ON THE CAUCHY-KOWALEVSKI THEOREM

In memory of Adrien Douady

by Marc CHAPERON

ABSTRACT. After a short review of the basic properties of analytic functions, we apply the infinite dimensional theory to get a simple proof of the Cauchy-Kowalevski theorem, in an infinite dimensional version which seems to be new.

INTRODUCTION

Most mathematicians no longer teach differential calculus in Banach spaces, though the theory has proved increasingly useful since Henri Cartan's first lectures on the subject [1]. Paradoxically, Cartan himself never included in his lectures the basic properties of analytic functions in Banach spaces, which his student Adrien Douady had written down in a simple and aesthetic way [3].

After a short introduction to this elegant theory, we recall (in the analytic case) Joel Robbin's proof of the existence theorem for differential equations via the implicit function theorem in Banach spaces [5] and explain how to adapt it to get the Cauchy-Kowalevski theorem. The painless proof casts some light on the reason why this result is true only in the analytic category.

Throughout the paper, E, F denote two Banach spaces over $\mathbf{K} = \mathbf{R}$ or $\mathbf{C}$.

1. ANALYTIC FUNCTIONS IN ARBITRARY DIMENSION

POWER SERIES We let $L^0(E,F) = L^0_s(E,F) = F$ and, for each positive integer n, we endow the space $L^n(E,F)$ of continuous n–linear maps $a_n : E^n \to F$ with its standard norm $|a_n| := \sup_{|x_1|=\cdots=|x_n|=1} |a_n(x_1,\ldots,x_n)|$, which makes it into a Banach space. Denoting by $L^n_s(E,F)$ the closed subspace

consisting of symmetric n–linear maps, we associate to each $a_n \in L^n_s(E,F)$ the homogeneous polynomials $E \ni x \mapsto a_n x^\ell \in L^{n-\ell}_s(E,F)$, $0 \le \ell \le n$, defined as follows: for $x, x_{\ell+1}, \dots, x_n \in E$, we have that $a_n x^\ell(x_{\ell+1}, \dots, x_n)$ is the value of $a_n(x_1, \dots, x_n)$ when $x_j = x$ for $1 \le j \le \ell$ (hence $a_n x^0 = a_n$). As when $E = \mathbf{K}$, the j–th derivative of $a_n x^n$ is $\frac{n!}{(n-j)!} a_n x^{n-j}$ for $j \le n$ and 0 otherwise. In particular, its n–th derivative is the constant $n! a_n$ and, if $\mathbf{K} = \mathbf{C}$, the homogeneous polynomial $x \mapsto a_n x^n$ is *holomorphic*, meaning that it is differentiable and that its derivative at each point is $\mathbf{C}$–linear.

A *power series* on E with values in F is a series of functions u_n of E into F whose general term is a homogeneous polynomial $u_n(x) = a_n\, x^n$, $a_n \in L^n_s(E,F)$. We shall call it the power series $\sum_{n\in\mathbf{N}} a_n\, x^n$ or $\sum a_n\, x^n$.

The *strict convergence radius* $\rho \in [0, +\infty]$ of the power series $\sum a_n\, x^n$ is the supremum of those $r \ge 0$ satisfying $\sum |a_n|\, r^n < \infty$. It is given by $\rho^{-1} = \limsup |a_n|^{\frac{1}{n}}$. When ρ is positive, the power series is called *convergent*; it converges at every point of the open ball $B_\rho(0)$, (normally) uniformly in $\bar{B}_r(0)$ for $0 \le r < \rho$. Hence, its sum $f : B_\rho(0) \to F$ is continuous and more:

PROPOSITION 1.1. *Every convergent power series $\sum a_n\, x^n$ on E with values in F, having strict convergence radius ρ, converges in the C^∞ sense in $B_\rho(0)$. Its sum $f : B_\rho(0) \to F$ is C^∞ and, if $\mathbf{K} = \mathbf{C}$, it is holomorphic. More precisely, for $j \in \mathbf{N}$, the power series $\sum \frac{(n+j)!}{n!}\, a_{n+j}\, x^n$ obtained by differentiating j times $\sum a_n\, x^n$ has the same strict convergence radius ρ as $\sum a_n\, x^n$, and its sum is $D^j f : B_\rho(0) \to L^j_s(E,F)$. Hence, $D^j f(0) = j! a_j$, showing that, in $B_\rho(0)$, the function f is the sum of its Taylor expansion at 0.*

Proof. For $a_n \in L^n_s(E,F)$ and $0 \le \ell \le n$, the symmetric form defining $x \mapsto a_n x^\ell$ has the same norm as a_n since $a_n \mapsto \big((x_1, \dots, x_\ell) \mapsto a_n x_1 \cdots x_\ell\big)$ is an isometric linear map of $L^n(E,F)$ onto $L^\ell\big(E, L^{n-\ell}(E,F)\big)$. $\square$

REMARK. For $\dim E > 1$, the ball of strict convergence $B_\rho(0)$, which depends on the norm and not just the topology of E, is definitely not the largest open subset in which the power series converges, as shown by the power series $\sum b_n x^n y^n$ on $\mathbf{K}^2$ when $\sum b_n z^n$ is convergent in one variable. However, convergence depends only on the topologies of E and F.

