We shall show that the derivatives of Φ can be obtained by formally differentiating inside the integral and applying the chain rule. However, one has to be a bit careful, since Φ maps into W and not into E.

Let us first consider the most difficult part, namely, the function $\Psi : U \rightarrow W$, where

$$\Psi(y)(t) = \int_0^t f(\tau, y(\tau)), \tag{9.4.52}$$

to which we want to apply the results of Lemma 9.4.3.

Define the mapping

$$\chi : \bar{J}_0 \rightarrow L(W, E) \tag{9.4.53}$$

by setting

$$\chi(\tau)y = y(\tau) \qquad \forall\, \tau \in \bar{J}_0,\ \forall\, y \in W. \tag{9.4.54}$$

Obviously, $\chi(\tau) \in L(W, E)$ and $\|\chi(\tau)\| \leq 1$. With the help of χ we shall be able to apply the chain rule.

First define

$$F : \bar{J}_0 \times U \rightarrow E \tag{9.4.55}$$

by setting

$$F(\tau, y) = f(\tau, \chi(\tau)y). \tag{9.4.56}$$

Since $\chi(\tau)$ is a continuous linear mapping, the partial derivatives of F with respect to y exist up to order m and the chain rule yields

$$D_2F(\tau, y) = D_2f(\tau, \chi(\tau)y) \circ \chi(\tau) \tag{9.4.57}$$

and for $2 \leq k \leq m$

$$D_2^kF(\tau, y) = D_2^kf(\tau, \chi(\tau)y) \circ (\chi(\tau), \dots, \chi(\tau)), \tag{9.4.58}$$

where $D_2^kF(\tau, y)$ is viewed as an element of $L_k(W; E)$; explicitly there holds for $h_i \in W$, $1 \leq i \leq k$,

$$D_2^kF(\tau, y)(h_1, \dots, h_k) = D_2^kf(\tau, \chi(\tau)y)(\chi(\tau)h_1, \dots, \chi(\tau)h_k). \tag{9.4.59}$$

With the help of the just defined mapping χ, the equation (9.4.52) can be written as

$$\chi(t)\Psi(y) = \int_0^t F(\tau, y). \tag{9.4.60}$$

For fixed $t \in \bar{J}_0$, $\chi(t)\Psi(y)$ is therefore of class $C^m(U, E)$ and also of class $C^{m,\alpha}(U, E)$, if f is of class $C^{m,\alpha}$, in view of Lemma 9.4.3, and

$$D^k\{\chi(t)\Psi(y)\} = \int_0^t D_2^kF(\tau, y), \qquad 0 \leq k \leq m, \tag{9.4.61}$$

and all $t \in \bar{J}_0$.

We shall prove in Lemma 9.4.7 that then $\Psi \in C^{m,\alpha}(U, W)$, such that, in view of the chain rule,

$$\chi(t) \circ D^k\Psi(y) = D^k\{\chi(t)\Psi(y)\} \qquad 0 \leq k \leq m. \tag{9.4.62}$$

The function Φ in equation (9.4.50) can written as

$$\Phi(\xi, y) = y - C\xi - \Psi(y), \tag{9.4.63}$$

where $C \in L(E, W)$ maps $\xi \in E$ to the constant function in W the image of which is ξ, completing the proof that Φ is of class C^m resp. $C^{m,\alpha}$.

