$ as follows:

$$\mathfrak{M}_\Gamma = \bigcup \{\eta \otimes M_\xi(E_p); \eta \otimes \xi \in \Gamma\} = \{\eta \otimes \mu; \mu \in M_\xi(E_p), \eta \otimes \xi \in \Gamma\}.$$

Let the letter L denote linear hull. Then the main part of Theorem 3 reads: $\mathfrak{M}_{L(\Gamma)} \subset L(\mathfrak{M}_\Gamma)$.

A measure $\mu \in M(E_p)$ is called real, if $\mu(\varphi)$ is real for every real-valued function φ. It is obvious that we can take all the μ_ν in (2.4) real, if μ_0 is real.

In the applications of Theorem 3 the measure μ_0 will be a difference measure μ_0: $\varphi \to \varphi(x + t\xi^0) - \varphi(x)$. It is essential to note that even if μ_0 is a discrete measure, there may not exist discrete measures μ_ν such that (2.4) holds. This fact can be proved as follows. Let $f: \mathbf{R}^2 \to \mathbf{R}^2$, set $\xi^1 = (1, 0)$, $\xi^2 = (0, 1)$, $\xi^3 = (1, 1)$ and set $\Lambda = \{(\xi^\nu, \xi^\nu); \nu = 1, 2, 3\}$. Then $P\Lambda = \{(\theta, t\theta); \theta \in \mathbf{R}^2, t \in \mathbf{R}\}$. Define μ_0 by $\mu_0(\varphi) = \varphi(\xi^0) - \varphi(0)$. Assume that there exist discrete measures $\mu_\nu \in M_{\xi^\nu}(E_p)$ such that

$$\xi^0 \otimes \mu_0 = \sum_{\nu=1}^{3} \xi^\nu \otimes \mu_\nu. \tag{2.5}$$

Let h be an arbitrary (possibly non-measurable) solution of the functional equation $h(s+t) = h(s) + h(t)$ and set $f = (f_1, f_2) = (h(x_2), -h(x_1))$. If μ is discrete, any function is μ-measurable. The functions $\langle \xi^1, f \rangle = h(x_2)$, $\langle \xi^2, f \rangle = -h(x_1)$ and $\langle \xi^3, f \rangle = h(x_2) - h(x_1) = h(x_2 - x_1)$ are constant on lines parallel with ξ^1, ξ^2 and ξ^3 respectively. Thus the right-hand side of (2.5) is equal to zero. Using the fact that $h(0) = 0$, we see that the left-hand side of (2.5) is equal to

$$(\xi^0 \otimes \mu_0)(f) = \mu_0(\xi_1^0 h(x_2) - \xi_2^0 h(x_1)) = \xi_1^0 h(\xi_2^0) - \xi_2^0 h(\xi_1^0).$$

However, if the last expression were equal to zero for every (ξ_1^0, ξ_2^0), the function h would necessarily be linear. But this gives a contradiction, since we know that the equation $h(s+t) = h(s) + h(t)$ has non-measurable solutions. Thus for some ξ^0 there does not exist discrete measures μ_ν satisfying (2.5).

The applications of Theorem 3 depend on the following simple lemma.

LEMMA 2. *Assume that φ: $E_p \to \mathbf{R}$ is continuous, that $\mu \in M_\xi(E_p)$ and that* $\operatorname{supp} \mu \subset B_r = \{x;\ |x| \leqslant r\}$. *Then*

$$|\mu(\varphi)| \leqslant \|\mu\| \max \{|\varphi(x) - \varphi(x+t\xi)|;\ x \in B_r,\ x+t\xi \in B_r,\ t \in \mathbf{R}\}.$$

Proof. Let $x_1, ..., x_p$ be the components of x with respect to an orthogonal basis such that $\xi = (1, 0, ..., 0)$. Define φ_0: $E_p \to \mathbf{R}$ by

$$\varphi_0(x_1, ..., x_p) = \varphi(0, x_2, ..., x_p). \text{ Since } \mu \in M_\xi(E_p), \text{ we have } \mu(\varphi_0) = 0.$$

Thus $$|\mu(\varphi)| = |\mu(\varphi - \varphi_0)| \leqslant \|\mu\| \max \{|\varphi(x) - \varphi_0(x)|;\ x \in B_r\}.$$

But $\varphi_0(x)$ can of course be written $\varphi(x+t\xi)$ for each x. This proves the statement.

COROLLARY 1. *Assume that $(\xi^0, \eta^0) \in P\Lambda$, let f be a continuous function defined in $B_r = \{x;\ x \in E_p,\ |x| \leqslant r\}$ with values in F_q, and assume that*

$$|\langle \eta, f(x+t\xi) - f(x) \rangle| \leqslant C_1, \quad \textit{if } (\xi, \eta) \in \Lambda,\ x \in B_r,\ x+t\xi \in B_r,\ t \in \mathbf{R}.$$

Then there exist constants $\delta > 0$ and C_2 which depend only on Λ, such that

$$|\langle \eta^0, f(x+t\xi^0) - f(x) \rangle| \leqslant C_1 C_2 |\eta^0|, \quad \textit{if } x \in B_{\delta r},\ x+t\xi^0 \in B_{\delta r},\ t \in \mathbf{R}.$$

Proof. Fix $x \in B_{\delta r}$ and $x+t\xi^0 \in B_{\delta r}$ and define μ_0 by $\mu_0(\varphi) = \varphi(x+t\xi^0) - \varphi(x)$. Then $\operatorname{supp} \mu_0 \subset B_{\delta r}$ and $\|\mu_0\| = 2$. By Theorem 3 we can find C depending only on Λ and $(\xi^\nu, \eta^\nu) \in \Lambda$ and $\mu_\nu \in M_{\xi^\nu}(E_p)$ such that $\operatorname{supp} \mu_\nu \subset B_{C\delta r}$, $\Sigma_\nu \|\mu_\nu\| \leqslant 2C|\eta^0|$ and (2.4) holds. Using Lemma 2 we obtain, if $\delta < C^{-1}$

$$\begin{aligned}
&|\langle \eta^0, f(x+t\xi^0) - f(x) \rangle| \\
&\quad = |\mu_0(\langle \eta^0, f \rangle)| = |\textstyle\sum_\nu \mu_\nu(\langle \eta^\nu, f \rangle)| \\
&\qquad \leqslant \textstyle\sum_\nu \|\mu_\nu\| \sup \{|\langle \eta, f(x+t\xi) - f(x) \rangle|;\ (\xi, \eta) \in \Lambda,\ x \in B_r,\ x+t\xi \in B_r\} \\
&\qquad \leqslant 2C|\eta^0| C_1.
\end{aligned}$$

If we use Theorem 5 we can now prove part of Theorem 1.

COROLLARY 2. *Let* $\sigma \in \Sigma$, *let* $f: E_p \to F_q$ *and let* Λ *be a subset of* $E_p \times F_q$ *satisfying* (A). *Assume that* $\langle \eta, f \rangle \in K(\xi, \sigma)$ *for each* $(\xi, \eta) \in \Lambda$. *Then* $f \in K(\sigma)$.

Proof. Since (A) implies $(\hat{A})$ Corollary 3 of Theorem 5 shows that f is continuous. Corollary 1 then shows that on compact sets

$$|f(x+y)-f(x)| \leqslant C\sigma(\varepsilon), \quad \text{if } |y| \leqslant \delta\varepsilon,$$

where δ is the constant in Corollary 1. Since σ is subadditive, this implies that $f \in K(\sigma)$.

The remainder of this section is devoted to a proof of Theorem 3.

Denote by $\hat{\mu}$ the Fourier transform of μ:

$$\hat{\mu}(z) = \mu(e^{-i\langle x, z \rangle}), \quad z \in E_p.$$

If $\mu \in M_\xi(E_p)$, we clearly have $\hat{\mu}(z)=0$ for each z such that $\langle z, \xi \rangle = 0$. Of course the converse also holds, i.e. if $\hat{\mu}(z)=0$ on the hyperplane $\langle z, \xi \rangle = 0$, then $\mu \in M_\xi(E_p)$. To see this we need only observe that a continuous function φ: $E_p \to \mathbf{R}$ such that $\varphi(x+t\xi)$ is independent of t for every x can be approximated uniformly on compact sets by linear combinations of exponential functions $e^{i\langle z, x \rangle}$ such that $\langle z, \xi \rangle = 0$. The Fourier transform of an arbitrary element of $F_q \otimes M(E_p)$ is defined by $\widehat{\eta \otimes \mu} = \eta \otimes \hat{\mu}$ and linear extension to all of $F_q \otimes M(E_p)$. The result is an element of $F_q \otimes A$, where A denotes the set of complex-valued real analytic functions on E_p. The set of Fourier transforms of elements in $M(E_p)$ is denoted $\hat{M}(E_p)$.