ANALYTIC MAPS A map f of an open subset U of E into F is called *analytic* when, for all $x_0 \in U$, there exists a convergent power series $\sum a_n\, x^n$ such that $f(x) = \sum a_n\, (x - x_0)^n$ in a neighbourhood of x_0. By Proposition 1.1,

this implies that f is C^∞ (and, if $\mathbf{K} = \mathbf{C}$, holomorphic) and can be expressed in a neighbourhood of every point $x_0 \in U$ as the sum of its Taylor expansion at x_0. As in one variable, the following fundamental result can be deduced from Cauchy's formula (see for example [2], chapitre 5, théorème principal):

PROPOSITION 1.2. *If $\mathbf{K} = \mathbf{C}$, a function f of an open subset U of E into F is analytic if and only if it is holomorphic. When a sequence (g_n) of holomorphic functions of U into F converges locally uniformly to a function g, the latter is holomorphic.*

COROLLARY 1.3 (analyticity of inverse maps). *If an analytic local map $f : (E, x_0) \to F$ has invertible derivative at x_0, then its local inverse $\big(F, f(x_0)\big) \to (E, x_0)$ is analytic. Hence, given a third Banach space Λ over $\mathbf{K}$ and an analytic local map $g : \big(\Lambda \times E, (\lambda_0, x_0)\big) \to (F, 0)$, if the partial $\partial_2 g(\lambda_0, x_0) : E \to F$ is invertible, the implicit function $\varphi : (\Lambda, \lambda_0) \to (E, x_0)$ whose graph coincides with $g^{-1}(0)$ near (λ_0, x_0) is analytic.*

Proof. As the inverse of a $\mathbf{C}$–linear isomorphism is $\mathbf{C}$–linear, this is obvious if $\mathbf{K} = \mathbf{C}$. The real case follows by complexification [2], as well as

COROLLARY 1.4. *The sum of a convergent power series is analytic in its ball of strict convergence. The composed map of two analytic maps is analytic.*

2. ANALYTIC LOCAL CAUCHY PROBLEMS

2.1 CAUCHY'S THEOREM ON ORDINARY DIFFERENTIAL EQUATIONS

Let f be an analytic map on an open subset $\mathrm{dom} f$ of $\mathbf{K} \times F$, taking its values in F. Given $(t_0, u_0) \in \mathrm{dom} f$, we are interested in the *local analytic solutions* of the *Cauchy problem*

$$\left\{ \begin{array}{l} \dfrac{du}{dt} = f(t, u) \\ u(t_0) = u_0, \end{array} \right. \tag{1}$$

i.e. analytic germs $\varphi : (\mathbf{K}, t_0) \to F$ which satisfy the *initial condition* $\varphi(t_0) = u_0$ and are solutions of the *differential equation* $\dfrac{du}{dt} = f(t, u)$, meaning that $\varphi'(t) = f(t, \varphi(t))$.

THEOREM (CAUCHY). *Under these hypotheses, the Cauchy problem* (1) *has a unique local analytic solution.*

Proof. If $\mathbf{K} = \mathbf{R}$, the complexified map of a solution of (1) is a solution of the complexified Cauchy problem. Now, if the theorem is true in the complex case, the solution φ of the complexified Cauchy problem must be the complexified map of a solution of (1) (hence existence and uniqueness in the real case), since otherwise $t \mapsto \overline{\varphi(\bar{t})}$ would be another solution, contradicting uniqueness. This reduces the question to the complex case.

If $\mathbf{K} = \mathbf{C}$, (1) is equivalent to the equation

$$v(t) = f\big(t, u_0 + \textstyle\int_{t_0}^t v(\tau)\, d\tau\big) \tag{2}$$

in the unknown local holomorphic function $v = \dfrac{du}{dt} : (\mathbf{C}, t_0) \to F$, where $\int_{t_0}^t v(\tau)\, d\tau$ denotes the local primitive of v which vanishes at t_0.