Notice that any y satisfying $\Phi(\xi, y) = 0$ is an integral curve of f with initial value ξ, i.e., $y = x[\xi]$ restricted to the interval $\bar{J}_0$. Since by assumption $y_0 = x[\xi_0] \in U$, we now omit the reference that the integral curves have to be restricted to $\bar{J}_0$, it is tempting to use the implicit function theorem to deduce that there is a smaller ball $B_{\rho_0}(\xi_0)$, $\rho_0 \le \rho$, and a function $\varphi \in C^{m,\alpha}(B_{\rho_0}(\xi_0), U)$ satisfying the equation

$$\Phi(\xi, \varphi(\xi)) = 0 \qquad \forall\, \xi \in B_{\rho_0}(\xi_0). \tag{9.4.64}$$

We would then deduce that the flow $x = x(t, \xi)$ could be expressed as

$$x(t, \xi) = \varphi(\xi)(t) = \chi(t) \circ \varphi(\xi) \qquad \forall\, (t, \xi) \in J_0 \times B_{\rho_0}(\xi_0). \tag{9.4.65}$$

Moreover, differentiating (9.4.64) with respect to ξ we would obtain, in view of (9.4.63),

$$D\varphi - C - D\Psi(\varphi) \circ D\varphi = 0, \tag{9.4.66}$$

i.e., $D\varphi(\xi)(t) = \chi(t) \circ D\varphi(\xi)$ would be of class C^1 with respect t and would satisfy the differentiated flow equation with initial value $D\varphi(\xi)(0) = \mathrm{id}_E$, in view of (9.4.50), (9.4.66), (9.4.61) and (9.4.59).

In order to apply the implicit function theorem $A = D_2\Phi(\xi_0, y_0)$ has to be a topological isomorphism. From (9.4.63) we deduce

$$A = \mathrm{id}_W - D\Psi(y_0) \tag{9.4.67}$$

Let $h \in W$ satisfy $Ah = 0$, then we infer from the preceding equation and (9.4.61)

$$h(t) = \int_0^t D_2F(\tau, y_0)h = \int_0^t D_2f(\tau, y_0)h(\tau), \tag{9.4.68}$$

in view of (9.4.57), and due to Gronwall's lemma it follows $h = 0$.

If we can also prove that A is surjective, then the open mapping theorem, Theorem 6.4.7 on page 12, would imply $A \in L_{\mathrm{top}}(W, W)$.

Thus let $g \in W$, we then have to find $h \in W$ such that $Ah = g$ or equivalently,

$$h(t) = \int_0^t D_2f(\tau, y_0)h(\tau) + g(t) \qquad \forall\, t \in J_0. \tag{9.4.69}$$

If g were of class C^1, then this equation could be solved immediately, since h would be required to be a solution of the affine equation

$$\dot{h} = D_2f(t, y_0)h + \dot{g}, \tag{9.4.70}$$

which exists, cf. Remark 9.3.3 on page 139.

From the equation (9.4.65) and the remarks following it, we deduce that $y = D_2x(t,\xi)$ is a solution of the affine equation

$$\dot{y} = D_2f(t,x) \circ y \tag{9.4.79}$$

with initial condition $y(0,\xi) = \mathrm{id}_E$, and $y, \dot{y} \in C^0(J_0 \times B_{\rho_0}(\xi_0), L(E,E))$. The original flow $x = x(t,\xi)$ satisfies $x \in C^{0,1}(J_0 \times B_\rho(\xi_0), E)$, cf. part (1) of the proof. Assuming $m = 1$ we see that the coefficients of the affine equation are continuous or Hölder continuous of class $C^{0,\alpha}$ in both variables depending, if $f \in C^1$ or $f \in C^{1,\alpha}$. Stipulating that $\alpha = 0$, if f only satisfies $f \in C^1$, we conclude from Theorem 9.3.4 on page 140 that $y \in C^{0,\alpha}(J_0 \times B_{\rho_0}(\xi_0), L(E,E))$.

On the other hand, since x satisfies the flow equation, we infer $\dot{x} \in C^{0,1}(J_0 \times B_\rho(\xi_0), E)$ and hence $x \in C^{1,\alpha}(J_0 \times B_{\rho_0}(\xi_0), E)$. Employing the flow equation once again, we also obtain $\dot{x} \in C^{1,\alpha}(J_0 \times B_{\rho_0}(\xi_0), E)$.

