The Flow of ODEs

Fabian Immler* and Christoph Traut

Institut für Informatik, Technische Universität München
{immler, trautc}@in.tum.de

Abstract. Formal analysis of ordinary differential equations (ODEs) and dynamical systems requires a solid formalization of the underlying theory. The formalization needs to be at the correct level of abstraction, in order to avoid drowning in tedious reasoning about technical details. The *flow* of an ODE, i.e., the solution depending on initial conditions, and a dedicated type of bounded linear functions yield suitable abstractions. The dedicated type integrates well with the type-class based analysis in Isabelle and we prove advanced properties of the flow, most notably, differentiable dependence on initial conditions via the variational equation and a rigorous numerical algorithm to solve it.

1 Introduction

Ordinary differential equations (ODEs) are ubiquitous for modeling continuous problems in e.g., physics, biology, or economics. A formalization of the theory of ODEs allows us to verify algorithms for the analysis of such systems. A popular example, where a verified algorithm is highly relevant, is Tucker's proof on the topic of a strange attractor for the Lorenz equations [9]. This proof relies on the output of a computer program, that computes bounds for analytical properties of the so-called *flow* of an ODE.

The flow is the solution as a function depending on an initial condition. We formalize the flow and prove conditions for analytical properties like continuity of differentiability (the derivative is of particular importance in Tucker's proof). Most of these properties seem very "natural", as Hirsch, Smale and Devaney call them in their textbook [2]. However, despite being "natural" properties and fairly standard results, they are delicate to prove: In the textbook, the authors present these properties rather early, but

> "postpone all of the technicalities [...], primarily because understanding this material demands a firm and extensive background in the principles of real analysis."

In this paper, we show that it is feasible to cope with these technicalities in a formal setting and confirm that Isabelle/HOL supplies a sufficient background of real analysis.

* Supported by the DFG RTG 1480 (PUMA)

We present our Isabelle/HOL library for reasoning about the flow of ODEs. The main results are formalizations of continuous and differentiable dependence on initial conditions. The differentiable dependence is characterized by a particular ODE, the variational equation, and we show how to use existing rigorous numerical algorithms to solve it (section 4). The variational equation is posed on the space of linear functions. We introduce a separate type for this space in order to profit from the type class based formalization of mathematics in Isabelle/HOL.

We are not aware of any other formalization that covers this foundational part of the theory of ODEs in similar detail.

2 Overview

We will first (in section 3) present the "interface" to our theory, i.e., the definitions and assumptions that are needed for formalizing our main results. Any potential user of the library needs in principle only know about these concepts. Because the general topic is very theoretical and foundational work, we present a practical application right afterwards in section (section 4)

Only then, we go into the details of the techniques that we used to make this formalization possible. Mathematics and analysis is formalized in Isabelle mostly based on type classes and filters, as has been presented earlier in earlier work [3]. We follow this path to formalize the foundations of our work:

Several proofs needed the notion of a uniform limit. We cast this notion into the "Isabelle/HOL approach to limits": we define it using a filter. This gives a versatile formalization and one can profit from the existing infrastructure for filters in limits. This will be presented in section 5

The derivative of the flow is a linear function. The space of linear functions forms a Banach space. In order to profit from the structure and properties that hold in a Banach space (which is a type class in Isabelle/HOL), we needed to introduce a type of bounded linear functions. We will present this type and further applications of its formalization in section 6.

In section 7, we present the technical lemmas that are needed to prove continuity and differentiability of the Flow in order to give an impression of the kind of reasoning that is required.

All of the theorems we present here and in the following are formalized in Isabelle/HOL [8], the source code can be found in the development version of the Archive of Formal Proof[1].

3 The Flow of a Differential Equation

In this section, we introduce the concept of *flow* and *existence interval* (which guarantees that the flow is well-defined) and present our main results (without

[1] `http://www.isa-afp.org/devel-entries/Ordinary_Differential_Equations.shtml`

proofs at first, we will present some of the lemmas leadings to the proofs in section 7).