Robbin's idea is to use a small parameter $\varepsilon \in \mathbf{C}$. For $\varepsilon \neq 0$, setting

$$t = t_0 + \varepsilon T\,,\ V(T) = v(t)\,,\ d^{-1}V(T) := \textstyle\int_0^T V(\tau)\, d\tau, \tag{3}$$

one has $\int_{t_0}^t v(\tau)\, d\tau = \varepsilon d^{-1}V(T)$; hence, (2) writes

$$V(T) = f\left(t_0 + \varepsilon T, u_0 + \varepsilon d^{-1}V(T)\right), \tag{4}$$

an equation which still means something for $\varepsilon = 0$.

Let $\mathcal{H}_b(\mathbf{D}, F)$ denote the Banach space (Proposition 1.2) of all bounded holomorphic maps of the open unit disk $\mathbf{D} \subset \mathbf{C}$ into F, equipped with the norm of uniform convergence $|\cdot|_\infty$. For nonzero $\varepsilon \in \mathbf{C}$, the "microscope" (3) defines a bijection of the set of solutions $V \in \mathcal{H}_b(\mathbf{D}, F)$ of (4) onto the set of those solutions v of (2) which are defined and bounded in the open disk of radius $|\varepsilon|$ centred at t_0. Therefore, to establish Cauchy's theorem, we should just prove that there exist $\eta > 0$ and $r > 0$ such that, for $|\varepsilon| < \eta$, the equation (4) has a unique solution $V \in \mathcal{H}_b(\mathbf{D}, F)$ satisfying $|V - V_0|_\infty < r$, where $V_0 \in \mathcal{H}_b(\mathbf{D}, F)$ denotes the constant $f(t_0, u_0)$.

LEMMA 2.1. *The formula* $\Phi(\varepsilon, V)(T) := f\left(t_0 + \varepsilon T, u_0 + \varepsilon d^{-1}V(T)\right)$ *defines a local holomorphic map* $\Phi : \big(\mathbf{C} \times \mathcal{H}_b(\mathbf{D}, F), (0, V_0)\big) \to \big(\mathcal{H}_b(\mathbf{D}, F), V_0\big)$ *such that* $\Phi(0, V) = V_0$ *and therefore* $\partial_V \Phi(0, V_0) = 0$.

This yields Cauchy's theorem: by the implicit function theorem, there exist $\eta > 0$ and $r > 0$ such that, for $|\varepsilon| < \eta$, the equation $V - \Phi(\varepsilon, V) = 0$, i.e. (4), has a unique solution $V = \varphi(\varepsilon) \in \mathcal{H}_b(\mathbf{D}, F)$ satisfying $|V - V_0|_\infty < r$.

Proof of Lemma 2.1. Φ is obtained by composing two holomorphic maps:

- the polynomial map $\Pi : \mathbf{C} \times \mathcal{H}_b(\mathbf{D}, F) \to \mathcal{H}_b(\mathbf{D}, \mathbf{C} \times F)$ defined by $\Pi(\varepsilon, V)(T) = \big(t_0 + \varepsilon T, u_0 + \varepsilon d^{-1} V(T)\big)$

(indeed, Π is a holomorphic polynomial with values in $\mathcal{H}_b(\mathbf{D}, \mathbf{C} \times F)$ because d^{-1} is a continuous endomorphism of $\mathcal{H}_b(\mathbf{D}, F)$ since we have $|d^{-1}V(T)| \leq |T|\,|V|_\infty \leq |V|_\infty$, $V \in \mathcal{H}_b(\mathbf{D}, F), T \in \mathbf{D}$)

- the local map $f_* : \big(\mathcal{H}_b(\mathbf{D}, \mathbf{C} \times F), (t_0, u_0)\big) \to \big(\mathcal{H}_b(\mathbf{D}, F), V_0\big)$, defined by $f_* W = f \circ W$.

Indeed, if f is well-defined and bounded on the open ball $B_\rho(t_0, u_0)$ of radius $\rho > 0$ centred at (t_0, u_0), then, for each $W \in \mathcal{H}_b(\mathbf{D}, \mathbf{C} \times F)$ with $|W - (t_0, u_0)|_\infty < \rho$, the map $f \circ W$ is well-defined, holomorphic and bounded on $\mathbf{D}$. To see that f_* is holomorphic near the constant (t_0, x_0), notice that if $D^2 f$ satisfies $|D^2 f(t, x)| \leq c < \infty$ on $B_\rho(t_0, u_0)$, then, for all $W, \delta W \in \mathcal{H}_b(\mathbf{D}, \mathbf{C} \times F)$ with $|W - (t_0, u_0)|_\infty < \rho$ and $|W + \delta W - (t_0, u_0)|_\infty < \rho$, Taylor's formula yields

