Journal of Pure and Applied Algebra 24 (1982) 59–71
North-Holland Publishing Company

CONTINUOUS EXTRAPOLATION TO TRIANGULAR MATRICES CHARACTERIZES SMOOTH FUNCTIONS

Stephen H. SCHANUEL*
Department of Mathematics, State University of New York at Buffalo, Buffalo, NY 14214, USA

Communicated by F.W. Lawvere
Received June 1980

Foreword

A complex function is analytic if and only if it has continuous extrapolations to all finite dimensional **C**-algebras, in a way compatible with algebra homomorphisms. Analogously, a real function is smooth (i.e. infinitely differentiable) if and only if it has compatible continuous extrapolations to all those finite-dimensional **R**-algebras A in which a^2+1 is invertible for all $a\in A$. In both cases the extrapolation is unique, and is itself analytic (resp. smooth). The requirement of continuity can be weakened to a local boundedness condition, but not abandoned. Similarly, suitable subcategories of algebras would suffice, but they must include some noncommutative algebras, since triangular matrices are the heart of the step from continuity to differentiability.

Global functions (compatible families of functions on algebras) are expressible in terms of the identity and "semisimple part" functions. An analogy between this representation and the role of z, $\bar{z}$ (or, Re(z)) among scalar functions $\mathbf{C}\to\mathbf{C}$ includes an analogue for the Cauchy–Riemann equations.

The methods are elementary, primarily classical techniques of matrix calculation.

1. Continuous extrapolation characterizes entire functions

Several methods are known for extrapolating analytic (= entire) functions $f:\mathbf{C}\to\mathbf{C}$ to various **C**-algebras A, most of them having been first discovered for the case $A=M_n(\mathbf{C})$, the algebra of $n\times n$ matrices. For instance, the formulas

$$f(a)=\frac{1}{2\pi i}\int(\zeta-a)^{-1}f(\zeta)\,d\zeta$$

* Partially supported by NSF Grant MCS-7904162.

0022-4049/82/0000–0000/\$02.75

(the integration over a circle enclosing the spectrum of a), and

$$f(a)=\sum \frac{f^{(n)}(0)}{n!}a^n$$

are useful for extrapolating to Banach algebras, while

$$f(a)=r(a)$$

(where $f(z)=p(z)q(z)+r(z)$ with p, r polynomials, q analytic, and $p(a)=0$) is helpful for finite dimensional A.

Our problem is to clarify two observations: only analytic functions seem to extrapolate in a well-behaved way; and for these the extrapolation is unique.

For the second of these observations there are at least two general explanations: one from the point of view of extrapolating all entire functions to a single algebra A, and a second from the point of view of extrapolating a single entire function to all finite dimensional algebras. Taking the first of these viewpoints, let $\mathbf{E}$ be the $\mathbf{C}$-algebra of entire functions $f:\mathbf{C}\to\mathbf{C}$, and let A^A be the $\mathbf{C}$-algebra of all functions $A\to A$, where A is a finite dimensional algebra. (In accordance with standard terminology, algebras have a unit element; and homomorphisms of algebras preserve the unit. In addition – because the interest here is in the finite dimensional case – algebra will mean finite dimensional algebra, unless the context clearly implies otherwise.) The following proposition summarizes the facts about simultaneously extrapolating all entire functions.

Proposition 1. *For each finite-dimensional* $\mathbf{C}$*-algebra* A*, there is a unique homomorphism of algebras*

$$(\ )_A:\mathbf{E}\to A^A,$$

denoted $f\mapsto f_A$*, such that the identity function goes to the identity function. Furthermore:*

(1) *the family* $\langle f_A\rangle$ *is natural with respect to algebra homomorphisms: If* $\varphi:A\to B$ *is a homomorphism of (finite-dimensional) algebras, the diagram below commutes:*

$$\begin{array}{ccc} A & \xrightarrow{f_A} & A \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\varphi} \\ B & \xrightarrow{f_B} & B; \end{array}$$

(2) $f_{\mathbf{C}}=f$;

(3) f_A *is analytic (as a function* $\mathbf{C}^n\to\mathbf{C}^n$*, after identifying* A *with* $\mathbf{C}^n$ *by any linear isomorphism).*

Proof. The mapping $\mathrm{Alg}(\mathbf{E},A)\to A$, by evaluating at the identity function, is a bijection. (Thus to finite dimensional A, $\mathbf{E}$ is indistinguishable from $\mathbf{C}[X]$, the free algebra on one generator.) To see this, recall that for any entire f and any monic polynomial p, there are entire q and polynomial r such that

$$f(z)=p(z)q(z)+r(z)$$

(induction on the degree of p; the case of degree one is $f(z)=(z-\lambda)q(z)+f(\lambda)$). Choosing p such that $p(a)=0$ shows $f\mapsto r(a)$, hence our mapping is injective, and one uses this formula to define for each a an algebra homomorphism, showing surjectivity.