Since $(t_0,\xi_0) \in \mathcal{D}(f)$ were arbitrary, the theorem is thus proved in the case $m = 1$.

Let us emphasize that in order to prove the regularity of the flow it is not necessary that the vector field is globally defined in a cartesian product $J \times \Omega$, it suffices to assume that f is defined in an open set $\Lambda \subset \mathbb{R} \times E$ containing $\mathcal{D}(f)$ satisfying the condition

(9.4.80) For any $(t_0,\xi_0) \in \mathcal{D}(f)$ there exists an open interval J containing 0 and t_0, an open set $\Omega \subset E$ and $\rho > 0$ such that $J \times \Omega \subset \Lambda$ and $x(t,\xi) \in \Omega$ for all $(t,\xi) \in J \times B_\rho(\xi_0)$.

(3) *The inductive step.*

Assume now that the theorem has already been proved for $m-1$, $m > 1$, and let f be of class $C^{m,\alpha}$. Then the flow of the vector field $f = f(t,x)$ is of class $C^{m-1,\alpha}(\mathcal{D}(f), E)$ by the inductive hypothesis. As we have proved in part (2), the function $y = D_\xi x$ is then well defined and is a solution of the affine equation (9.4.79), which can be viewed as a flow equation

$$\begin{aligned} &\dot{y} = F(t,\xi,y), \\ &y(0,\xi) = \mathrm{id}_E, \end{aligned} \tag{9.4.81}$$

The vector field F is defined in $\mathcal{D}(f) \times L(E,E)$ which is an open set contained in $J \times \Omega \times L(E,E)$ and $F \in C^{m-1,\alpha}(\mathcal{D}(f) \times L(E,E))$. In order to apply the inductive hypothesis it is advisable first to look at the modified flow problem

$$\begin{aligned} &\dot{z} = (0, F(t,\xi,z^2)), \\ &z(0,\xi,\eta) = (\xi,\eta), \end{aligned} \tag{9.4.82}$$

for arbitrary initial conditions $(\xi,\eta) \in \Omega \times L(E,E)$, where $z = z(t,\xi,\eta) = (z^1, z^2)$. The vector field Φ is now of the form $\Phi = (0,F)$.

Since the vector field F is affine, we conclude from Remark 9.3.3 on page 139 that $\mathcal{D}(\Phi) = \mathcal{D}(f) \times L(E,E)$. Let us check that the open set *are C^m-diffeomorphisms, where, in case of $x[t]$, we only consider $t \in J$ satisfying $\mathcal{D}(f)[t] \neq \emptyset$. If $f \in C^{m,\alpha}$, $0 < \alpha \leq 1$, then the diffeomorphisms are also of class $C^{m,\alpha}$.*

Proof. „(9.4.105)“ Let $\xi \in \mathcal{D}(f)[t]$. We already know that $x[t]$ is injective and of class C^m resp. $C^{m,\alpha}$, hence it suffices to prove, in view of Theorem 8.2.4 on page 107, that $Dx[t](\xi)$ is a topological isomorphism

$$Dx[t](\xi) \in L_{\text{top}}(E, E). \tag{9.4.107}$$

Let $J_0 = [0, t]$, where we assume without loss of generality $t > 0$, and set

$$A(\tau) = Dx[\tau](\xi) \qquad \tau \in J_0. \tag{9.4.108}$$

Then $A \in C^{m-1}(J_0, L(E, E))$[3] and is a solution of the linear initial value problem

$$\begin{aligned} \dot{A} &= D_2 f(\tau, x) \circ A, \\ A(0) &= I = \mathrm{id}_E, \end{aligned} \tag{9.4.109}$$

cf. (9.4.40), hence $A \in C^1(J_0, L(E, E))$.