$$\begin{aligned} &\big|f\big(W(T) + \delta W(T)\big) - f\big(W(T)\big) - Df\big(W(T)\big)\delta W(T)\big| = \\ &\quad = \Big|\int_0^1 (1-s)\, D^2 f\big(W(T) + s\delta W(T)\big)\delta W(T)^2\, ds\Big| \leq \frac{c}{2}\,|\delta W|_\infty^2\,,\ T \in \mathbf{D}\,, \end{aligned}$$

implying that f_* is differentiable at W and that $Df_*(W)$ is the complex endomorphism of $\mathcal{H}_b(\mathbf{D}, \mathbf{C} \times F)$ given by $\big(Df_*(W)\delta W\big)(T) = Df\big(W(T)\big)\delta W(T)$. $\square$

2.2 The Cauchy-Kowalevski theorem

Hypotheses and notation Let $J^1(E, F) := E \times F \times L(E, F)$ and let f be an analytic map on an open subset $\mathrm{dom} f$ of $\mathbf{K} \times J^1(E, F)$, taking its values in F. Given $t_0 \in \mathbf{K}$ and an analytic germ $u_0 : (E, x_0) \to F$ satisfying $\big(t_0, j^1 u_0(x_0)\big) \in \mathrm{dom} f$, where $j^1 u_0(x_0) := \big(x_0, u_0(x_0), Du_0(x_0)\big)$, we are interested in the *local analytic solutions* of the *Cauchy problem*

$$\begin{cases} \partial_t u = f(t, x, u, \partial_x u) \\ u(t_0, x) = u_0(x), \end{cases} \tag{5}$$

i.e. analytic germs $\varphi : \big(\mathbf{K} \times E, (0, x_0)\big) \to F$ such that, setting $\varphi_t(x) := \varphi(t, x)$,

- φ is a solution of the *partial differential equation* $\partial_t u = f(t, x, u, \partial_x u)$, meaning that $\partial_t \varphi(t, x) = f\left(t, j^1 \varphi_t(x)\right)$
- the *initial condition* $\varphi_{t_0} = u_0$ is satisfied.

Theorem (Cauchy-Kowalevski). *Under these hypotheses, the Cauchy problem* (5) *has a unique local analytic solution.*

REDUCTION OF THE PROBLEM Denoting by g the analytic function defined on an open subset $\operatorname{dom} g \ni 0$ of $\mathbf{K} \times J^1(E,F)$ by

$$g(t,x,y,z) := f\left(t_0+t, x_0+x, u_0(x_0+x)+y, Du_0(x_0+x)+z\right) ,$$

(5) is equivalent to the local Cauchy problem

$$\left\{ \begin{array}{l} \partial_t w = g(t,x,w,\partial_x w) \\ w(0,x) = 0 \end{array} \right. \tag{6}$$

in the unknown function $w(t,x) := u(t_0+t, x_0+x) - u_0(x_0+x)$ near $0 \in \mathbf{K} \times E$. Replacing $w(t,x)$ by $w(t,x) - tg(0)$ and $g(t,x,y,z)$ by $g(t,x,y+tg(0),z) - g(0)$, we may assume

$$g(0) = 0. \tag{7}$$

Solving (6) in an open convex subset $C \ni 0$ of $\mathbf{K} \times E$ is equivalent to finding an analytic map $v : C \to F$ (the partial derivative $v := \partial_t w$) such that

$$v(t,x) = g\left(t,x,\partial_t^{-1} v(t,x), \partial_x \partial_t^{-1} v(t,x)\right) , \tag{8}$$

where, as in [4],

$$\partial_t^{-1} v(t,x) = \textstyle\int_0^t v(\tau,x)\, d\tau .$$

Again, we introduce a small parameter $\varepsilon \in \mathbf{K}$, whose use is subtler than before, due to derivation with respect to x: for $\varepsilon \neq 0$, setting

$$(t,x) = (\varepsilon^2 T, \varepsilon X) , \ V(T,X) = v(t,x) ,$$

one has $\partial_t^{-1} v(t,x) = \varepsilon^2 \partial_T^{-1} V(T,X)$ and $\partial_x \partial_t^{-1} v(t,x) = \varepsilon \partial_X \partial_T^{-1} V(T,X)$, hence (8) writes

$$V(T,X) = g\left(\varepsilon^2 T, \varepsilon X, \varepsilon^2 \partial_T^{-1} V(T,X), \varepsilon \partial_X \partial_T^{-1} V(T,X)\right) . \tag{9}$$

THE FUNCTION SPACE The presence of $\partial_X \partial_T^{-1} V(T,X)$ in (9) makes it harder to find a function space for which the analogue of Lemma 2.1 holds.