Now for any set X, it follows that

$$\mathrm{Alg}(\mathbf{E}, A^X)=(\mathrm{Alg}(\mathbf{E}, A))^X\to A^X$$

is also a bijection. Specializing to $X=A$ gives for each function $g\in A^A$ a unique algebra homomorphism $\mathbf{E}\to A^A$ such that the identity function goes to g; the case $g=$ identity function on A is the main assertion of the proposition. (1) follows by replacing X by A and A by B; (2) from the observation that the inclusion $\mathbf{E}\to\mathbf{C}^{\mathbf{C}}$ is a homomorphism.

To prove (3), it is easier to pass to another extrapolation procedure, substitution into the Taylor series for f. The uniqueness shows this is the desired algebra homomorphism $\mathbf{E}\to A^A$, but simple estimates show that this f_A is analytic, concluding the proof.

f_A will be called the *canonical extrapolation* of the entire function f to the algebra A. In the useful terminology introduced by Eilenberg and MacLane, the proposition asserts that the family $\langle f_A\rangle$ is a *natural transformation* $U\to U$, where U is the 'underlying space' functor from the category of finite-dimensional $\mathbf{C}$-algebras (and homomorphisms) to the category of finite dimensional vector spaces over $\mathbf{C}$ (and analytic maps). A word should be said about the reason for insisting on naturality as the sensible requirement for regarding a family $\langle f_A\rangle$ as an extrapolation of f. It is simply the generalization of the basic idea of extrapolation: to extrapolate a function $f:\mathbf{R}\to\mathbf{R}$ to $\mathbf{C}$, for instance, is precisely to find a function $f_{\mathbf{C}}:\mathbf{C}\to\mathbf{C}$ such that the diagram below commutes, and $f_{\mathbf{R}}=f$:

$$\begin{array}{ccc} \mathbf{R} & \xrightarrow{f_{\mathbf{R}}} & \mathbf{R} \\ \downarrow & & \downarrow \\ \mathbf{C} & \xrightarrow{f_{\mathbf{C}}} & \mathbf{C}. \end{array}$$

Taking now the alternative viewpoint, extrapolation of a single function to all (finite dimensional) algebras, the claim is that continuity implies analyticity.

Theorem 2. *Let $f_A:A\to A$ be a family of continuous maps, one for each finite dimensional $\mathbf{C}$-algebra A, natural with respect to algebra homomorphisms. Then all f_A (in particular $f_{\mathbf{C}}$) are analytic, and the family $\langle f_A\rangle$ is just the one obtained by canonical extrapolation of $f_{\mathbf{C}}$.*

Proof. The naturality allows one to compute, letting T be the algebra of two-by-two upper triangular matrices,

$$F_T\begin{bmatrix}\lambda & 1\\ 0 & \mu\end{bmatrix}=\begin{bmatrix}f(\lambda) & \dfrac{f(\lambda)-f(\mu)}{\lambda-\mu}\\ 0 & f(\mu)\end{bmatrix} \quad \text{for } \lambda\neq\mu.$$

(Here we abbreviate $f_{\mathbf{C}}$ as f.) The calculation uses only the algebras $\mathbf{C}$, $\mathbf{C}\times\mathbf{C}$ and T, and naturality with respect to the two projections $\mathbf{C}\times\mathbf{C}\to\mathbf{C}$ and the homomorphism $\mathbf{C}\times\mathbf{C}\to T$ by

$$(\sigma,\tau)\to\begin{bmatrix}\sigma & \dfrac{\sigma-\tau}{\lambda-\mu}\\ 0 & \tau\end{bmatrix}.$$

Now letting λ tend to μ, the continuity of f_T guarantees that $f=f_{\mathbf{C}}$ is differentiable, hence analytic. The proof that the family $\langle f_A\rangle$ coincides with the family arising from f is equally easy. Embed A in a full matrix algebra M, and then observe that matrices with distinct eigenvalues are dense, so it suffices to show that f_M agrees with the 'extension' of $f=f_{\mathbf{C}}$ on these. But if $x\in M$ is such, the subalgebra B generated by x is isomorphic to a product of copies of $\mathbf{C}$, and the naturality with respect to the projections finishes the proof.