The claim we want to make in this section is the flow as definition is a suitable abstraction for initial value problems. But beware: do not get deceived by simplicity of statements: as already mentioned in the introduction, these are all "natural" properties, but the proofs (also in the textbook) require many technical lemmas.

First of all, let us introduce the concepts we are interested in. We consider open sets T, X and an autonomous[2] ODE with right hand side f

$$\dot{x}(t) = f(x(t)), \text{ where } f : \mathbb{R}^n \to \mathbb{R}^n \text{ is a function from } X \text{ to } X \tag{1}$$

Under mild assumptions (which we will make more precise later in definition 27), there exists a solution $\varphi(t)$, which is unique for an initial condition $x(t_0) = x_0$. To emphasize the dependence on the initial condition, we write $\varphi(x_0, t)$ for the solution of equation (1). This solution depending on initial conditions is called the *flow* of the differential equation:

Definition 1 (Flow). *The* flow $\varphi(x_0, t)$ *is the (unique) solution of the ODE (1) with initial condition* $x\ 0 = x_0$

The solution does not necessarily exist for every $t \in T$. For example, solutions can *explode* in finite time s: if $\lim_{t \to s} \varphi(x_3, t) = \infty$, then the flow is only defined for $t < s$ as is illustrated in figure 1 for $\varphi(x_3, \_)$. We therefore need to define a notion of (maximal) existence interval.

Definition 2 (Maximal Existence Interval). *The* maximal existence interval *of the ODE (1) is the open interval*

$$\textsf{ex-ivl}(x_0) :=]\alpha; \beta[$$

for $\alpha, \beta \in \mathbb{R} \cup \{\infty, -\infty\}$*, such that* $\varphi(x_0, t)$ *is a solution for* $t \in \textsf{ex-ivl}$*. Moreover for every other interval* I *and every solution* $\psi(x_0, t)$ *for* $t \in I$*, one has* $I \subseteq J$ *and* $\forall t \in I.\ \psi(x_0, t) = \varphi(x_0, t)$.

We claim that the flow φ (together with $\textsf{ex-ivl}$, which guarantees the flow to be well-defined) is a very nice way to talk about solutions, because after guaranteeing that they are well-defined, these constants have many nice properties, which can be stated without further assumptions.

3.1 Composition of solutions

A first nice property is the abstract property of the generic notion of flow. This notion makes it possible to easily state composition of solutions and to algebraically reason about them. As illustrated in figure 1, flowing from x_1 for time $s + t$ is equivalent to first flowing for time s, and from there flowing for time t.

This only works if the flow is defined also for the intermediate times (the theorem can not be true for $\varphi(x_0, t + (-t))$ if $t \notin \textsf{ex-ivl}$).

[2] this means that f does not depend on t. Many of our results are also formalized for non-autonomous ODEs, but the presentation is clearer, and reduction is possible.

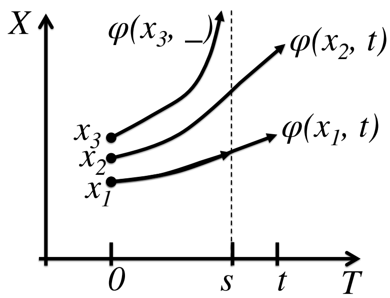

Fig. 1: The flow for different initial values

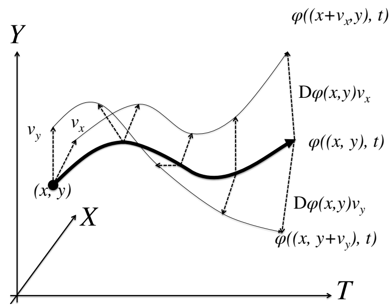

Fig. 2: Illustration of the derivative of the flow

Theorem 3 (Flow property).

$$\{s, t, s+t\} \subseteq \textsf{ex-ivl}\,(x_0) \implies \varphi(x_0, s+t) = \varphi(\varphi(x_0, s), t)$$

3.2 Continuity of the Flow

In the previous lemma, the assumption that the flow is defined (i.e., that the time is contained in the existence interval) was important. Let us now study the domain $\Omega = \{(x, t) \mid \ t \in \textsf{ex-ivl}\,(x)\} \subseteq X \times T$ of the flow in more detail. Ω is called the *state space*.