PROPOSITION 2.2. *Given Banach spaces E_1, F_1, let B be the open unit ball of E_1. The set $\mathcal{F}(E_1,F_1)$ of all $V : B \to F_1$ of the form $V(x) = \sum_{n\in\mathbf{N}} V_n x^n$ with $V_n \in L_s^n(E_1,F_1)$ and $\sum_{n\in\mathbf{N}} |V_n| < \infty$ is a Banach space over $\mathbf{K}$ for the norm $|V|_1 := \sum_{n\in\mathbf{N}} |V_n|$ and its elements are analytic functions.*

Proof. The power series whose sum belongs to $\mathcal{F}(E_1,F_1)$ have strict convergence radius ≥ 1. Therefore, by Corollary 1.4, the elements of $\mathcal{F}(E_1,F_1)$ are analytic. As each $V \in \mathcal{F}(E_1,F_1)$ identifies to the sequence consisting of the coefficients $V_n = \frac{1}{n!} D^n V(0)$ of the power series defining it, a standard argument proves that $\mathcal{F}(E_1,F_1)$ is a Banach space. $\square$

NOTATION. Let $\mathcal{F}_0(E_1,F_1)$ be the closed subspace of $\mathcal{F}(E_1,F_1)$ consisting of all V with $V(0)=0$. The following result, implicit in [4] (p. 44, estimate line -4), has no C^∞ analogue. This may be viewed as "the" reason why the Cauchy-Kowalevski theorem is true only in the analytic category:

PROPOSITION 2.3. *Let E_1 denote the Banach space $\mathbf{K}\times E$* endowed with the norm $|(T,X)| := |T|+|X|$. *For $V\in\mathcal{F}_0(E_1,F)$, one has $\partial_T^{-1}V\in\mathcal{F}_0(E_1,F)$, $\partial_X\partial_T^{-1}V\in\mathcal{F}_0\big(E_1,L(E,F)\big)$, $|\partial_T^{-1}V|_1\le\frac{1}{2}|V|_1$ and $|\partial_X\partial_T^{-1}V|_1\le|V|_1$.*

Proposition 2.3, whose easy proof is given in Appendix A, provides the analogue of the first point in the proof of Lemma 2.1. We turn to the second point, again inspired by [4] (Proposition 1.11 p. 42) and proved in Appendix B:

PROPOSITION 2.4. *Given three Banach spaces E_1,F_1,F_2 and an analytic local map $g:(F_1,0)\to(F_2,0)$, the formula $g_*(W):=g\circ W$ defines a local analytic map $g_*:\big(\mathcal{F}_0(E_1,F_1),0)\to\big(\mathcal{F}_0(E_1,F_2),0\big)$.*

PROOF OF THE THEOREM Here is the analogue of Lemma 2.1:

LEMMA 2.5. *With the notation of Proposition 2.3, the formula*

$$\Phi(\varepsilon,V)(T,X) := g\left(\varepsilon^2T,\varepsilon X,\varepsilon^2\partial_T^{-1}V(T,X),\varepsilon\partial_X\partial_T^{-1}V(T,X)\right)$$

defines an analytic map $\Phi:\big(\mathbf{K}\times\mathcal{F}_0(E_1,F),\{0\}\times\mathcal{F}_0(E_1,F)\big)\to\big(\mathcal{F}_0(E_1,F),0\big)$.

Proof. Proposition 2.3 asserts that we have $\partial_T^{-1}\in L\big(\mathcal{F}_0(E_1,F),\mathcal{F}_0(E_1,F)\big)$ and $\partial_X\partial_T^{-1}\in L\left((\mathcal{F}_0(E_1,F),\mathcal{F}_0\big(E_1,L(E,F)\big)\right)$, implying that the formula $\Pi(\varepsilon,V)(T,X) := \big(\varepsilon^2T,\varepsilon X,\varepsilon^2\partial_T^{-1}V(T,X),\varepsilon\partial_X\partial_T^{-1}V(T,X)\big)$ defines an analytic map $\Pi:\big(\mathbf{K}\times\mathcal{F}_0(E_1,F),\{0\}\times\mathcal{F}_0(E_1,F)\big)\to\big(\mathcal{F}_0\big(E_1,\mathbf{K}\times J^1(E,F)\big),0\big)$. Since $\Phi=g_*\circ\Pi$, we conclude by Proposition 2.4. $\square$