In Sections 3 and 4, the precise role played by the non-commutativity will be clarified, while Section 5 shows that the hypothesis of continuity can be considerably weakened.

2. Continuous extrapolation characterizes smooth functions

Theorem 2 immediately implies a corresponding but rather uninteresting result for $\mathbf{R}$-algebras. Any family $\langle f_A\rangle$ defined for all finite dimensional $\mathbf{R}$-algebras, natural with respect to all $\mathbf{R}$-algebra homomorphisms, is in particular so for $\mathbf{C}$-algebras, hence comes from an entire function, which is then forced also, by naturality with respect to the inclusion $\mathbf{R}\to\mathbf{C}$, to carry $\mathbf{R}$ to itself. Thus one gets a bijection from the set of entire functions commuting with complex conjugation to the set of natural transformations from the underlying space functor on finite dimensional $\mathbf{R}$-algebras to itself.

More interesting is to restrict further the class of $\mathbf{R}$-algebras on which our function is required to live, and there is a sensible way to do this. Call a finite-dimensional $\mathbf{R}$-algebra A *triangulable* if it satisfies any (and hence all) of the following equivalent conditions:

(i) A can be embedded in an algebra of triangular matrices,

(ii) For every $a\in A$, a^2+1 is invertible in A,

(iii) For every $a \in A$, there are real scalars $\lambda_1, \dots, \lambda_m$ with

$$\prod (a - \lambda_i) = 0.$$

(iv) For each $a \in A$, there is a homomorphism of **R**-algebras

$$C^\infty(\mathbf{R}) \to A$$

such that the identity function goes to a.

The equivalence, in the order (iv) ⇒ (ii) ⇒ (i) ⇒ (iii) ⇒ (iv), is reasonably straightforward. For (ii) ⇒ (i), observe that (ii) is inherited on $A/\mathrm{rad}\,A$; but then $A/\mathrm{rad}\,A$, as a product of matrix algebras over reals, complexes, and quaternions, must necessarily be just a product of copies of **R**. Hence every simple module is one-dimensional, from which (i) follows.

Proposition 1*. *For each finite-dimensional triangulable* **R**-*algebra* A, *there is a unique homomorphism of algebras*

$$(\ )_A : C^\infty(\mathbf{R}) \to A^A$$

such that the identity function goes to the identity function. Furthermore:

(1) *The family* $\langle f_A \rangle$ *is natural with respect to algebra homomorphisms.*
(2) $f_{\mathbf{R}} = f$.
(3) f_A *is smooth* (*i.e.* C^∞ *as a function* $\mathbf{R}^n \to \mathbf{R}^n$).

Proof. Except for (3), the proof goes exactly as in the complex case, restricting p to range over monic polynomials with all roots real. For (3), we need a rather stronger version of the divisibility of smooth functions by linear polynomials, Hadamard's Lemma. Let $g : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}$ be C^∞. Then there is a (unique) C^∞-function $h : \mathbf{R}^2 \times \mathbf{R}^n \to \mathbf{R}$ such that

$$g(x, v) = (x - \lambda) h(x, \lambda, v) + g(\lambda, v).$$

(This is proved by writing $g(x, v) - g(\lambda, v) = \int_\lambda^x g_1(t, v)\,\mathrm{d}t$.) Then by induction one has, for each smooth $f : \mathbf{R} \to \mathbf{R}$,

$$f(x) = \left(\prod (x - \lambda_i)\right) g(x, \lambda_1, \dots, \lambda_n) + r(x),$$

where r is a polynomial in x with coefficients smooth functions of $(\lambda_1, \dots, \lambda_n)$. This is enough to finish the proof in the case $A = T_n$, the algebra of $n \times n$ triangular matrices, since the eigenvalues $\lambda_1, \dots, \lambda_n$ of $a \in A$ are smooth functions of a, being just the diagonal entries. The general case reduces to this case by naturality, since every triangulable A embeds in some T_n. (This trick avoids the necessity of a 'smooth elementary symmetric functions theorem'; such a theorem exists [3], but is more difficult.)