For the first "natural" property, we consider an element in the state space. $(t, x) \in \Omega$ means that we can follow a solution starting at x for time t. It is "natural" to expect that solutions starting close to x can be followed for times that are close to t. In topological parlance, the state space is open.

Theorem 4 (Open State Space). $\textsf{open}\,\Omega$

In the previous theorem, the state space allows us to reason about the fact that solutions are *defined* for close times and initial values. For quantifying how deviations in the initial values are propagated by the flow, Grönwall's lemma is an important tool that is used in several proofs. Because of its importance in the theory of dynamical systems, we list it here as well, despite it being a rather technical result. Starting from an *implicit* inequality $g\ t \leq C + K \cdot \int_0^t g(s)\,\mathrm{d}s$ involving a continuous, nonnegative function $g : \mathbb{R} \to \mathbb{R}$, it allows one to deduce an *explicit* bound for g:

Lemma 5 (Grönwall).

$$0 < C \implies 0 < K \implies \textsf{continuous-on}\ [0; a]\ g \implies$$
$$\forall t.\ 0 \leq g(t) \leq C + K \cdot \int_0^t g(s)\,\mathrm{d}s \implies$$
$$\forall t \in [0; a].\ g(t) \leq C \cdot e^{K \cdot t}$$

Grönwall's lemma can be used to show that solutions deviate *at most* exponentially fast: $\exists K.\ |\varphi(x,t) - \varphi(y,t)| < |x-y| \cdot e^{K \cdot |t|}$ (see also Lemma 30). Therefore, by choosing x and y close enough, one can make the distance of the solutions arbitrarily small. In other words, the flow is a continuous function on the state space:

Theorem 6 (Continuity of Flow). $\textsf{continuous-on}\ \Omega\ \varphi$

3.3 Differentiability of the Flow

Continuity just states that small deviation in the initial values result in small deviations of the flow. But one can be more precise on how initial deviations propagate. Consider figure 2, which depicts a solution starting at (x, y) and its evolution up to time t, as well as two other solutions evolving from initial values that have been perturbed via vectors v_x and v_y, respectively. A nice property of the flow is that it is differentiable: the way initial deviations propagate can be approximated by a linear function. So instead of solving the ODE for perturbed initial values, one can approximate the resulting perturbation with the linear function: $\mathsf{D}\varphi \cdot v \approx \varphi((x, y), t) - \varphi((x, y) + v, t)$. More formally, our main result is the formalization of the fact that the derivative of the flow exists and is continuous.

Theorem 7 (Differentiability of the Flow). *For every $(x, t) \in \Omega$ There exists a linear function $W(x, t)$, which is the derivative of the flow at (x, t):*

$$\exists W.\ \mathsf{D}\varphi|_{(x,t)} = W(x,t) \wedge \textsf{continuous-on}\ \Omega\ W$$

4 Rigorous Numerics for the Derivative of the Flow

In this section, we show that the formalization is not something abstract and detached, but something that can actually be computed with: The derivative W of the flow can be characterized as the solution of a linear, matrix-valued ODE, a byproduct of the (constructive) proof of differentiability in lemma 36: The derivative with respect to x, written W_x, is the solution to the following ODE[3]

$$\dot{W}(t) = \mathsf{D}f|_{\varphi(x_0,t)} \cdot W(t)$$

with initial condition $W(0)$ is the identity matrix.