We can now prove the necessity of the condition $(\xi^0, \eta^0) \in P\Lambda$ in Theorem 3 as follows. Take a measure $\mu_0 \in M_{\xi^0}(E_p)$ such that $\hat{\mu}(z) = \langle \xi^0, z \rangle + o(z)$ (e.g. $\mu(\varphi) = i(\varphi(\xi^0) - \varphi(0))$). Assume that the equality (2.4) holds, take the Fourier transform of both members, and finally single out the linear part of each term. Since $\hat{\mu}_\nu(z)=0$ when $\langle z, \xi^\nu \rangle = 0$, the linear part of $\hat{\mu}_\nu$ must also vanish in the plane $\langle z, \xi^\nu \rangle = 0$, i.e. the linear part can be written $b_\nu \langle z, \xi^\nu \rangle$ with some complex constant b_ν. Thus we obtain the relation

$$\eta^0 \otimes \langle z, \xi^0 \rangle = \sum_{\nu=1}^{n} \eta^\nu \otimes b_\nu \langle z, \xi^\nu \rangle, \quad z \in E_p. \tag{2.6}$$

It is obvious that the existence of such constants b_ν is equivalent to $(\xi^0, \eta^0) \in P\Lambda$.

A natural attempt to prove Theorem 3 would be to take b_ν such that (2.6) holds and set

$$u_\nu(z) = b_\nu \langle z, \xi^\nu \rangle (\hat{\mu}_0(z) / \langle z, \xi^0 \rangle). \tag{2.7}$$

This expression defines an analytic function, since $\mu_0 \in M_{\xi^0}(E_p)$. Moreover,

$$\left.\begin{aligned} \sum_{\nu=1}^{n} \eta^\nu \otimes u_\nu &= \eta^0 \otimes \mu_0 \quad \text{and} \\ u_\nu(z) &= 0, \quad \text{when} \quad \langle z, \xi^\nu \rangle = 0. \end{aligned}\right\} \tag{2.8}$$

However, the trouble is that u_ν need not be the Fourier transform of a measure (it is always the Fourier transform of a distribution of order one). The idea of our proof is to modify the functions u_ν so that they become Fourier transforms of measures with compact support without losing the properties (2.8).

Assume that
$$\eta^0 \otimes \xi^0 = \sum_{j=1}^{n} b_j \eta^j \otimes \xi^j. \tag{2.9}$$

The assertion of Theorem 3 is that there exist $g_j \in \hat{M}(E_p)$ such that

$$\left.\begin{aligned} \eta^0 \otimes \hat{\mu}_0 &= \sum_{j=1}^{n} \eta^j \otimes g_j \quad \text{and} \\ g_j(z) &= 0, \quad \text{when} \quad \langle z, \xi^j \rangle = 0, j = 1, \ldots, n. \end{aligned}\right\} \tag{2.10}$$

Let the dimension of the linear hull of $\eta^1, \ldots, \eta^n$ be r. Then we can determine r of the g_j, say $g_{n-r+1}, \ldots, g_n$, from (2.10) in terms of $g_1, \ldots, g_{n-r}$. Set $n-r=m$. Then the first equation in (2.10) becomes
$$g_i = \sum_{j=1}^{m} a_{ij} g_j + a_{i0} \hat{\mu}_0, \quad i = m+1, \ldots, n. \tag{2.11}$$

Thus (2.10) is equivalent to

$$\left.\begin{aligned} \sum_{j=1}^{m} a_{ij} g_j &= -a_{i0} \hat{\mu}_0, \quad \text{when} \quad \langle z, \xi^i \rangle = 0, i = m+1, \ldots, n, \\ g_i(z) &= 0, \quad \text{when} \quad \langle z, \xi^i \rangle = 0, i = 1, \ldots, m. \end{aligned}\right\} \tag{2.12}$$

In constructing solutions g_j to (2.12) we of course have to use an assumption on a_{ij} corresponding to (2.9). This assumption can be expressed in a convenient form as follows: there exist analytic functions u_j such that (2.12) holds with g_j replaced by u_j. In fact, since the functions u_j defined by (2.7) satisfy (2.8), these functions must also satisfy (2.12). Thus our problem is to find $g_j \in \hat{M}(E_p)$ such that

$$\sum_{j=1}^{m} a_{ij} g_j = \sum_{j=1}^{m} a_{ij} u_j, \quad \text{when} \quad \langle z, \xi^i \rangle = 0, i = 1, \ldots, n, \tag{2.13}$$

where $a_{ij} = \delta_j^i$ when i and $j \leqslant m$, and a_{ij} are the same as in (2.12) when $m < i \leqslant n$.

For the existence of $g_j \in \hat{M}(E_p)$ satisfying (2.13) it is obviously necessary that the restriction of $U_i = \sum_{j=1}^{m} a_{ij} u_j$ to the plane $\langle z, \xi^i \rangle = 0$ be the Fourier transform of a measure. The functions U_i have in fact a stronger property which we will now formulate. Take a non-negative function ψ of class C^1 on the real line, vanishing when $|t| > 1$ and satisfying $\int \psi \, dt = 1$, and set $\psi_r(t) = r^{-1}\psi(t/r)$, if $r > 0$. Let w be the Fourier transform of ψ_r, and set $w_i(z) = w(\langle z, \xi^i \rangle)$. Then the function $\langle z, \xi^i \rangle w_i(z) \in \hat{M}(E_p)$ (it is actually the Fourier transform of the measure which takes φ into $(1/i)\int \varphi(t\xi)\psi_r'(t)\,dt)$. We assert that $w_i U_i \in \hat{M}(E_p)$, $i=1,\dots,n$. When $i\leqslant m$, we have $U_i = u_i$ and hence

$$w_i U_i = w_i u_i = b_i(\langle z, \xi^i\rangle w_i)(\hat{\mu}_0(z)/\langle z, \xi^0\rangle), \quad i\leqslant m, \tag{2.14}$$

which belongs to $\hat{M}(E_p)$, since each of the expressions within brackets belongs to $\hat{M}(E_p)$. When $i>m$, we have $\sum_{j=1}^m a_{ij}u_j = -a_{i0}\hat{\mu}_0 + u_i$, since u_i satisfy (2.8) and therefore must satisfy (2.11). Hence

$$w_i U_i = -a_{i0}\hat{\mu}_0 w_i + w_i u_i \in \hat{M}(E_p), \quad m<i\leqslant n. \tag{2.15}$$

If $h=\hat{\mu}$, $\mu\in M(E_p)$, we define $\|h\|$ as $\|\mu\|$. If $g, h\in\hat{M}(E_p)$, we have of course $\|gh\|\leqslant \|g\|\cdot\|h\|$. We shall need an estimate of $\|w_i U_i\|$. It is immediately seen that $\|w_i\|=1$ and that $\|\langle z, \xi^i\rangle w_i\| = (1/r)\int |\psi'(t)|\,dt = C_0/r$. It is an elementary fact that if supp $\mu_0\subset B_r$, then

$$\|\hat{\mu}_0(z)/\langle z, \xi^0\rangle\| \leqslant 2r|\xi^0|^{-1}\|\mu_0\|.$$

Thus, if supp $\mu_0\subset B_r$, then

$$\|w_i u_i\| \leqslant |b_i|\,\|\langle z, \xi^i\rangle w_i\|\,\|\hat{\mu}_0/\langle z, \xi^0\rangle\| \leqslant |b_i|2C_0|\xi^0|^{-1}\|\mu_0\|.$$

Using (2.14) and (2.15) we obtain

$$\|w_i U_i\| \leqslant (|b_i|2C_0|\xi^0|^{-1} + |a_{i0}|)\|\mu_0\|, \quad i\leqslant n.$$

The constants b_i depend on ξ^0 and η^0 as well as on Λ. However, there exists a constant C which depends only on Λ, such that for each $(\xi^0, \eta^0)\in P\Lambda$ there exist b_i and $(\xi^i, \eta^i)\in\Lambda$ such that (2.9) holds and $|b_i|\leqslant C|\xi^0|\,|\eta^0|$. Perhaps the simplest way to see this is to replace Λ by a finite subset Λ_1 such that $P\Lambda = P\Lambda_1$, which is always possible. Similarly we may take C such that $|a_{i0}|\leqslant C|\eta^0|$. Thus we obtain

$$\|w_i U_i\| \leqslant C|\eta^0|\,\|\mu_0\|, \tag{2.16}$$

with a constant C which depends only on Λ.

We shall need the following lemma.