The analogue of Theorem 2 holds, but of course the proof must be altered, since differentiability does not imply smoothness.

Theorem 2*. *Let $f_A : A \to A$ be a family of continuous maps, one for each triangulable* **R***-algebra A, natural with respect to algebra homomorphisms. Then all f_A are smooth, and $\langle f_A \rangle$ is the canonical extrapolation of $f_{\mathbf{R}}$.*

Proof. Let $\mathbf{R}[\varepsilon]$ be the algebra of dual numbers ($\varepsilon^2 = 0$). Then if A is triangulable, so is $A \otimes_{\mathbf{R}} \mathbf{R}[\varepsilon]$. Thus one can define the *formal derivative* $\langle g_A \rangle$ of the natural transformation $\langle f_A \rangle$ by composing

$$A \to A \otimes \mathbf{R}[\varepsilon] \to A \otimes \mathbf{R}[\varepsilon] \to A$$

to get g_A, where the maps are given by

$$a \mapsto a \otimes 1 + 1 \otimes \varepsilon, \qquad f_{A \otimes \mathbf{R}[\varepsilon]}, \qquad a \otimes 1 + b \otimes \varepsilon \mapsto b.$$

More suggestively, with redundant tensors and subscripts dropped, $f(a+\varepsilon) = f(a) + g(a)\varepsilon$; so that g is the derivative of f with respect to 'infinitesimal variation' of a, i.e. adding an element of square zero which commutes with a. (In the definition the ε is generic, but an obvious use of naturality would show that the formula holds even in the form:

$$f_A(a+\varepsilon) = f_A(a) + g_A(a)\varepsilon$$

for any $\varepsilon \in A$ such that $\varepsilon^2 = [\varepsilon, a] = 0$.)

Now the calculation with triangular matrices, as in the complex case, shows that $f_{\mathbf{R}}$ is differentiable, with derivative $g_{\mathbf{R}}$. One needs only to use in addition the naturality with respect to the homomorphism

$$\mathbf{R}[\varepsilon] \to T \quad \text{by} \quad \varepsilon \mapsto \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Thus the procedure iterates, and $f_{\mathbf{R}}$ is C^∞. The rest of the proof is exactly as in the complex case.

3. 'Semisimple part' generates the Whitney algebra of global functions

Before discussing further topological and analytic properties of functions, it is wise to abandon the rather cumbersome 'family of functions' terminology used in the earlier sections. A *global function, with domain* $\mathscr{A}$ is a natural transformation $U \to U$, where U is the 'underlying set' functor from a category $\mathscr{A}$ of finite dimensional (real or complex) algebras to sets. Thus a global function F is a family of functions $\langle F_A \rangle$, indexed by algebras $A \in \mathscr{A}$, natural with respect to homomorphisms $\varphi \in \mathscr{A}$. A global function F is said to be continuous, smooth, locally bounded, etc., just in case each F_A has the corresponding property, where the structure (e.g. topology) on A comes from the vector space structure on A. (The idea that a function should have a single set as its fixed domain is a relatively recent one, helpful in clarifying and codifying earlier work, but usefully abandoned on many occasions. For instance, the exponential function is often best viewed as a single function, which can be evaluated at numbers, matrices, etc.)

For now, the following full subcategories of the categories of finite dimensional **R**-algebras and **C**-algebras will suffice as domains. All algebras, commutative algebras, and the one-object category **C**, in the complex case; and triangulable algebras, commutative triangulable algebras, and **R** alone in the real case. (*Full* subcategory simply means that all the algebra homomorphisms between algebras in the subcategory are allowed.) It will be simplest to treat the real and complex cases simultaneously, by adopting the following terminology.

K is the scalar field (**R** or **C**). *Global function,* unadorned, means global function on all algebras if $K=\mathbf{C}$, on all triangulable algebras if $K=\mathbf{R}$. *Smooth* means analytic if $K=\mathbf{C}$, C^∞ if $K=\mathbf{R}$. (Commutative) *algebra* means (commutative) triangulable algebra if $K=\mathbf{R}$. Note that a global function with domain (the one-object category) K is simply an arbitrary function $K \to K$. (Since the only K-algebra homomorphism $K \to K$ is the identity, the naturality requirement is vacuous.) So a global function with domain K is called a *scalar function.*

The earlier results can now be restated: a global function is smooth if (and only if) it is continuous; and restriction to K gives a bijection from smooth global functions to smooth scalar functions.