As $\Phi(0,V)=0$ yields $\Phi(0,0)=0$ and $\partial_V\Phi(0,0)=0$, the implicit function theorem implies that there exist $\eta>0$ and $r>0$ such that, for $|\varepsilon|<\eta$, the equation $V-\Phi(\varepsilon,V)=0$, i.e. (9), has a unique solution $V=\varphi(\varepsilon)\in\mathcal{F}_0(E_1,F)$ satisfying $|V|_1<r$. Now, if v is a solution of (8) which is analytic near 0, then, for $|\varepsilon|\in(0,\eta)$ small enough, the solution $V(T,X)=v(\varepsilon^2T,\varepsilon X)$ of (9) is well-defined in B, belongs to $\mathcal{F}_0(E_1,F)$ and satisfies $|V|_1<r$, proving the local existence and uniqueness of the solution of (8), hence of (5). $\square$

REMARKS. With the notation of Lemma 2.5, for $k \in \mathbf{N}^*$ and $V \in \mathcal{F}$, the k–th order Taylor polynomial $j_0^k\Phi(\varepsilon, V)$ of $\Phi(\varepsilon, V)$ at 0 depends only on ε and $j_0^k V$. Denoting it by $\hat{\Phi}^k(\varepsilon, j_0^k V)$ and replacing Φ by $\hat{\Phi}^k$ in what we have just done, we get that the solution of (5) is *formally* unique.

One can avoid Proposition 2.4 and just show that the map $V \mapsto \Phi(\varepsilon, V)$ is a contraction of the closed unit ball of $\mathcal{F}_0(E_1, F)$ for small enough ε. However, this requires the same ingredients as the proof of Proposition 2.4.

Appendices

A. PROOF OF PROPOSITION 2.3

For $V \in \mathcal{F}_0(E_1, F)$, we have that $V(T, X) = \sum_{n\in\mathbf{N}^*} V_n(T, X)^n$ with $V_n := \frac{1}{n!}D^n V(0,0) \in L_s^n(E_1, F)$, i.e. $V(T, X) = \sum_{n\in\mathbf{N}^*}\sum_{k=0}^{n}\begin{pmatrix} n \\ k \end{pmatrix} T^k\, V_n(1,0)^k(0,X)^{n-k}$ since $(T, X) = T(1,0) + (0, X)$; hence, near 0,

$$\begin{aligned}
\partial_T^{-1}V(T,X) &= \sum_{n\in\mathbf{N}^*}\sum_{k=0}^{n}\begin{pmatrix} n \\ k \end{pmatrix}\frac{T^{k+1}}{k+1}\, V_n(1,0)^k(0,X)^{n-k} \\
&= \sum_{n\in\mathbf{N}^*}\frac{1}{n+1}\sum_{k=0}^{n}\begin{pmatrix} n+1 \\ k+1 \end{pmatrix} T^{k+1}\, V_n(1,0)^k(0,X)^{n-k} \qquad (10)\\
&= \sum_{n\geq 2}\frac{1}{n}\sum_{k=1}^{n}\begin{pmatrix} n \\ k \end{pmatrix} T^k\, V_{n-1}(1,0)^{k-1}(0,X)^{n-k}
\end{aligned}$$

and, denoting by $(0, dX)$ the injection $E \ni \delta X \mapsto (0, \delta X) \in E_1$,

$$\begin{aligned}
\partial_X\partial_T^{-1}V(T,X) &= \sum_{n\geq 2}\frac{1}{n}\sum_{k=1}^{n-1}\begin{pmatrix} n \\ k \end{pmatrix}(n-k)\, T^k\, V_{n-1}(1,0)^{k-1}(0,X)^{n-k-1}(0,dX) \\
&= \sum_{n\geq 2}\sum_{k=1}^{n-1}\begin{pmatrix} n-1 \\ k \end{pmatrix} T^k\, V_{n-1}(1,0)^{k-1}(0,X)^{n-k-1}(0,dX) \qquad (11)\\
&= \sum_{n\in\mathbf{N}^*}\sum_{k=1}^{n}\begin{pmatrix} n \\ k \end{pmatrix} T^k\, V_n(1,0)^{k-1}(0,X)^{n-k}(0,dX)\,.
\end{aligned}$$