LEMMA 3. *Let a_{ij} be real numbers, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant m$. Then there exist polynomials $P_{jk}=P_{jk}(x_1,\dots,x_n)$, $1\leqslant j\leqslant m$, $1\leqslant k\leqslant n$, of degree at most n, such that*

$$P_{jk}(x) = 0, \quad \textit{if } x_k = 0,\ 1\leqslant j\leqslant m,\ 1\leqslant k\leqslant n \tag{2.17}$$

$$\sum_{j=1}^{m}\sum_{k=1}^{n} a_{ij}P_{jk}(x)\,a_{kl} = a_{il}, \quad \text{if} \quad x_i = 1,\ 1\leqslant i\leqslant n,\ 1\leqslant l\leqslant m. \tag{2.18}$$

Using this lemma it is simple to complete the proof of Theorem 3. Define functions $p_{jk}\in\hat{M}(E_p)$ by

$$p_{jk}(z)=P_{jk}(w_1(z),\ldots,w_n(z)),\quad z\in E_p,$$

and define g_j by

$$g_j(z)=\sum_{k=1}^{n}p_{jk}(z)\,U_k(z)=\sum_{k=1}^{n}\sum_{l=1}^{m}p_{jk}(z)\,a_{kl}\,u_l(z),\quad z\in E_p.$$

Then $p_{jk}U_k\in\hat{M}(E_p)$ for every j and k, since P_{jk} is divisible by x_k and $w_k U_k\in\hat{M}(E_p)$. Moreover,

$$\sum_{j=1}^{m}a_{ij}\,g_j(z)=\sum_j\sum_k\sum_l a_{ij}\,p_{jk}(z)\,a_{kl}\,u_l(z)=\sum_{l=1}^{m}a_{il}\,u_l(z),$$

when $\langle z,\xi^i\rangle=0$ by virtue of (2.18), since $w_i(z)=1$ when $\langle z,\xi^i\rangle=0$. This proves (2.13) which has been shown to imply (2.4).

It remains to prove the estimates of $\|\mu_j\|$ and $\operatorname{supp}\mu_j$. Writing $q_{jk}=p_{jk}/w_k$, we have $p_{jk}U_k=q_{jk}w_kU_k$ and $\|p_{jk}U_k\|\leqslant\|q_{jk}\|\,\|w_kU_k\|=\|p_{jk}\|\,\|w_kU_k\|$, since $\|w_k\|=1$. An estimate of $\|w_kU_k\|$ has been given in (2.16). The norm $\|p_{jk}\|$ depends formally only on the numbers a_{ij}, i.e. on the vectors $\eta^1,\ldots,\eta^n$, but these may depend on ξ^0 and η^0. However, since we may assume that Λ is finite, it is clear that $\|p_{jk}\|$ can be estimated by a constant which depends only on Λ (actually only on $\{\eta;(\xi,\eta)\in\Lambda\}$). Using (2.16) and the definition of g_j we then get

$$\sum_{j=1}^{m}\|\mu_j\|=\sum_{j=1}^{m}\|g_j\|\leqslant mnC|\eta^0|\,\|\mu_0\|\leqslant C_1|\eta^0|\,\|\mu_0\|,$$

where C_1 depends only on Λ. To prove the estimate of $\operatorname{supp}\mu_j$ we use the well-known fact that $\operatorname{supp}\varrho_1\subset B_r$ and $\operatorname{supp}\varrho_2\subset B_s$ implies $\operatorname{supp}(\varrho_1*\varrho_2)\subset B_{r+s}$. We use the symbol $\mathfrak{F}^{-1}$ to denote the inverse Fourier transform. It suffices to find C depending only on Λ such that $\operatorname{supp}\mathfrak{F}^{-1}p_{jk}\subset B_{Cr}$. It is seen from the definition of w_i that $\operatorname{supp}\mathfrak{F}^{-1}w_i\subset B_{r|\xi^i|}$. Since the degree of each of the polynomials P_{jk} is at most n, we must then have

$$\operatorname{supp}\mathfrak{F}^{-1}p_{jk}\subset B_{nrc},$$

where $c=\sup\{|\xi|;(\xi,\eta)\in\Lambda\}$, and Λ is again assumed to be finite. The proof of Theorem 3 is complete.

Proof of Lemma 3. Let $T=(t_{ij})$ be a non-singular $m\times m$ matrix with inverse $S=(s_{ij})$. If the matrix $P=(P_{jk})$ of polynomials and the matrix $A=(a_{ij})$ of real numbers satisfy (2.17) and (2.18), it is easily seen that the matrices

$$P'=SP=\Big(\sum_{j=1}^{m}s_{ij}P_{jk}\Big)\quad\text{and}\quad A'=AT=\Big(\sum_{j=1}^{m}a_{ij}t_{jk}\Big)$$

satisfy the same relations. We may choose T so that the last row in the matrix A' becomes $(0,\ldots,0,1)$. We assume that such a transformation has already been made, so that $(a_{n1},\ldots,a_{nm})=(0,\ldots,0,1)$.

We will prove the lemma by induction on the number n. When $n=1$ the statement is trivial. We assume that the statement is proved for $n-1$. Then there exist polynomials $Q_{jk}=Q_{jk}(x_1, \ldots, x_{n-1})$, $j\leqslant m$, $k\leqslant n-1$, of degree at most $n-1$, such that $Q_{jk}=0$ if $x_k=0$, and

$$\sum_{j=1}^{m}\sum_{k=1}^{n-1} a_{ij}Q_{jk}a_{kl}=a_{il}, \quad \text{if} \quad x_i=1, i\leqslant n-1, l\leqslant m.$$

Setting $Q_{jn}=0$ when $j=1, \ldots, m$ we note that the Q_{jk} satisfy (2.17) and (2.18) for $i\leqslant n-1$.

By the induction hypothesis we can also find polynomials $R_{jk}=R_{jk}(x_1, \ldots, x_{n-1})$, $j\leqslant m-1$, $k\leqslant n-1$, of degree at most $n-1$, such that $R_{jk}=0$ when $x_k=0$, and

$$\sum_{j=1}^{m-1}\sum_{k=1}^{n-1} a_{ij}R_{jk}a_{kl}=a_{il}, \quad \text{if} \quad x_i=1, i\leqslant n-1, l\leqslant m-1. \tag{2.19}$$

Take $R_{m1}=\ldots=R_{mn-1}=0$, $R_{mn}=1$, and

$$R_{jn}=-\sum_{k=1}^{n-1}R_{jk}a_{km}, \quad \text{when} \quad j\leqslant m-1.$$

Then, since $a_{nm}=1$, we have

$$\sum_{k=1}^{n}R_{jk}a_{km}=\sum_{k=1}^{n-1}R_{jk}a_{km}+R_{jn}\cdot 1=0, \quad \text{if} \quad j\leqslant m-1. \tag{2.20}$$

We claim that (2.18) holds with $P_{jk}=R_{jk}$. (Note that R_{jk} does not satisfy (2.17), since $R_{mn}=1$.) First let $l<m$. Then $R_{jk}a_{kl}=0$ if $k=n$ or $j=m$, and hence the assertion follows from (2.19) in this case. When $l=m$ we use (2.20) and obtain

$$\sum_{j=1}^{m}\sum_{k=1}^{n} a_{ij}R_{jk}a_{km}=\sum_{j=1}^{m-1}0+a_{im}\sum_{k=1}^{n}R_{mk}a_{km}=a_{im}$$

for every x and $i\leqslant n$.

Define the polynomials $P_{jk}=P_{jk}(x_1, \ldots, x_n)$ of degree at most n by

$$P_{jk}=(1-x_n)Q_{jk}+x_nR_{jk}, \; j\leqslant m, \; k\leqslant n.$$

Then it is easily seen that P_{jk} satisfy (2.17) and (2.18). This completes the proof of the lemma.

3. Consequences of the weaker algebraic condition ($\hat{\text{A}}$)

We begin by proving a statement concerning infinite differentiability. Using the Fourier transform we reduce this statement to a very simple statement concerning the rate of decrease of a function $E_p\to F_q$ at infinity. The proof of the corresponding result for continuity (Corollary 3 of Theorem 5) is more difficult. This result, including the sharp estimate for the modulus of continuity, will be deduced from a theorem on decomposition of measures analogous to Theorem 3.