The simplest discontinuous global function with smooth restriction to scalars is λ, the *semisimple part function.* Given any $a \in A$, we can write a uniquely as a sum $a=\lambda_A(a)+(a-\lambda_A(a))$, where $\lambda_A(a)$ is semisimple (has minimal polynomial with no repeated roots), $a-\lambda_A(a)$ is nilpotent, and $\lambda_A(a)$ commutes with a. λ is clearly natural with respect to algebra homomorphisms, since these preserve semisimplicity, nilpotence, and commutativity. So λ is a global function; and λ_K is the identity function.

While it is unnecessary for our purposes, it is worth observing that the central roles of the global functions identity and semisimple part can be accounted for on purely algebraic grounds: $\{X,\lambda\}$ is the center of the monoid (under composition) of global functions. Since it won't be used in what follows, the proof is omitted.

Proposition 3. *For each global function F, the following are equivalent:*

(i) $F \in \{X,\lambda\}$;

(ii) $F \circ G = G \circ F$ *for all global functions G (i.e. $F_A \circ G_A = G_A \circ F_A$ for all A, or all commutative A);*

(iii) $F \circ (\alpha X+\beta)=(\alpha X+\beta)\circ F$ *for all $\alpha,\beta \in K$, and* $F\circ(X^2)=(X^2)\circ F$;

(iv) F_A *is an algebra homomorphism for all commutative A.*

All global functions can be expressed in terms of λ. Indeed, by naturality, a global function F is determined by its (compatible) components F_A for *cyclic* algebras $A=K[X]/_{p(X)}$, where p is a monic polynomial with all roots in K. Hence we get the following description of the *Whitney K-algebra of global functions:*

$$W \cong \varprojlim_{p} K[X]/p(X) \cong \prod_{\mu \in K} \varprojlim_{m} K[X]/(X-\mu)^m \cong K^K[[X-\lambda]].$$

The first inverse limit is over all monic polynomials p with all roots in K, ordered by divisibility; the second over natural numbers m. In $K^K[[X-\lambda]]$, K^K is the K-algebra of all functions $K \to K$, and $\lambda \in K^K$ is the identity function on K. It will be simplest for our purposes to write power series with factorials in the denominators. Every global function F has thus a canonical representation:

$$F = \sum f_n(\lambda) \frac{(X-\lambda)^n}{n!}$$

(where the subscript A on λ has been dropped). The interpretation of $f_n(\lambda(a))$ is not problematic: since a semisimple element generates a subalgebra isomorphic to a product of copies of K, the value of a (global) function at semisimple elements is determined by its value at scalars. Note that for any fixed algebra A, the sum on the right is finite, since $(a-\lambda(a))^n = 0$ for $n \geq \dim A$.

It is frequently easiest to compute $F_A(a)$ by thinking of F as an element of $\varprojlim K[X]/p(X)$. Choose a polynomial $p(X)$, monic with all roots in K, such that $p(a)=0$. Then select $r(X) \in K[X]$ such that the derivatives $r^{(k)}(X)$ satisfy

$$r^{(k)}(\mu) = f_k(\mu) \quad \text{for } 0 \leq k < k_\mu,$$

where k_μ is the multiplicity of μ as a root of $p(X)$. Then $F_A(a) = r(a)$.

While we won't need it, it is worth remarking that the composition of global functions corresponds to an easily-described 'formal composition' of Whitney series. Identifying a global function with its series, we have

$$\left(\sum f_n(\lambda) \frac{(X-\lambda)^n}{n!} \right) \circ G = \sum (f_n \circ g_0)(\lambda) \frac{(G-g_0(\lambda))^n}{n!},$$

where $G - g_0(\lambda)$ is a power series with no constant term, so that the sums to be collected are finite.

The formal 'infinitesimal' differentiation introduced in Section 2 (for the real case, but one can replace **R** by **C** everywhere) corresponds exactly to differentiation with respect to X in the Whitney algebra:

$$\frac{\partial}{\partial X} \left(\sum f_n(\lambda) \frac{(X-\lambda)^n}{n!} \right) = \sum f_{n+1}(\lambda) \frac{(X-\lambda)^n}{n!}.$$

(By naturality, it suffices to verify the correspondence for the case $A = K[t]/t^n$, where it is an easy computation.) Hence the formal derivative of a global function F is denoted by $\partial F/\partial X$.