We encode this matrix valued variational equation into a vector valued one, use an existing rigorous numerical algorithm for solving ODEs in Isabelle [5] to compute bounds on the solutions. Re-interpreting the result as bounds on matrices, we obtain bounds on the solution of the variational equation. As a concrete example, we use the van der Pol system: $\dot{x} = y; \quad \dot{y} = (1 - x^2)y - x$ for initial condition $(x_0, y_0) = (1.25, 2.27)$.

[3] here, $\cdot$ stands for matrix multiplication

The overall setup for the computation is as follows: We have an executable specification of the Euler method (which is formally verified to produce rigorous enclosures for the solution of an ODE) and use Isabelle's code generator [1] to generate SML code from this specification. We chose to compute the evolution until time $t = 2$ with a discrete grid of 500 time steps. The computation takes about 3 minutes on an average laptop computer. As a result, we get the following inclusion for the variational equation:

Theorem 8.

$$W(2) \in \begin{pmatrix} [0.18; 0.23] & [0.41; 0.414] \\ [-0.048; -0.041] & [0.26; 0.27] \end{pmatrix}$$

The left column of the matrix shows the propagation of a deviation in the x direction: a $(1, 0)$ deviation is propagated to a $([0.18; 0.23], [-0.048; -0.041])$ deviation: it gets smaller but remains mostly in the x direction. For the right column, a deviation in the y direction $(0, 1)$ is propagated to a $([0.41; 0.414]; [0.26; 0.27])$ deviation: it contracts as well, but it gets rotated towards the x direction.

5 Uniform Limit as Filter

Filters have proved to be useful to describe all kinds of limits and convergence [3]. We use filters to define uniform convergence. For details about filters, please consider the source code and the paper [3]. In the formalization, the uniform limit *uniform-limit* $X\ f\ l\ F$ is parameterized by a filter F, here we just present the explicit formulations for the *sequentially* and *at* filters.

A sequence of functions $f_n : \alpha \to \beta$ for $n \in \mathbb{N}$ is said to converge uniformly on $X : \mathcal{P}(\alpha)$ against the *uniform limit* $l : \alpha \to \beta$, if

Definition 9.

$$\begin{aligned}&\textit{uniform-limit}\ X\ f\ l\ \textit{sequentially} := \\ &\quad \forall \varepsilon > 0.\ \exists N.\ \forall x \in X.\ \forall n \geq N.\ |f_n\ x - l\ x| < \varepsilon\end{aligned}$$

Note the difference to pointwise convergence, where one would exchange the order of the quantifiers $\exists N$ and $\forall x \in X$.

With the ($\textit{at}\ z$) filter, we can also handle uniform convergence of a family of functions $f_y : \alpha \to \beta$ as y approaches z:

Definition 10.

$$\begin{aligned}&\textit{uniform-limit}\ X\ f\ l\ (\textit{at}\ z) := \\ &\quad \forall \varepsilon > 0.\ \exists \delta > 0. \quad \forall y.\ |y - z| < \delta \implies (\forall x \in X.\ \textit{dist}\ (f_y\ x)\ (l\ x) < \varepsilon)\end{aligned}$$

The advantage of the filter approach is that many important lemmas can be expressed for arbitrary filters, for example the uniform limit theorem, which states that the uniform limit of a (via filter F generalized) sequence f_n of continuous functions is continuous.

Theorem 11 (Uniform Limit Theorem).

$$(\forall n \in F.\ \textit{continuous-on}\ X\ f_n) \implies \textit{uniform-limit}\ X\ f\ l\ F \implies \textit{continuous-on}\ X\ f_n$$

A frequently used criterion to show that an infinite series of functions converges uniformly is the Weierstrass M-test. Assuming majorants M_n for the functions f_n and assuming that the series of majorants converges, it allows one to deduce uniform convergence of the partial sums towards the series.