THEOREM 4. *Assume that* $\Lambda \subset E_p \times F_q$ *satisfies* $(\hat{A})$ *and that* $f: E_p \to F_q$ *satisfies*

$$\langle \eta, f \rangle \in C^\infty(\xi) \text{ for each } (\xi, \eta) \in \Lambda. \tag{3.1}$$

Then $f \in C^\infty$. *Conversely, if* (3.1) *implies that* $f \in C^\infty$, *then* $(\hat{A})$ *holds.*

LEMMA 4. *Assume that* Λ *is a finite subset of* $E_p \times F_q$ *satisfying* $(\hat{A})$ *and that* s *is a natural number. Then there exists a constant* C *which is independent of* u *och* v *such that*

$$|v|\,|u|^s \leqslant C \sum_{(\xi,\eta)\in\Lambda} |\langle \eta, v\rangle|\,|\langle \xi, u\rangle|^s, \quad u \in E_p, v \in F_q. \tag{3.2}$$

Proof. The function on the right-hand side is positively homogeneous with respect to v of degree 1 and with respect to u of degree s, and so is the left-hand side. Thus it is sufficient to prove (3.2) when $|v| = |u| = 1$. But if $|u| = |v| = 1$ the bilinear form $\Phi(\xi, \eta) = \langle u, \xi\rangle\langle v, \eta\rangle$ can not vanish on all of Λ in view of the condition $(\hat{A})$. Hence the continuous function on the right-hand side of (3.2) must have a positive lower bound on the compact set $|u| = |v| = 1$. This completes the proof.

Proof of Theorem 4. Take $\psi: F_q \to \mathbf{R}$, $\psi \in C^\infty$ such that $\psi = 1$ on some open set and $\psi = 0$ outside some compact set. It will be enough to prove that $g = \psi f \in C^\infty$. It is clear that $\langle \eta, g\rangle \in C^\infty(\xi)$ for each $(\xi, \eta) \in \Lambda$. Since g has compact support we can form the Fourier transform $\hat{g}$ of g. We may assume that Λ is finite. By partial integration we obtain in a well-known way for any natural number s a constant C_s such that

$$|\langle \eta, \hat{g}(z)\rangle|\,|\langle \xi, z\rangle|^s \leqslant C_s,$$

for every $(\xi, \eta) \in \Lambda$ and $z \in E_p$. Applying Lemma 4 we then obtain

$$|\hat{g}(z)|\,|z|^s \leqslant CC_s, \quad z \in E_p,\ s = 1, 2, \ldots .$$

It is well known that this implies that $g \in C^\infty$. To prove the converse statement we assume that there exist $u \neq 0$ and $v \neq 0$ such that $\langle u, \xi\rangle\langle v, \eta\rangle = 0$ for each $(\xi, \eta) \in \Lambda$. Then, if h is an arbitrary continuous function $\mathbf{R} \to \mathbf{R}$, the function $f: E_p \to F_q$ defined by $f(x) = vh(\langle u, x\rangle)$ satisfies $D_\xi\langle \eta, f\rangle = 0$ for every $(\xi, \eta) \in \Lambda$, and hence $\langle \eta, f\rangle \in C^\infty(\xi)$ for every $(\xi, \eta) \in \Lambda$. But of course f does not in general belong to C^∞.

We now turn to the decomposition theorem for measures, which is analogous to Theorem 3 but valid under the weaker assumption $(\xi^0, \eta^0) \in \hat{P}\Lambda$. The symbol $\hat{P}$ is defined at the end of section 1.

We will denote the convolution of μ and $\nu \in M(E_p)$ by $\mu * \nu$. The convolution can be defined e.g. by

$$\mu * \nu(\varphi) = \mu_y(\nu_x(\varphi(x+y))),$$

where the subscripts have an obvious meaning. If k is a natural number we write μ^{*k} to denote the convolution $\mu * \mu * \ldots * \mu$ (k factors). We will frequently use the fact that $\widehat{\mu * \nu} = \hat{\mu}\hat{\nu}$. This shows that $M_\xi(E_p)$ is an ideal in the ring $M(E_p)$.

THEOREM 5. *Assume that* $\Lambda \subset E_p \times F_q$ *and that* $(\xi^0, \eta^0) \in \hat{P}\Lambda$. *Then there exists a constant* C, *which depends only on* Λ, *with the following properties. For each* $\mu_0 \in M_{\xi^0}(E_p)$ *such that* $\operatorname{supp} \mu_0 \subset B_r = \{x;\ |x| \leqslant r\}$ *there exists* $(\xi^\nu, \eta^\nu) \in \Lambda$ *and* $\mu_\nu \in M_{\xi^\nu}(E_p)$, $\nu = 1, \ldots, n$, such *that* $\operatorname{supp} \mu_\nu \subset B_{Cr}$, $\sum_{\nu=1}^n \|\mu_\nu\| \leqslant C\|\mu_0\|^q |\eta^0|$, *and*

$$\eta^0 \otimes \mu_0^{*q} = \sum_{\nu=1}^n \eta^\nu \otimes \mu_\nu. \tag{3.3}$$

If μ_0 *is discrete, we can take all the measures* μ_ν *discrete. Conversely, if for each* $\mu_0 \in M_{\xi^0}(E_p)$ *there exist* $(\xi^\nu, \eta^\nu) \in \Lambda$ *and* $\mu_\nu \in M_{\xi^\nu}(E_p)$ *such that* (3.3) *holds, then* $(\xi^0, \eta^0) \in \hat{P}\Lambda$.

For the proof we need the following lemma, which is precisely the assertion of Theorem 5 in the case where the dimension q is equal to 1. Denote by $L\{\xi^1, \ldots, \xi^n\}$ the linear hull of $\{\xi^1, \ldots, \xi^n\}$.

LEMMA 5. *Assume that* $\xi^0 \in L\{\xi^1, \ldots, \xi^n\}$, $\xi^\nu \in E_p$, *and that* $\mu_0 \in M_{\xi^0}(E_p)$, *and* $\operatorname{supp} \mu_0 \subset B_r = \{x;\ |x| \leqslant r\}$. *Then there exist a constant* C, *which depends only on* $\xi^1, \ldots, \xi^n$, *and measures* $\mu_\nu \in M_{\xi^\nu}(E_p)$ *such that*

$$\|\mu_\nu\| \leqslant 2\|\mu_0\| \text{ and } \operatorname{supp} \mu_\nu \subset B_{Cr} \text{ for each } \nu, \text{ and } \mu_0 = \sum_{\nu=1}^n \mu_\nu. \tag{3.4}$$

If μ_0 *is discrete, we may choose* μ_ν *discrete.*

If $\Gamma \subset E_p$, set $M_\Gamma = \bigcup \{M_\xi(E_p);\ \xi \in \Gamma\}$. Denoting again the linear hull by L we can formulate one part of Lemma 5 as follows: $M_{L(\Gamma)} \subset L(M_\Gamma)$.

Proof of Lemma 5. Note that Theorem 3 implies all the assertions of the lemma except the fact that μ_ν can be taken discrete if μ_0 is discrete. We will nevertheless give a direct proof here of all the assertions of the lemma. We may assume that $\xi^1, \ldots, \xi^n$ are linearly independent. If $n < p$ we take $\xi^{n+1}, \ldots, \xi^p$ such that $\xi^1, \ldots, \xi^p$ form a basis for E_p. Let P_m, $m = 1, \ldots, p$, be the linear operator $E_p \to E_p$ which takes $\sum_{\nu=1}^p c_\nu \xi^\nu$ into $\sum_{\nu=m}^p c_\nu \xi^\nu$, and let P_{p+1} be the 0-operator. Define μ_ν by

$$\mu_\nu(\varphi) = \mu_0(\varphi \circ P_\nu - \varphi \circ P_{\nu+1}), \quad \nu = 1, \ldots, n.$$

If $\varphi(x+t\xi^m)$ is independent of t for every x, the function $\varphi\circ P_m-\varphi\circ P_{m+1}$ is identically zero. This shows that $\mu_m\in M_{\xi^m}(E_p)$. Similarly the function $\varphi\circ P_{n+1}$ is constant in the direction ξ^0, since $\xi^0\in L\{\xi^1, ..., \xi^n\}$. This shows that $\mu_0(\varphi\circ P_{n+1})=0$. Hence

$$\mu_0(\varphi)=\mu_0(\varphi)-\mu_0(\varphi\circ P_{n+1})=\sum_{\nu=1}^{n}\mu_\nu(\varphi).$$

It is obvious that μ_ν satisfies the estimates (3.4) and that μ_ν is discrete if μ_0 is discrete.

Proof of Theorem 5. Let B be a subspace of F_q such that $\eta^0\notin B$. Denote by A the linear hull of $\{\xi; (\xi,\eta)\in\Lambda, \eta\notin B\}$. We claim that

$$\xi^0\in A, \text{ if } (\xi^0,\eta^0)\in\hat{P}\Lambda. \tag{3.5}$$

In fact, if $\xi^0\notin A$, one could choose $u\in E_p$ and $v\in F_q$ such that u annihilates A and v annihilates B, but $\langle u,\xi^0\rangle\langle v,\eta^0\rangle\neq 0$. By the definition of A one would then have $\langle u,\xi\rangle\langle v,\eta\rangle=0$ for every $(\xi,\eta)\in\Lambda$. This, however, contradicts the assumption $(\xi^0,\eta^0)\in\hat{P}\Lambda$.