All this looks quite formal; the next two sections demonstrate its use in investigating smoothness of global functions. (The introduction by Whitney of the algebra $K^K[[X-\lambda]]$ was for a rather different but related purpose: to combine all the Taylor series of a (scalar) smooth real function into a single entity.)

4. Smooth equals commutative-differentiable plus $\partial/\partial\lambda = 0$

The conventions of Section 3 remains in force: *algebra* means finite dimensional (and triangulable, if $K = \mathbf{R}$). Restriction to the subcategory of commutative algebras gives an isomorphism from the (infinite dimensional) algebra of global functions to the algebra 'global functions with domain commutative algebras', and this isomorphism respects composition of functions. Thus one may safely identify a global function with its restriction to the category of commutative algebras. But there is a sharp distinction to be made in topological and analytic properties: for instance λ is discontinuous, but its restriction to commutative algebras is smooth (even linear).

That λ is discontinuous follows from the theorems in Sections 1 and 2: if continuous, it would be smooth, but it agrees with the identity function on scalars, which contradicts uniqueness of extrapolation. But it is easier to see it directly, on the algebra T of 2×2 triangular matrices:

$$\lambda_T\begin{bmatrix} t & 1 \\ 0 & 0 \end{bmatrix} = \begin{cases} \begin{bmatrix} t & 1 \\ 0 & 0 \end{bmatrix} & \text{for } t \neq 0, \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{for } t = 0. \end{cases}$$

Proposition 4. *The following are equivalent, for any global function F:*

(i) F *is* commutative-differentiable (*i.e. F_A is differentiable for each commutative algebra A*).

(ii) *For every A, and every $a \in A$, the function*

$$K \to A \qquad \mu \mapsto F_A(a+\mu)$$

is differentiable (*or differentiable at* 0).

(iii) *For every $m \geq 0$, the function*

$$K \to K[t]/t^m = A \qquad \mu \mapsto F_A(t+\mu)$$

is differentiable.

(iv) *In the Whitney series $\sum f_n(\lambda)(X-\lambda)^n/n!$ for F, all f_n are differentiable.*

Proof. (i)$\Rightarrow$(ii) Without loss of generality (by naturality with respect to the inclusion of the subalgebra generated by a) A is commutative; and then $K \to A$ is a composite of differentiable maps.

(ii)$\Rightarrow$(iii) is just restriction to a special case.

(iii)$\Rightarrow$(iv) $F_A(t+\mu) = \sum_{n=0}^{m-1} f_n(\mu)t^n/n!$, and $\langle t^n/n!\rangle$ is a linear basis for A.

(iv)$\Rightarrow$(i) F_A is a finite sum of differentiable maps, since λ_A is linear for commutative A, and multiplication is differentiable.

Note that the proof works as well if 'differentiable' is replaced by other regularity conditions, e.g. continuous, C^k, analytic, algebraic, polynomial (all as maps K or K^n to K or K^n). The main point: commutative-'regularity' is equivalent to regularity along lines $a+K$.

For commutative-differentiable global functions we have thus an additional derivation. On the Whitney series it is

$$\sum f_n(\lambda)\frac{(X-\lambda)^n}{n!} \mapsto \sum f'_n(\lambda)\frac{(X-\lambda)^n}{n!},$$

which could be described as 'partial derivative with respect to λ for fixed $(X-\lambda)$'. But the concrete description in terms of the global functions is more illuminating:

$$F'_A(a)=\lim \frac{F_A(a+\mu)-F_A(a)}{\mu}$$

where the limit is over $\mu\in K$, tending to 0. Thus $F\mapsto F'$ should better be described as the 'scalar derivative', as opposed to the 'infinitesimal derivative', which exists for all functions. (Proposition 4 asserts that the scalar derivative exists if and only if the Whitney series derivative $\sum f'_n(\lambda)(X-\lambda)^n/n!$ does; but the proof shows that they correspond.)