Lemma 12 (Weierstrass M-Test).

$$\forall n.\ \forall x \in X.\ |f_n\ x| \le M_n \implies \sum_{n \in \mathbb{N}} M_n < \infty \implies$$
$$\textit{uniform-limit}\ X\ (n \mapsto x \mapsto \sum_{i \le n} f_i\ x)\ (x \mapsto \sum_{i \in \mathbb{N}} f_i\ x)\ \textit{sequentially}$$

6 Bounded Linear Functions

We introduce a type of bounded linear functions (or equivalently *continuous* linear functions) in order to be able to profit from the hierarchy of mathematical type classes in Isabelle/HOL.

6.1 Type Classes for Mathematics in Isabelle/HOL

In Isabelle/HOL, many of the mathematical concepts (in particular spaces with a certain structure) are formalized using type classes.

The advantage of type class based reasoning is that most of the reasoning is generic: formalizations are carried out in the context of type classes and can then be used for all types inhabiting that type class. For generic formalizations, we use Greek letters α, β, γ and name their type class constraints in prose (i.e., if we write that we "consider a topological space" α, then this result is formalized generically for every type α that fulfills the properties of a topological space).

The spaces we consider are topological spaces with *open* sets, (real) vector spaces with addition $+ : \alpha \to \alpha \to \alpha$ and scalar multiplication $\cdot : \mathbb{R} \to \alpha \to \alpha$. Normed vector spaces come with a norm $|(\_)| : \alpha \to \mathbb{R}$. A vector space with multiplication $* : \alpha \to \alpha \to \alpha$ that is compatible with addition $(a + b) * c = a*c+b*c$ is an algebra and can also be endowed with a norm. Complete normed vector spaces are called Banach spaces.

6.2 A Type of Bounded Linear Functions

An important concept is that of a linear function. For vector spaces α and β, a linear function is a function $f : \alpha \to \beta$ that is compatible with addition and scalar multiplication.

Definition 13.

$$\textit{linear}\ f := \forall x\, y\, c.\ f(c \cdot x + y) = c \cdot f(x) + f(y)$$

We need topological properties of linear functions, we therefore now assume normed vector spaces α and β. One usually wants linear functions to be continuous, and if α and β are vector spaces of finite dimension, any linear function $\alpha \to \beta$ is continuous. In general, this is not the case, and one usually assumes *bounded* linear functions. The norm of the result of a bounded linear function is linearly bounded by the norm of the argument:

Definition 14.

$$\textit{bounded-linear}\ f := \textit{linear}\ f \wedge \exists K.\ \forall x.\ |f(x)| \leq K * |x|$$

We now cast bounded linear functions $\alpha \to \beta$ as a type $\alpha \to_{bl} \beta$ in order to make it an instance of Banach space.

Definition 15.

$$\textbf{\textit{typedef}}\ \alpha \to_{bl} \beta := \{f : \alpha \to \beta \mid \textit{bounded-linear}\ f\}$$

6.3 Instantiations

For defining operations on type $\alpha \to_{bl} \beta$, the Lifting and Transfer package [4] is an essential tool: operations on the plain function type $\alpha \to \beta$ are automatically lifted to definitions on the type $\alpha \to_{bl} \beta$ when supplied with a proof that functions in the result are *bounded-linear* under the assumption that argument functions are *bounded-linear*. We write application of a bounded linear function $f : \alpha \to_{bl} \beta$ with an element $x : \alpha$ as follows.

Definition 16 (application of bounded linear functions).

$$(f \cdot x) : \beta$$

We present the definitions of operations involving the type $\alpha \to_{bl} \beta$ by presenting them in an extensional form using $\cdot$. Bounded linear functions with pointwise addition and pointwise scalar multiplication form a vector space.

Definition 17 (Vector Space Operations). *For $f, g : \alpha \to_{bl} \beta$ and $c : \mathbb{R}$,*

$$(f + g) \cdot x := f \cdot x + g \cdot x$$

$$(c \cdot f) \cdot x := c \cdot (f \cdot x)$$

The usual choice of a norm for bounded linear functions is the operator norm: the maximum of the image of the bounded linear function on the unit ball. With this norm, $\alpha \to_{bl} \beta$ forms a normed vector space and we prove that it is Banach if α and β are Banach.