Let $N=N(\eta^0,\Lambda)$ be the set of all $\mu\in M(E_p)$ such that

$$\eta^0\otimes\mu=\sum_{\nu=1}^{n}\eta^\nu\otimes\mu_\nu$$

for some $\mu_\nu\in M_{\xi^\nu}(E_p)$, and $(\xi^\nu,\eta^\nu)\in\Lambda$. Then N is a linear subset of $M(E_p)$ (N is in fact an ideal in the ring $M(E_p)$). Our assertion is that $\mu_0^{*q}\in N$, if $\mu_0\in M_{\xi^0}(E_p)$. Let $N_0=N_0(\eta^0,\Lambda)$ be the linear hull of the set of all measures ϱ of the form

$$\varrho=\varrho_1*\dots*\varrho_k, \tag{3.6}$$

where $\varrho_\nu\in M_{\xi^\nu}(E_p)$, $(\xi^\nu,\eta^\nu)\in\Lambda$, $\nu=1, ..., k$, and $\eta^0\in L\{\eta^1, ..., \eta^k\}$. Since each $M_{\xi^\nu}(E_p)$ is an ideal, N_0 is an ideal. We assert that $N_0\subset N$. Since N is a linear set, it is enough to verify that each ϱ of the form (3.6) belongs to N. But this is obvious, since $\eta^0=\sum_{\nu=1}^{k}b_\nu\eta^\nu$ for some b_ν, and hence

$$\eta^0\otimes\varrho=\sum_{\nu=1}^{n}\eta^\nu\otimes b_\nu\varrho,$$

where ϱ belongs to each $M_{\xi^\nu}(E_p)$, since $M_{\xi^\nu}(E_p)$ is an ideal.

We now assert that $M_{\xi^0}(E_p)^{*q}\subset N_0$. This will prove (3.3), since of course $\mu_0^{*q}\in M_{\xi^0}(E_p)^{*q}$. To shorten the formulas we write M_ξ instead of $M_\xi(E_p)$. We will in fact prove the following stronger statement. If $(\xi^\nu,\eta^\nu)\in\Lambda$ for $\nu=1, ..., k$ and $\eta^1, ..., \eta^k$ are linearly independent, then

$$M_{\xi^0}^{*q-k}*M_{\xi^1}*\dots*M_{\xi^k}\subset N_0. \tag{3.7}$$

We will prove (3.7) by induction on k for decreasing k. If $k=q$, then (3.7) is obvious. Next, let $k<q$. If $\eta^0\in L\{\eta^1, ..., \eta^k\}$, then (3.7) is again obvious, since in this case we have even

$$M_{\xi^1} * \ldots * M_{\xi^k} \subset N_0.$$

If $\eta^0 \notin L\{\eta^1, \ldots, \eta^k\}$, we apply (3.5) with $B = L\{\eta^1, \ldots, \eta^k\}$. Thus we find $(\theta^j, \zeta^j) \in \Lambda$, $j = 1, \ldots, s$, such that $\zeta^j \notin B$ for each j and $\xi^0 \in L\{\theta^1, \ldots, \theta^s\}$. Then for each j the $k+1$ vectors $\eta^1, \ldots, \eta^k$, ζ^j are linearly independent, and hence the induction assumption gives:

$$M_{\xi^0}^{*q-k-1} * M_{\xi^1} * \ldots * M_{\xi^k} * M_{\theta^j} \subset N_0, \quad j = 1, \ldots, s. \tag{3.8}$$

However, since $\xi^0 \in L\{\theta^1, \ldots, \theta^s\}$, we have $M_{\xi^0} \subset M_{\theta^1} + \ldots + M_{\theta^s}$ by Lemma 5, which shows that (3.8) implies (3.7). This completes the proof that $M_{\xi^0}^{*q} \subset N_0 \subset N$.

The proof shows that μ_0^{*q} is the sum of at most p^q terms of the type (3.6), where each ϱ_i arises from a decomposition of μ_0 of the type considered in Lemma 5. Thus each ϱ_i can be chosen so that $\|\varrho_i\| \leqslant 2\|\mu_0\|$ and $\operatorname{supp} \varrho_i \subset B_{Cr}$, if $\operatorname{supp} \mu_0 \subset B_r$, where C depends only on $\{\xi; (\xi, \eta) \in \Lambda\}$. Then

$$\operatorname{supp} \varrho = \operatorname{supp} \varrho_1 * \ldots * \varrho_k \subset B_{kCr} \subset B_{qCr},$$

and

$$\|\varrho\| \leqslant 2^k \|\mu_0\|^k \leqslant 2^q (1 + \|\mu_0\|^q).$$

Our proof shows that $\eta^0 \otimes \mu_0^{*q}$ is the sum of at most p^q sums of the form $\sum_{\nu=1}^q b_\nu \eta^\nu \otimes \varrho$, where ϱ is a measure of the form (3.6). To estimate $\|\mu_\nu\|$ it therefore only remains to estimate b_ν. However, it is obvious that we can take b_ν so that $|b_\nu| \leqslant C|\eta^0|$, where C depends only on the set $\{\eta; (\xi, \eta) \in \Lambda\}$. Thus we obtain the estimate

$$\|\mu_\nu\| \leqslant C(1 + \|\mu_0\|^q)|\eta^0|,$$

with a new constant C. For homogeneity reasons we must in fact have $\|\mu_\nu\| \leqslant C\|\mu_0\|^q |\eta^0|$. Finally, if μ_0 is discrete, we can take all the measures ϱ_i discrete by Lemma 5, and this of course makes all the measures μ_ν discrete as well.

We now prove the necessity of the condition $(\xi^0, \eta^0) \in \hat{P}\Lambda$. If $(\xi^0, \eta^0) \notin \hat{P}\Lambda$, we can choose u and v such that $\langle u, \xi \rangle \langle v, \eta \rangle = 0$ for every $(\xi, \eta) \in \Lambda$, and $\langle u, \xi^0 \rangle \langle v, \eta^0 \rangle \neq 0$. Choose bases so that $u = (1, 0, \ldots, 0)$, $v = (1, 0, \ldots, 0)$. Then $\xi_1 \eta_1 = 0$ for every $(\xi, \eta) \in \Lambda$ and $\xi_1^0 \eta_1^0 \neq 0$. If the formula (3.3) holds, then in particular

$$\eta_1^0 \mu_0^{*q} = \sum_{\nu=1}^n \eta_1^\nu \mu_\nu.$$

Choose for μ_0 the measure $\mu_0(\varphi) = \varphi(\xi^0) - \varphi(0)$, and take φ depending only on x_1, i.e. $\varphi(x) = h(x_1)$. Then $\eta_1^\nu \mu_\nu(\varphi) = 0$ for each ν, since either $\eta_1^\nu = 0$ or $\mu_\nu(\varphi) = 0$. But if $h(x_1) = x_1^q$, for instance, we have

$$\eta_1^0 \mu_0^{*q}(\varphi) = \eta_1^0 q!\ (\xi_1^0)^q \neq 0,$$

which gives a contradiction. The proof of Theorem 5 is complete.

Set $\Delta_\xi \varphi(x) = \varphi(x+\xi) - \varphi(x)$ and define $\Delta_\xi^k \varphi(x)$ recursively by $\Delta_\xi^j \varphi(x) = \Delta_\xi \Delta_\xi^{j-1} \varphi(x)$. Applying Theorem 5 with $\mu_0(\varphi) = \Delta_{t\xi}\varphi(x)$, t real, we immediately obtain (in the same way as Corollary 1 of Theorem 3 was deduced from Theorem 3):

COROLLARY 1. *Assume that* $(\xi^0, \eta^0) \in \hat{P}\Lambda$ *and let* f *be an arbitrary function*

$$B_r = \{x;\, x \in E_p,\ |x| \leqslant r\} \to F_q$$

such that

$$|\Delta_{t\xi}\langle \eta, f(x)\rangle| \leqslant C_1, \quad \textit{when} \quad (\xi, \eta) \in \Lambda,\ x \in B_r,\ x + t\xi \in B_r.$$

Then there exist constants $\delta > 0$ *and* C_2 *which depend only on* Λ *such that*

$$|\Delta_{t\xi^0}^q \langle \eta^0, f(x)\rangle| \leqslant C_1 C_2 |\eta^0|, \quad \textit{when} \quad x \in B_{\delta r},\ t\xi^0 \in B_{\delta r}. \tag{3.9}$$

Note that we obtain the estimate (3.9) for quite arbitrary functions, since we can choose the measures μ_ν in Theorem 5 discrete.