Now the description of (globally) smooth global functions in terms of the Whitney series is simple, since they are precisely the canonical extrapolations of smooth scalar functions. Thus a global function is smooth if and only if $f'_n=f_{n+1}$ for all n; that is, if and only if it is commutative-differentiable and the scalar and infinitesimal derivatives are equal! Finally, we may phrase the situation in one more way. The derivation 'partial derivative with respect to λ with X fixed' – as opposed to the partial with $X-\lambda$ fixed – will be denoted by $\partial/\partial\lambda$ on the formal series, or $\partial/\partial\lambda$ on the commutative-differentiable global functions. But on the formal series it is clear that $\partial/\partial\lambda$ is just the difference between the scalar and infinitesimal derivatives. Summing up:

Theorem 5. *For each global function F, the following are equivalent:*

(i) *F is smooth (i.e. F_A is smooth for all A).*

(ii) *F is commutative-differentiable and $\partial F/\partial\lambda=0$.*

(iii) *F is commutative-differentiable and its scalar and infinitesimal derivatives agree.*

(iv) *For each commutative A, and $a\in A$, $\varepsilon\in A$ with $\varepsilon^2=0$,*

$$F_A(a+\varepsilon)=F_A(a)+\left(\lim \frac{F_A(a+\mu)-F_A(a)}{\mu}\right)\varepsilon,$$

with the limit over scalars tending to 0.

Theorem 5 seems at first rather surprising. The existence of the infinitesimal derivative is no restriction at all on a global function, and the existence of the scalar derivative (commutative-differentiability) is a very weak condition, especially in the real case. Yet the equality of these two derivatives is a very strong condition. The situation is analogous to the relation for a scalar function $\mathbf{C} \to \mathbf{C}$, between real-differentiability and complex-differentiability. The Cauchy–Riemann equations can be (and often are) interpreted as expressing the fact that a function of x and y is really a function of $z = x + iy$, by introducing either of two derivations: The 'partial derivative with respect to $\bar{z}$, with z held fixed', or the 'partial derivative with respect to x, with z held fixed'. The vanishing of either of these yields from the weak condition of real differentiability the strong condition of complex differentiability, hence analyticity. Likewise the condition $\partial/\partial\lambda = 0$ can be thought of as asserting that F does not depend on the (nonsmooth) function λ, but only on the (smooth) function X, and is therefore smooth.

5. Bornological implies smooth

The previous section was devoted to the question: What conditions on the behavior of F_A on *commutative* algebras A are sufficient for (global) smoothness of F? We now return to the question of Sections 1 and 2: What conditions on the behavior of F_A on *all* algebras A suffice for smoothness? It turns out that continuity is far stronger than necessary: bornological implies smooth. That is, if each F_A carries bounded sets to bounded sets, all F_A are smooth (and then F is of course just the canonical extrapolation of F_K). Rather weaker conditions suffice: locally bounded at scalars, or not too badly unbounded locally along straight lines.

Theorem 6. *Suppose that a global function F satisfies one of the following properties:*

(i) *for each A, and each $\mu \in K$, there is a neighborhood V of μ in A such that $F_A(V)$ is bounded;*

(ii) *for some $n \geq 0$, and for each A and $a, b \in A$, the function*

$$K \to A \quad \text{by} \quad \mu \mapsto F_A(a + \mu b)$$

is $O(t^{-n})$.

Then F is smooth.

The implied constant in $O(t^{-n})$ may depend on A, a, b, and of course on the selection of norm for the vector space underlying A.

Proof. Either of the conditions (i), (ii) is stable under replacing F by:

$F +$ polynomial (global) function, or

$F \circ$ (translation by scalar), or

$\partial F/\partial X$, the formal (= infinitesimal) derivative of F.

Of these, the only one not quite straightforward is the stability of (i) under replacement of F by $\partial F/\partial X$. One must show that the composite (as defined in Section 2)

$$A \to A\otimes K[\varepsilon] \to A\otimes K[\varepsilon] \to A$$

is bounded in some neighborhood of $\mu\cdot 1\in A$, knowing that the middle map has this property. Unfortunately, the first map takes μ, not to μ, but to $\mu+1\otimes\varepsilon$. Fortunately, naturality with respect to the automorphism $1_A\otimes\varphi_t$, where $\varphi_t(\alpha+\beta\varepsilon)=\alpha+t\beta\varepsilon$, serves the desired purpose, since this automorphism will, for sufficiently small t, carry a neighborhood of $\mu+1\otimes\varepsilon$ inside a neighborhood of μ.