By continuity $A(\tau) \in L_{\text{top}}(E, E)$ for small τ, $0 \leq \tau < \epsilon$. Let $J_\epsilon = [0, \epsilon)$ and

$$B(\tau) = A^{-1}(\tau) \qquad \tau \in J_\epsilon, \tag{9.4.110}$$

then

$$\dot{B} = -A^{-1}\dot{A}A^{-1} = -B\dot{A}B, \tag{9.4.111}$$

cf. exercise 4 of Exercises 3.5.5 of Analysis I, i.e., B solves the initial value problem

$$\begin{aligned} \dot{B} &= -B \circ D_2 f(\tau, x) \\ B(0) &= I \end{aligned} \tag{9.4.112}$$

in J_ϵ.

This is also a linear differential equation in the sense of Definition 9.3.2 on page 139 even though the product is written in the form $B \circ D_2 f(\tau, \xi)$ instead of $D_2 f(\tau, x) \circ B$. For linear initial value problems the solutions are defined on the whole interval in which the vector field is defined, cf. Remark 9.3.3 on page 139, the solution $B = B(\tau)$ of (9.4.112) exists for all $\tau \in J_0$, in particular for $\tau = t$, and there holds $B \in C^{m-1}(J_0, L(E, E))$.

Thus, if we could prove that

$$A(\tau)B(\tau) = B(\tau)A(\tau) = I \qquad \forall \tau \in J_0, \tag{9.4.113}$$

then (9.4.107) would immediately follow.

[3] Notice that A is defined in an open interval containing J_0.

Proof. It suffices to prove „(9.5.4)“. Differentiating

$$\varphi(t) = \Phi(t) \circ \Phi(\tau)^{-1}, \tag{9.5.8}$$

we deduce

$$\dot{\varphi} = \dot{\Phi}(t) \circ \Phi(\tau)^{-1} = A\Phi(t) \circ \Phi(\tau)^{-1} = A\varphi, \tag{9.5.9}$$

and of course there holds $\varphi(\tau) = \mathrm{id}_E$. ☐

Variation of constants

Consider now the inhomogeneous differential equation

$$\begin{aligned} \dot{x} &= A(t)x + \psi \\ x(\tau) &= \xi. \end{aligned} \tag{9.5.10}$$

Let $y = y(t)$ be a *special* solution of the inhomogeneous equation

$$\begin{aligned} \dot{y} &= A(t)y + \psi \\ y(\tau) &= 0, \end{aligned} \tag{9.5.11}$$

then the solution of (9.5.10) is represented as

$$x(t) = \Lambda(t, \tau)\xi + y(t). \tag{9.5.12}$$

In order to find a solution of (9.5.11) we use the ansatz

$$y(t) = \Lambda(t, \tau)\xi(t) \tag{9.5.13}$$

trying to determine $\xi(t)$ such that y is a solution.

Setting $\xi(t) \equiv$ const, then (9.5.13) is just a solution of the homogeneous equation with initial value ξ, cf. (9.5.5). Therefore the ansatz with variable ξ is often called *variation of constants.*

Differentiating the equation (9.5.13) yields

$$\begin{aligned} \dot{y} = \dot{\Lambda}\xi + \Lambda\dot{\xi} = A\Lambda\xi + \Lambda\dot{\xi} &= Ay + \Lambda\dot{\xi} \\ &\overset{!}{=} Ay + \psi, \end{aligned} \tag{9.5.14}$$

in view of (9.5.4), where the symbol „$\overset{!}{=}$“ means „should be equal to“.

Thus, if y is supposed to be a solution of (9.5.11), then there must hold

$$\Lambda\dot{\xi} = \psi \tag{9.5.15}$$

or equivalently,

$$\dot{\xi} = \Lambda(t, \tau)^{-1}\psi(t), \tag{9.5.16}$$

but this can also be expressed as

$$\xi(t) = \int_\tau^t \Lambda(s, \tau)^{-1}\psi(s) \underset{(9.5.7)}{=} \int_\tau^t \Lambda(\tau, s)\psi(s), \tag{9.5.17}$$