By (10), one has $\partial_T^{-1}V(T,X) = \sum_{n\geq 2}(\partial_T^{-1}V)_n(T,X)^n$ near 0, where

$$(\partial_T^{-1}V)_n(T_1,X_1)\cdots(T_n,X_n)=$$
$$\frac{1}{n}\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{k=1}^{n}\left(\begin{array}{c}n\\k\end{array}\right)T_{\sigma(1)}\cdots T_{\sigma(k)}\,V_{n-1}(1,0)^{k-1}(0,X_{\sigma(k+1)})\cdots(0,X_{\sigma(n)}),$$

and therefore

$$\begin{aligned}
&|(\partial_T^{-1}V)_n(T_1,X_1)\cdots(T_n,X_n)|\leq\\
\leq\ &\frac{1}{n}|V_{n-1}|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{k=1}^{n}\left(\begin{array}{c}n\\k\end{array}\right)|T_{\sigma(1)}|\cdots|T_{\sigma(k)}||X_{\sigma(k+1)}|\cdots|X_{\sigma(n)}|\\
\leq\ &\frac{1}{n}|V_{n-1}|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)|T_{\sigma(1)}|\cdots|T_{\sigma(k)}||X_{\sigma(k+1)}|\cdots|X_{\sigma(n)}|=\\
\leq\ &\frac{1}{n}|V_{n-1}|(|T_1|+|X_1|)\cdots(|T_n|+|X_n|)\,;
\end{aligned}$$

hence, $|(\partial_T^{-1}V)_n|\leq\dfrac{1}{n}|V_{n-1}|$ and $\sum_{n\geq 2}|(\partial_T^{-1}V)_n|\leq\sum_{n\in\mathbf{N}^*}\frac{1}{n+1}|V_n|\leq\frac{1}{2}|V|_1$, which does yield $\partial_T^{-1}V\in\mathcal{F}_0(E_1,F)$ and $|\partial_T^{-1}V|_1\leq\frac{1}{2}|V|_1$.

By (11), one has $\partial_X\partial_T^{-1}V(T,X)=\sum_{n\in\mathbf{N}^*}(\partial_X\partial_T^{-1}V)_n(T,X)^n$ near 0, where

$$(\partial_X\partial_T^{-1}V)_n(T_1,X_1)\cdots(T_n,X_n)=$$
$$\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{k=1}^{n}\left(\begin{array}{c}n\\k\end{array}\right)T_{\sigma(1)}\cdots T_{\sigma(k)}\,V_n(1,0)^{k-1}(0,X_{\sigma(k+1)})\cdots(0,X_{\sigma(n)})(0,dX),$$

and therefore

$$\begin{aligned}
&|(\partial_X\partial_T^{-1}V)_n(T_1,X_1)\cdots(T_n,X_n)|\leq\\
\leq\ &|V_n|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{k=1}^{n}\left(\begin{array}{c}n\\k\end{array}\right)|T_{\sigma(1)}|\cdots|T_{\sigma(k)}||X_{\sigma(k+1)}|\cdots|X_{\sigma(n)}|\\
\leq\ &|V_n|\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{k=0}^{n}\left(\begin{array}{c}n\\k\end{array}\right)|T_{\sigma(1)}|\cdots|T_{\sigma(k)}||X_{\sigma(k+1)}|\cdots|X_{\sigma(n)}|=\\
\leq\ &|V_n|(|T_1|+|X_1|)\cdots(|T_n|+|X_n|)\,;
\end{aligned}$$

hence, $|(\partial_X\partial_T^{-1}V)_n|\leq|V_n|$ and $\displaystyle\sum_{n\in\mathbf{N}^*}|(\partial_X\partial_T^{-1}V)_n|\leq\sum_{n\in\mathbf{N}^*}|V_n|=|V|_1$, which does yield $\partial_X\partial_T^{-1}V\in\mathcal{F}_0\big(E_1,L(E,F)\big)$ and $|\partial_X\partial_T^{-1}V|_1\leq|V|_1$. $\square$

B. Proof of Proposition 2.4

Let $W(Z)=\sum_{n\in\mathbf{N}^*}W_nZ^n$ and $g(w)=\sum_{n\in\mathbf{N}^*}g_nw^n$ be the Taylor expansions of W and g at 0. Since $g\circ W(Z)=\sum_{n\in\mathbf{N}^*}g_n\,W(Z)^n$ near

0, the Taylor expansion $\sum_{n\in\mathbf{N}^*} g_{*n}W^n$ of g_* at 0 must be given by $g_{*n}W^n(Z) = g_n\, W(Z)^n$, i.e.

$$g_{*n}(W_1,\dots,W_n)(Z) = g_n\,(W_1(Z),\dots,W_n(Z))\,. \tag{12}$$

There remains to show that (12) defines a map $g_{*n} \in L^n_s\big(\mathcal{F}_0(E_1,F_1),\mathcal{F}_0(E_1,F_2)\big)$ for all $n\in\mathbf{N}^*$ and that the power series $\sum_{n\in\mathbf{N}} g_{*n}W^n$ converges (what follows is essentially Proposition 1.11 page 42 of [4] and should be classical).