Definition 18 (Norm in Banach Space). *For* $f : \alpha \rightarrow_{bl} \beta$*,*

$$|f| := \max \{|f \cdot y| \mid |y| \leq 1\}$$

One can also compose bounded linear functions according to $(f \circ g) \cdot x = f \cdot (g \cdot x)$. Bounded linear operators—that is bounded linear functions $\alpha \rightarrow_{bl} \alpha$ from one type α into itself—form a Banach algebra with composition as multiplication and the identity function as neutral element:

Definition 19 (Banach Algebra of Bounded Linear Operators). *For* $f, g : \alpha \rightarrow_{bl} \alpha$*,*

$$(f * g) \cdot x := (f \circ g) \cdot x$$

$$1 \cdot x := x$$

6.4 Applications

Now we can profit from many of the developments that are available for Banach spaces or algebras. Here we present some useful applications: The exponential function is defined generically for *banach-algebra* and can therefore be used for bounded linear functions as well. Furthermore, the type of bounded linear functions can be used to describe derivatives in arbitrary vector spaces and therefore allows one to naturally express (and conveniently prove) basic results from analysis: the Leibniz rule for differentiation under the integral sign and conditions for (total) differentiability of multidimensional functions. Note that not everything in this section is directly necessary for the formalizations of our main results, it is rather intended to show the versatile use of a separate type for bounded linear functions in Isabelle/HOL.

Exponential of operators The exponential function for bounded linear functions is a useful concept and important for the analysis linear ODEs. Here we present that the solution of linear autonomous homogeneous differential equations can be expressed using the exponential function. For a Banach algebra α, the exponential function is defined using the usual power series definition (B^k is a k fold multiplication $B * \cdots * B$):

Definition 20 (Exponential Function). *For a Banach algebra* α *and* $B : \alpha$*,*

$$e^B := \sum_{k=0}^{\infty} \frac{1}{k!} \cdot B^k$$

We prove the following rule for the derivative of the exponential function

Lemma 21 (Derivative of Exponential). $\frac{\mathrm{d}\, e^{x \cdot A}}{\mathrm{d}x} = e^{x \cdot A} \cdot A$

Proof. After unfolding the definition of derivative $\frac{\mathrm{d}\, e^{x \cdot A}}{\mathrm{d}x} = \lim_{h \to 0} \frac{e^{(x+h)\cdot A} - e^{x \cdot A}}{h}$, the crucial step in the proof is to exchange the two limits (one is explicit in $\lim_{h\to 0}$, and the other one is hidden as the limit of the series definition 20 of the exponential). Exchange of limits can be done similar to Theorem 11, while uniform convergence is guaranteed according to the Weierstrass M-Test from Lemma 12. □

With this rule for the derivative and an obvious calculation for the initial value, one can show the following

Lemma 22 (Solution of linear initial value problem). $\varphi_{x_0,t_0}(t) := \left(e^{(t-t_0)\cdot A}\right)(x_0)$ *is the unique solution to the ODE* $\dot{\varphi}\ t = A\ (\varphi\ t)$ *with initial condition* $\varphi(t_0) = x_0$.

Total Derivatives The total derivative (or Fréchet derivative) is a generalization of the ordinary derivative (of functions $\mathbb{R} \to \mathbb{R}$) for arbitrary normed vector spaces. To illustrate this generalization, recall that the ordinary derivative yields the slope of the function: if $f'(x) = m$, then

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = m \tag{2}$$

Moving the m under the limit, one sees that the (linear) function $h \mapsto h \cdot m$ is a good approximation for the difference of the function value at nearby points x and $x + h$:

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - h \cdot m}{h} = 0$$

This concept can be generalized by replacing $h \mapsto h \cdot m$ with an arbitrary (bounded) linear function A. In the following equation, A is a good linear approximation.