COROLLARY 2. *If* $\Lambda \subset E_p \times F_q$ *satisfies* $(\hat{A})$ *and if* $f\colon E_p \to F_q$ *is a locally bounded function such that* $\langle \eta, f\rangle$ *is constant on lines parallel with* ξ *for each* $(\xi, \eta) \in \Lambda$, *then* f *is a (vector valued) polynomial of degree at most* $q-1$.

Proof. Applying Corollary 1 with μ_0 equal to the difference measure $\mu_0(\varphi) = \varphi(\xi^0) - \varphi(0) = \Delta_{\xi^0}\varphi$, we find that $\Delta_{\xi^0}^q \langle \eta^0, f\rangle = 0$ for arbitrary ξ^0 and η^0. It is well known that a locally bounded function $\psi(t)$ of one variable such that $\Delta_s^q \psi(t) = 0$ for all s and t must be a polynomial of degree at most $q-1$. Hence $t \to \langle \eta^0, f(x+t\xi^0)\rangle$ is a polynomial of degree at most $q-1$ for each ξ^0 and η^0. It is easy to see that this implies that f is a polynomial of degree at most $q-1$.

Using Corollary 2 we can replace the condition that f is continuous in the corollary of Theorem 2 by the weaker condition that f is bounded.

If Λ does not satisfy $(\hat{A})$, of course no regularity property for f follows from the assumptions. In fact, relative to suitable bases in E_p and F_q we then have $\xi_1\eta_1 = 0$ for each $(\xi, \eta) \in \Lambda$, and thus any function f of the form $f(x) = (f_1(x_1), 0, \ldots, 0)$ satisfies the assumptions.

Let us say that the function f is continuous in the direction ξ, if $f(x+t\xi)$ tends to $f(x)$ uniformly on compact sets when t tends to zero.

COROLLARY 3. *Assume that* $\Lambda \subset E_p \times F_q$ *satisfies* $(\hat{A})$, *that* $f\colon E_p \to F_q$ *is locally bounded, and that* $\langle \eta, f\rangle$ *is continuous in the direction* ξ *for each* $(\xi, \eta) \in \Lambda$. *Then* f *is continuous.*

To prove Corollary 3 it remains only to estimate the first difference of a function of one variable in terms of bounds for the qth difference and for the function itself. A very sharp estimate of this kind is given in Theorem 6. However, if we only want to prove that f is continuous without caring about the modulus of continuity, we can manage with a much weaker estimate, such as the following. Let k be a natural number and let ψ be a real-valued function on a finite interval $I \subset \mathbf{R}$. Set

$$\omega_k(\psi, \varepsilon) = \sup \{ |\Delta_t^k \psi(x)|;\, x \in I,\, x + kt \in I,\, |t| \leqslant \varepsilon \}.$$

LEMMA 6. *There exist constants C and ε_0, which depend only on the length of the interval I, and a constant C_δ depending only on δ, such that*

$$\omega_1(\psi, \varepsilon) \leqslant C(C_\delta \omega_k(\psi, \varepsilon) + \delta \max |\psi|), \quad \textit{if} \quad 0 < \varepsilon < \varepsilon_0 \quad \textit{and} \quad \delta > \varepsilon. \tag{3.10}$$

This inequality shows that $\lim_{\varepsilon \to 0} \omega_1(\psi, \varepsilon) = 0$, if $\lim_{\varepsilon \to 0} \omega_k(\psi, \varepsilon) = 0$. Corollary 1 and Lemma 6 together imply Corollary 3. We do not prove Lemma 6 here. By simple considerations one can prove (3.10) with $C_\delta = \delta^{-a}$, $a > 0$. Theorem 6 implies (3.10) with $C_\delta = \log(1/\delta)$.

If we use Theorem 6 we can obtain exact information about the modulus of continuity of f.

COROLLARY 4. *If Λ satisfies $(\hat{A})$ and $\langle \eta, f \rangle \in K(\xi, \sigma)$ for every $(\xi, \eta) \in \Lambda$, then $f \in K(\hat{\sigma})$.*

Proof. We may assume that Λ contains only a finite number of elements. Then for each compact set $K \subset E_p$ we can choose C (independent of ε) such that

$$|\Delta_{t\xi} \langle \eta, f(x) \rangle| \leqslant C\sigma(\varepsilon), \quad \text{when } (\xi, \eta) \in \Lambda,\ x \in K,\ x + t\xi \in K,\ |t\xi| \leqslant \varepsilon.$$

Applying Corollary 1 with B_r equal to an arbitrary ball with radius ε contained in K we see that for any compact set $K \subset E_p$ there exist C and $\delta > 0$ (independent of ε) such that

$$|\Delta_{t\xi^0}^q \langle \eta^0, f(x) \rangle| \leqslant C\sigma(\varepsilon), \quad \text{when} \quad x \in K,\ x + t\xi^0 \in K,\ |t\xi^0| \leqslant \delta\varepsilon.$$

Let K_0 be another compact set, and assume that K is so large that $K_0 + \{x;\, |x| \leqslant 1\} \subset K$. If x is fixed in K_0, we then have an estimate of the qth difference of the function

$$h\colon t \to \langle \eta^0, f(x + t\xi^0) \rangle$$

on an interval I of length at least $2/|\xi^0|$. Thus we have with a new constant C

$$|\Delta_\varepsilon^q h(t)| \leqslant C\sigma(\varepsilon), \quad t \in I.$$

(We frequently use the fact that $\sigma(a\varepsilon) \leqslant (a+1)\sigma(\varepsilon)$ in view of the subadditivity.) By Theorem 6 with $k = 1$ this implies that $h \in K(\hat{\sigma})$. Here it is essential that $\hat{\sigma}$ is definied so that $\varepsilon \leqslant C\hat{\sigma}(\varepsilon)$ even if $\sigma(\varepsilon)=0$. Applying this result for ξ^0 and η^0 in a pair of bases for E_p and F_q we obtain $f \in K(\hat{\sigma})$.

The assumption that f is locally bounded, which is included in the definition of $K(\xi, \sigma)$, can not be omitted. This is seen from the example considered after Theorem 3. With the same notation, $\langle \xi^\nu, f\rangle$ is constant on lines parallel with ξ^ν for $\nu=1, 2, 3$, and Λ satisfies $(\hat{A})$, if $\Lambda=\{(\xi^\nu, \eta^\nu); \nu=1, 2, 3\}$. But f is not continuous if h is a non-measurable solution of $h(s+t)=h(s)+h(t)$.

Finally we make a comment on the number q in the term μ_0^{*q} in Theorem 5. Assume that $\hat{\mu}_0(z)=\langle \xi^0, z\rangle + o(z)$ and that formula (3.3) holds. Taking the Taylor expansion of degree m of the Fourier transform of both sides we obtain

$$\eta^0 \otimes \langle \xi^0, z\rangle^m = \sum_{\nu=1}^{n} \eta^\nu \otimes r_\nu(z) \langle \xi^\nu, z\rangle, \tag{3.11}$$

for some homogeneous real-valued polynomials $r_\nu(z)$ of degree $m-1$. If

$$\Lambda = \{(\xi^\nu, \eta^\nu); \nu=1, 2, \ldots, n\}$$

is fixed and m is fixed, the dimension of the linear space of expressions on the right-hand side of (3.11) is at most $n\binom{m+p-2}{m-1}$, since the number of terms in each $r_\nu(z)$ is at most $\binom{m+p-2}{m-1}$. The dimension of the linear space spanned by all the functions $\eta^0 \otimes \langle \xi^0, z\rangle^m$ for $\xi^0 \in E_p$, $\eta^0 \in F_q$ is $q\binom{m+p-1}{m}$. Thus, if for each $\xi^0 \in E_p$ and $\eta^0 \in F_q$ there exist $r_\nu(z)$ such that (3.11) holds, we must have $n\binom{m+p-2}{m-1} \geqslant q\binom{m+p-1}{m}$ or equivalently $n \geqslant q(m+p-1)/m$. However, if $m<q$ and $p>1$, then $q(m+p-1)/m > q+p-1$, which implies that $n > q+p-1$. On the other hand, for any p and q there exist Λ consisting of precisely $p+q-1$ elements such that $\hat{P}\Lambda = E_p \times F_q$ (Lemma 7). These observations prove that the number q in the term μ_0^{*q} in Theorem 5 cannot be replaced by any smaller number depending only on p and q.