Clearly, g_{*n} is n–linear and symmetric. To see that it sends $\mathcal{F}_0(E_1,F_1)^n$ into $\mathcal{F}_0(E_1,F_2)$, we inject into (12) the Taylor expansion $W_j(Z) = \sum_{\ell\in\mathbf{N}^*} W_{j,\ell}Z^\ell$ of W_j at 0 for $1\le j\le n$: denoting by $\sum_{k\in\mathbf{N}^*} g_{*n}(W_1,\dots,W_n)_kZ^k$ the Taylor expansion of $g_{*n}(W_1,\dots,W_n)$, we get

$$g_{*n}(W_1,\dots,W_n)_kZ^k = \sum_{\ell\in\mathbf{N}^{*n},\,|\ell|=k} g_n\left(W_{1,\ell_1}Z^{\ell_1},\dots,W_{n,\ell_n}Z^{\ell_n}\right)$$

and therefore

$$\begin{gathered} g_{*n}(W_1,\dots,W_n)_k(Z_1,\dots,Z_k) = \\ \frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_k}\sum_{\ell\in\mathbf{N}^{*n},\,|\ell|=k} g_n\left(W_{1,\ell_1}(Z_{\sigma(1)}\cdots Z_{\sigma(\ell_1)}),\dots,W_{n,\ell_n}(Z_{\sigma(k-\ell_n+1)}\cdots Z_{\sigma(k)})\right) \end{gathered}$$

hence

$$\begin{gathered} |g_{*n}(W_1,\dots,W_n)_k(Z_1,\dots,Z_k)| \le \\ \frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_k}\sum_{\ell\in\mathbf{N}^{*n},\,|\ell|=k} |g_n||W_{1,\ell_1}|\cdots|W_{n,\ell_n}||Z_{\sigma(1)}|\cdots|Z_{\sigma(k)}| = \\ \le |g_n|\sum_{\ell\in\mathbf{N}^{*n},\,|\ell|=k}|W_{1,\ell_1}|\cdots|W_{n,\ell_n}||Z_1|\cdots|Z_k| \end{gathered}$$

and finally

$$|g_{*n}(W_1,\dots,W_n)_k| \le |g_n|\sum_{\ell\in\mathbf{N}^{*n},\,|\ell|=k}|W_{1,\ell_1}|\cdots|W_{n,\ell_n}|.$$

From this, we deduce the inequality

$$\sum_{k\in\mathbf{N}}|g_{*n}(W_1,\dots,W_n)_k| \le |g_n|\sum_{\ell\in\mathbf{N}^{*n}}|W_{1,\ell_1}|\cdots|W_{n,\ell_n}| = |g_n||W_1|_1\cdots|W_n|_1\,,$$

which does prove that $g_{*n}\in L^n_s\big(\mathcal{F}_0(E_1,F_1),\mathcal{F}_0(E_1,F_2)\big)$ and $|g_{*n}|\le|g_n|$. Thus, the strict convergence radius of the power series $\sum_{n\in\mathbf{N}^*} g_{*n}W^n$ is at least equal to that of $\sum_{n\in\mathbf{N}^*} g_nw^n$ and therefore positive, proving our result. $\square$

REMARK. Unfortunately, this part of the proof requires some modest calculations with power series[1], as we have been unable to stick (as in the case of differential equations) to Cauchy's viewpoint on holomorphic maps and Hadamard's strong maxim: "The shortest way between two truths in the real domain passes through the complex domain".

REFERENCES

[1] CARTAN, H. *Calcul différentiel*, Hermann, Paris, 1967

[2] CHAPERON, M. *Calcul différentiel et calcul intégral*, deuxième édition, Dunod, Paris, 2008

[3] DOUADY, A. Le problème des modules pour les sous-espaces analytiques compacts d'un espace analytique donné, *Ann. Inst. Fourier (Grenoble) 16* (1966), 1–95

[4] MALGRANGE, B. *Systèmes différentiels involutifs*, Panoramas et Synthèses 19, Société mathématique de France, 2005

[5] ROBBIN, J. On the existence theorem for differential equations, *Proc. Amer. Math. Soc. 19* (1968), 1005–1006

(Reçu le 30 avril 2009)

Marc Chaperon

Institut de mathématiques de Jussieu, Université Paris Diderot-Paris 7
UFR de Mathématiques, Case 7012
Bâtiment Chevaleret
F–75205 PARIS CEDEX 13
e-mail: chaperon@math.jussieu.fr

[1] A more traditional proof involving majorant series can easily be cooked up with the same ingredients...