To complete the proof of smoothness under hypothesis (i), we may assume, after adjusting by a linear polynomial and translation, that the Whitney series for F satisfies $f_0(0)=f_1(0)=0$, and need only prove that $f'_0(0)$ exists and vanishes. (This shows that $f'_0=f_1$, and then $f'_n=f_{n+1}$ follows by stability under $\partial/\partial X$.) Now let $A=T_3(K)$, upper triangular 3×3 matrices, and

$$a=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix},\qquad b=\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix},$$

and evaluate $F_A(\alpha a+\beta b)$ with $\alpha,\beta\in K$, $\beta\neq 0$.

To do this, choose a polynomial $r(X)$ with $r(0)=r'(0)=0$ and $r(\beta)=f_0(\beta)$ ($r(X)=(f_0(\beta)/\beta^2)X^2$ will do nicely) and then

$$F_A(\alpha a+\beta b)=r(\alpha a+\beta b)=\frac{f_0(\beta)}{\beta^2}\begin{bmatrix}0&0&\alpha^2\\0&0&\alpha\beta\\0&0&\beta^2\end{bmatrix}.$$

One wants

$$\lim_{\beta\to 0}\frac{f_0(\beta)}{\beta}=0.$$

The upper right corner gives it, since otherwise $(f_0(\beta)/\beta)(\alpha^2/\beta)$ would be unbounded as $\alpha\to 0$, $\beta\to 0$ with $\beta\ll\alpha$, say $\beta=\alpha^3$; and that would contradict the boundedness of F_A near 0.

The proof under the hypothesis (ii) is nearly the same. Let A be the algebra $T_N(K)$, with $N=n+3$, and let a be the matrix with ones on the superdiagonal, zeroes elsewhere, and b the matrix with 1 in the lower right corner, zeroes elsewhere. After translation and adjusting F by a polynomial, we may assume $f_0(0)=f_1(0)=\cdots=f_{N-2}(0)=0$, and need only prove $f'_0(0)=0$. Evaluating:

$$F(a+\mu b)=\frac{f_0(\mu)}{\mu^{N-1}}(a+\mu b)^{N-1},$$

and the upper right corner of this matrix is $f_0(\mu)/\mu^{N-1}$. This being $O(\mu^{-n})$ says $f_0(\mu)=O(\mu^2)$, which gives the result.

Afterword

The importance of nilpotent elements in the practice of analysis is by no means new; see for instance [1]. Recent clarification of their role received a strong impetus from Lawvere's Chicago lectures on categorical dynamics, and was developed further by, among others, Dubuc, Joyal, Kock, Lawvere [2, 4, 5], Reyes, and Wraith. The emphasis in these developments has been on commutative algebra, in which the (finite dimensional) commutative triangulable **R**-algebras (finite products of Weil algebras) have played a strong part. What is suggested is that noncommutative algebra may also be useful in these investigations. One should explore functions of several noncommuting variables, and the underlying set functor as a ring object in the category of set-valued functors on (finite dimensional, and triangulable in the real case) algebras.

Acknowledgments

Thanks: to my class in linear algebra, whose efforts to understand the use of exponentials of matrices in solving differential equations encouraged me to pursue the matter a bit further; to colleagues at the University of Puerto Rico conference on topos theory, who expressed warm interest in the results; to W.R. Zame, whose knowledge of commutative Banach algebras led to showing that similar results were also possible for these [7]; but above all to F. William and Fatima Lawvere for persistent advice, encouragement, and generous assistance.

References

[1] P. Deligne, Equations Différentielles à Points Singuliers Réguliers, Lecture Notes in Math., No. 163 (Springer, Berlin, 1970).

[2] E.J. Dubuc, Sur les modeles de la geométrie différentielle synthétique [especially section 0.9], Cahiers de Top. et Geom. Diff. XX(3) (1979) 234-279.

[3] G. Glaeser, Functions composées différentiables, Ann. of Math. Ser. II 77 (1963) 193–209.

[4] A. Kock, A simple axiomatics for differentiation, Math. Scand. 40 (1977) 183–193.

[5] F.W. Lawvere, Categorical Dynamics, Lecture at Univ. of Chicago 1967, in: Various Publications Series No. 30 (Aarhus University, 1979).

[6] F.W. Lawvere, Some algebraic problems in the context of functorial semantics of algebraic theories, in: Lecture Notes in Math., No. 61 (Springer, Berlin, 1968) 41-61.

[7] S.H. Schanuel and W.R. Zame, Naturality of the functional calculus, to appear in J. Lond. Math. Soc.