We have seen that if $(\xi^0, \eta^0) \in \hat{P}\Lambda$, then there exist polynomials $r_\nu(z)$ of degree $q-1$ and $(\xi^\nu, \eta^\nu) \in \Lambda$ such that

$$\eta^0 \otimes \langle \xi^0, z\rangle^q = \sum_{\nu=1}^{n} \eta^\nu \otimes r_\nu(z) \langle \xi^\nu, z\rangle. \tag{3.12}$$

(The converse statement is of course also true.) An examination of our proof of Theorem 5 shows that its algebraic content is very close to a proof of this fact. To make this point clear we need to reformulate (3.12). Let $E_p^{\vee q} = E_p \vee \ldots \vee E_p$ (q factors) denote the symmetric tensor product of q copies of E_p (see [2] for a definition). The elements of $E_p^{\vee q}$ can be identified with homogeneous polynomials of degree q in p variables. Set $G = F_q \otimes E_p^{\vee q}$. Let G_Λ denote the set consisting of all elements of G of the form $\eta \otimes \xi \vee \xi^1 \vee \dots \vee \xi^{q-1}$, where $(\xi, \eta) \in \Lambda$ and $\xi^1, \dots, \xi^{q-1}$ are arbitrary elements of E_p. Let L denote linear hull. Then $\eta^0 \otimes (\xi^0)^{\vee q} \in L(G_\Lambda)$ means precisely the same as (3.12). It is easy to see that our proof of Theorem 5 shows that $(\xi^0, \eta^0) \in \hat{P}\Lambda$ implies $\eta^0 \otimes (\xi^0)^{\vee q} \in L(G_\Lambda)$. However, we wish to emphasize that one cannot deduce Theorem 5 from this fact. Theorem 5 depends also on the fact that—in somewhat vague terms—one can generate $L(G_\Lambda)$ from G_Λ by varying one of the ξ^ν at a time.

4. Conclusion of proof of Theorem 1

Corollary 2 of Theorem 3 and Corollary 4 of Theorem 5 together prove the first part of Theorem 1. It remains to prove the necessity of (1.3) or (1.4).

To prove that (1.3) or (1.4) holds means to prove that each of the following three statements hold

(a$_1$) $\sigma \prec \tau$

(a$_2$) Λ satisfies $(\hat{A})$

(a$_3$) $\hat{\sigma} \prec \tau$ or Λ satisfies (A).

If (1.5) implies that $f \in K(\tau)$, it is obvious that (a$_1$) holds. That (a$_2$) must hold follows from the remark following Corollary 2 of Theorem 5. It remains to prove (a$_3$). Let us assume that (A) does not hold; we then have to prove that $\hat{\sigma} \prec \tau$. We shall do this by constructing a function $f: E_p \to F_q$ such that

(b$_1$) $\langle \eta, f \rangle \in K(\xi, \sigma)$ for each $(\xi, \eta) \in \Lambda$, and

(b$_2$) $f \in K(\tau)$ implies $\hat{\sigma} \prec \tau$.

Since (A) does not hold there exists a non-trivial bilinear form $\Phi: (E_p, F_q) \to \mathbf{R}$ which vanishes on Λ. We can represent Φ in the form $\Phi(\xi, \eta) = \langle A\xi, \eta \rangle$, where A is a linear operator $E_p \to F_q$. Set

$$f(x) = (Ax/|Ax|)\hat{\sigma}(|Ax|), \quad x \in E_p.$$

Then (b$_2$) is obviously true. We claim that $\langle \eta, f \rangle \in K(\xi, \sigma)$ whenever $\langle A\xi, \eta \rangle = 0$; this will prove (b$_1$). We may assume $|Ax| < |A(x+t\xi)|$. Then

$$|\langle \eta, f(x+t\xi) - f(x) \rangle| = \left| \langle \eta, Ax \rangle \left(\frac{\hat{\sigma}(|A(x+t\xi)|)}{|A(x+t\xi)|} - \frac{\hat{\sigma}(|Ax|)}{|Ax|} \right) \right|$$

$$\leqslant |\langle \eta, Ax \rangle| \int_{|Ax|}^{|A(x+t\xi)|} s^{-2} \sigma(s)\, ds.$$

Using first the fact that σ is increasing and then that $\sigma(st) \leqslant (1+s)\sigma(t)$ we obtain, if $a<b$,

$$\int_a^b s^{-2}\sigma(s)\,ds \leqslant \sigma(b)\int_a^b s^{-2}ds = \sigma(b)\,(b-a)/(ab)$$

$$\leqslant \sigma(b-a)\left(1+\frac{b}{b-a}\right)\frac{b-a}{ab} \leqslant (2/a)\,\sigma(b-a).$$

With $a=|Ax|$, $b=|A(x+t\xi)|$, we have $b-a\leqslant |At\xi|$ and hence

$$|\langle \eta, f(x+t\xi)-f(x)\rangle| \leqslant 2|\eta|\sigma(|At\xi|) \leqslant 2|\eta|(1+|A\xi|)\sigma(|t|).$$

This completes the proof of Theorem 1.

We will now state a few consequences of Theorem 1.

There is a simple sufficient condition on Λ in order that $\hat{P}\Lambda = E_p \times F_q$. To formulate this condition we need the following definition. A finite or infinite set H of elements of a linear space is said to be *r-wise linearly independent*, if each subset of H consisting of at most r elements forms a linearly independent set. It is obvious that if $p\geqslant 2$ there are infinite subsets of E_p which are p-wise linearly independent.

LEMMA 7. *Assume that* $\Lambda = \{(\xi^\nu, \eta^\nu);\ \nu=1, \ldots, p+q-1\} \subset E_p \times F_q$, *where* $\{\xi^\nu\}$ *is p-wise and* $\{\eta^\nu\}$ *is q-wise linearly independent. Then* Λ *satisfies* $(\hat{A})$, *i.e.* $\hat{P}\Lambda = E_p \times F_q$.

Proof. Assume that $\langle u, \xi^\nu\rangle\langle v, \eta^\nu\rangle = 0$ for every ν. Then either $\langle u, \xi^\nu\rangle = 0$ for a set of p indices, or $\langle v, \eta^\nu\rangle = 0$ for a set of q indices ν. This shows that u or v must be equal to zero.

The above argument shows that $\hat{P}\Lambda \neq E_p \times F_q$ if Λ contains less than $p+q-1$ elements. If, in addition, $\{\xi^\nu\}$ is p-wise and $\{\eta^\nu\}$ is q-wise linearly independent, we can in fact prove that $\hat{P}\Lambda$ is as small as it can be, namely $\hat{P}\Lambda = \Lambda$. This statement together with Lemma 7 shows that in the case where $\{\xi^\nu\}$ is p-wise and $\{\eta^\nu\}$ is q-wise linearly independent we must have either $\hat{P}\Lambda = E_p \times F_q$ or $\hat{P}\Lambda = \Lambda$, depending on whether or not Λ contains at least $p+q-1$ elements. We will not prove this statement here.

COROLLARY 1. *Let* $\sigma \in \Sigma$ *and assume that* $\xi^\nu \in \mathbf{R}^p$, $\nu = 1, \ldots, m \geqslant p$, *form a p-wise linearly independent set of vectors. Let* g_ν *be functions from* $\mathbf{R}^p$ *to* $\mathbf{R}$ *such that* $g_\nu \in K(\xi^\nu, \sigma)$ *for each* ν. *Assume moreover that*

$$\sum_{\nu=1}^{m} c_{i\nu} g_\nu = 0, \quad i = 1, \ldots, p-1,$$

and that the $m+p-1$ *vectors in* $\mathbf{R}^m$ $(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1), (c_{1,1}, \ldots, c_{1,m}), \ldots, (c_{p-1,1}, \ldots, c_{p-1,m})$ *form an m-wise linearly independent set. Then each* $g_\nu \in K(\hat{\sigma})$.

Proof. Define $f\colon \mathbf{R}^p \to \mathbf{R}^m$ by $f=(g_1, \ldots, g_m)$ and take $\eta^\nu \neq 0$ parallel to the νth coordinate axis in $\mathbf{R}^m$ when $\nu \leqslant m$ and $\eta^\nu = (c_{\nu-m,1}, \ldots, c_{\nu-m,m})$, when $m < \nu \leqslant m+p-1$. Then for arbitrary

$\xi^{m+1}, \ldots, \xi^{m+p-1}$ we have $\langle \eta^\nu, f \rangle \in K(\xi^\nu, \sigma)$ for $\nu = 1, \ldots, m+p-1$. Hence $f \in K(\hat{\sigma})$ by Theorem 1 and Lemma 7.

In the case $p=2$ we obtain the following.

COROLLARY 2. *Let $\sigma \in \Sigma$ and let $\xi^\nu \in \mathbf{R}^2$, $\nu = 1, \ldots, m$, be pairwise linearly independent. Let g_ν be functions from $\mathbf{R}^2$ to $\mathbf{R}$ and assume that $g_\nu \in K(\xi^\nu, \sigma)$ for each ν and that $\sum_{\nu=1}^m g_\nu = 0$. Then $g_\nu \in K(\hat{\sigma})$ for each ν.*

Note that the last condition of Corollary 1 is trivially satisfied in this case, since all the coefficients $c_{1\nu}$ are equal to 1. Note also that the assertion of Corollary 2 is trivial if $m \leqslant 2$. The example considered after Theorem 1 corresponds to the case $m=3$ in Corollary 2.

In a recent paper [1] we applied a result very close to Corollary 2 (Lemma 7 in [1]). Here it was known of a continuous function $f: \mathbf{R}^p \to \mathbf{R}^p$ that $\langle \xi, f \rangle \in K(\xi, \sigma)$ for every non-zero $\xi \in \mathbf{R}^p$, and we wanted to estimate the modulus of continuity of f. The estimate that we gave in [1] was not the best possible. However, this situation is easily analysed by means of Theorem 1. Setting $\Lambda = \{(\xi, \xi); 0 \neq \xi \in \mathbf{R}^p\}$ we have of course $\hat{P}\Lambda = E_p \times F_q$, but $P\Lambda \neq E_p \times F_q$. This shows that $f \in K(\hat{\sigma})$ and that no stronger conclusion is possible.

5. The inequality for the moduli of smoothness

In the proof of Theorem 1 we used an inequality between the so-called moduli of smoothness of various orders, which was first proved by Marchaud [3]. This inequality will now be described. In order to make the paper self-contained we have included a proof here. A proof can also be found in Timan's book on approximation theory [4].

Let g be a real-valued function on a finite subinterval I of $\mathbf{R}$ and k a natural number. The kth order difference

$$\Delta_t^k g(x) = \sum_{j=0}^{k} \binom{k}{j} (-1)^{j+k} g(x+jt)$$

and the modulus of smoothness of order k

$$\omega_k(g, \varepsilon) = \sup \{ |\Delta_t^k g(x)|;\ x \in I,\ x+kt \in I,\ |t| \leqslant \varepsilon \}$$

have already been considered. When $k=1$, $\omega_k(g, \varepsilon)$ is of course the modulus of continuity of g. It is obvious that $\omega_k(g, \varepsilon) \geqslant 2^{k-n}\omega_n(g, \varepsilon)$ if $k \leqslant n$. The result of Marchaud is an estimate in the opposite direction.

THEOREM 6. *There exists a constant C which depends only on q and on the length of the interval I, such that if $1 \leqslant k \leqslant q$,*

$$\omega_k(g, \varepsilon) \leqslant C\varepsilon^k \left(\int_\varepsilon^1 t^{-k-1}\omega_q(g, t)\, dt + \sup |g| \right), \quad \varepsilon > 0. \tag{5.1}$$

Special cases of Theorem 6 have frequently been considered in the literature. For example, Zygmund considers so-called smooth functions in connection with trigonometric series [5]. A function is called uniformly smooth in an interval, if $(f(x+t)+f(x-t)-2f(x))/t$ tends to zero uniformly when t tends to zero. It is well known that such a function must have modulus of continuity $O(\varepsilon \log (1/\varepsilon))$. This simple assertion is of course implied by Theorem 6 ($k=1$, $q=2$, $\omega_2(g, \varepsilon) \leqslant C\varepsilon$).

The inequality (5.1) is closely related to the theory of best approximation. Note for example that the expression $\hat{\sigma}(\varepsilon)$ occurs in the converse of Jackson's theorem (see Timan [4], section 6.2.1).

Proof of Theorem 6. Let T_h be the translation operator defined by $T_h g(t) = g(t+h)$ and J the identity operator. Using the fact that $T_h^2 = T_{2h}$ we obtain (multiplication of operators defined as usual)

$$\begin{aligned}(T_{2h}-J)^k - 2^k(T_h-J)^k &= (T_h-J)^k((T_h+J)^k - 2^kJ)\\ &= (T_h-J)^{k+1}((T_h+J)^{k-1} + (T_h+J)^{k-2}\,2 + \dots + 2^{k-1}J).\end{aligned}$$

Since $(T_h-J)^k g = \Delta_h^k g$, this shows that for each t and h such that $t \in I$ and $t+2kh \in I$ we have

$$|\Delta_{2h}^k g(t) - 2^k \Delta_h^k g(t)| \leqslant k\, 2^{k-1}\, \omega_{k+1}(g, |h|).$$

By the triangle inequality we then obtain

$$|\Delta_h^k g(t)| \leqslant 2^{-k}\omega_k(g, |2h|) + (k/2)\omega_{k+1}(g, |h|), \quad \text{if } t \in I,\ t+2hk \in I. \tag{5.2}$$

Let $d(I)$ denote the length of I. Assume that $t \in I$, $t+hk \in I$ and that $|hk| \leqslant d(I)/3$. Then either $t-hk \in I$ or $t+2hk \in I$. Using this observation and the fact that $|\Delta_h^k g(t)| = |\Delta_{-h}^k g(t+h)|$ we obtain (5.2) for all t and h such that $t \in I$, $t+hk \in I$ and $|hk| \leqslant d(I)/3$. Hence

$$\omega_k(g, \varepsilon) \leqslant 2^{-k}\omega_k(g, 2\varepsilon) + (k/2)\omega_{k+1}(g, \varepsilon), \quad \text{if } 0 < \varepsilon < d(I)/3k. \tag{5.3}$$

Repeated use of (5.3) will now prove the theorem when $q = k+1$. Let us show this by an induction argument. Set

$$H_k(\varepsilon) = \varepsilon^k \left(\int_\varepsilon^1 s^{-k-1}\, \omega_{k+1}(g, s)\, ds + \sup |g| \right).$$

If $C > k^2$ and $\varepsilon < \frac{1}{2}$ we have, since $\omega_{k+1}(g, \varepsilon)$ is increasing,

$$(k/2)\,\omega_{k+1}(g,\varepsilon) \leqslant \omega_{k+1}(g,\varepsilon)\,C\varepsilon^k \int_\varepsilon^{2\varepsilon} s^{-k-1} ds$$

$$\leqslant C\varepsilon^k \int_\varepsilon^{2\varepsilon} s^{-k-1}\,\omega_{k+1}(g,s)\,ds \leqslant C(H_k(\varepsilon) - 2^{-k}H_k(2\,\varepsilon)). \tag{5.4}$$

Adding (5.3) and (5.4) gives

$$\omega_k(g,\varepsilon) - CH_k(\varepsilon) \leqslant 2^{-k}(\omega_k(g,2\varepsilon) - CH_k(2\varepsilon)), \quad \text{if} \quad 0<\varepsilon<\varepsilon_0, \tag{5.5}$$

where $\varepsilon_0 = \min\,(\frac{1}{2}, d(I)/3k)$. Since $\omega_k(g,\varepsilon) \leqslant 2^k \sup|g|$, it is clear that $\omega_k(g,\varepsilon) - CH_k(\varepsilon) \leqslant 0$ when $\varepsilon_0<\varepsilon<1$, if $C>(2/\varepsilon_0)^k$. This together with (5.5) proves the theorem in the special case where $q=k+1$.

Using this special case of the inequality we obtain, if $n \geqslant k+1$,

$$\int_\varepsilon^1 s^{-k-1}\,\omega_n(g,s)\,ds \leqslant \int_\varepsilon^1 s^{-k-1}\,Cs^n\left(\int_s^1 u^{-n-1}\,\omega_{n+1}(g,u)\,du + \sup|g|\right) ds$$

$$\leqslant C\int_\varepsilon^1 u^{-n-1}\,\omega_{n+1}(g,u)\int_\varepsilon^u s^{n-k-1} ds\,du + C\,\sup|g|$$

$$\leqslant C\int_\varepsilon^1 u^{-n-1}\,\omega_{n+1}(g,u)\,u^{n-k} du + C\,\sup|g|$$

$$= C\int_\varepsilon^1 u^{-k-1}\,\omega_{n+1}(g,u)\,du + C\,\sup|g|.$$

Combining this inequality with the special case already proved we obtain the general case inductively.

References

[1]. Boman, J., Differentiability of a function and of its compositions with functions of one variable. To appear in Math. Scand.

[2]. Bourbaki, N., *Algèbre, Chap.* IV, § 1, Exercise 1.

[3]. Marchaud, A., Sur les dérivées et sur les différences des fonctions de variables réelles. *J. Math. Pures Appl.*, 6 (1927), 337–425.

[4]. Timan, A. F., *Theory of approximation of functions of a real variable.* Pergamon Press, Oxford 1963.

[5]. Zygmund, A., Smooth functions. *Duke Math. J.*, 12 (1945), 47–76.

Received April 4, 1967