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - A \cdot h}{|h|} = 0 \tag{3}$$

Note that in the previous equation, we can (just formally) drop many of the restrictions on the type of f. We started with $f : \mathbb{R} \to \mathbb{R}$ in equation 2, but the last equation still makes sense for $f : \alpha \to \beta$ for normed vector spaces α, β. We call $A : \alpha \to_{bl} \beta$ the total derivative $\mathsf{D}f$ of f at a point x:

Definition 23 (Total Derivative). *For* $A : \alpha \to_{bl} \beta$ *in equation 3, we write*

$$\mathsf{D}f|_x = A$$

The total derivative is important for our developments as it is for example the derivative W of the flow in Theorem 7. It is only due to the fact that the resulting type $\alpha \to_{bl} \alpha$ is a normed vector space, that makes it possible to express continuity of the derivative or to express higher derivatives.

Another example, where interpreting the derivative as bounded linear function $\alpha \to_{bl} \beta$ is helpful, is when deducing the total derivative of a function f by looking at its partial derivatives f_1 and f_2 (that is, the derivatives w.r.t. one variable while fixing the other). One needs the assumption that the partial derivatives are continuous.

Lemma 24 (Total Derivative via Continuous Partial Derivatives). *For* $f : \alpha \to \beta \to \gamma$, $f_1 : \alpha \to \beta \to (\alpha \to_{bl} \gamma)$, $f_2 : \alpha \to \beta \to (\beta \to_{bl} \gamma)$

$$\begin{aligned}
&\forall x.\ \forall y.\ \mathsf{D}(x \mapsto f\ x\ y)|_x = f_1\ x\ y \implies \\
&\forall x.\ \forall y.\ \mathsf{D}(y \mapsto f\ x\ y)|_y = f_2\ x\ y \implies \\
&\mathit{continuous}((x,y) \mapsto f_1\ x\ y) \implies \\
&\mathit{continuous}((x,y) \mapsto f_2\ x\ y) \implies \\
&\mathsf{D}((x,y) \mapsto f\ x\ y)|_{(x,y)} \cdot (t_1, t_2) = (f_1\ x\ y) \cdot t_1 + (f_2\ x\ y) \cdot t_2
\end{aligned}$$

Leibniz rule Another example is a general formulation of the Leibniz rule. The following rule is a generalization of e.g., the rule formalized by Lelay and Melquiond [7] to general vector spaces. Here $[[a;b]]$ is a hyperrectangle in Euclidean space $\mathbb{R}^n$. The rule allows one to differentiate under the integral sign: the derivative of the parameterized integral $\int_a^b f\ x\ t\,\mathrm{d}t$ with respect to x can be expressed as the integral of the derivative of f. Note that the integral on the right is in the Banach space of bounded linear functions.

Lemma 25 (Leibniz rule). *For Banach spaces* α, β *and* $f : \alpha \to \mathbb{R}^n \to \beta$, $f_1 : \alpha \to \mathbb{R}^n \to (\alpha \to_{bl} \beta)$,

$$\begin{aligned}
&\forall t.\ \mathsf{D}(x \mapsto f\ x\ t)|_x = f_1\ x\ t \implies \\
&\forall x.\ (f\ x)\ \mathit{integrable\text{-}on}\ [[a;b]] \\
&\forall xt.\ t \in [[a;b]] \implies \mathit{continuous}((x,t) \mapsto f\ x\ t) \\
&\mathsf{D}\left(x \mapsto \int_a^b f\ x\ t\,\mathrm{d}t\right)|_x = \int_a^b f_1\ x\ t\,\mathrm{d}t
\end{aligned}$$

7 Proofs about the Flow

We will now go into the technical details of the proofs leading towards continuity and differentiability of the flow (Theorems 6 and 7). We still do not present the proofs: their structure is very similar to the textbook [2] proofs. Nevertheless, we want to present the detailed statements of the propositions, as they give a good impression on the kind of reasoning that was required.

7.1 Criteria for Unique Solution

First of all, we specify the common assumptions to guarantee existence of a unique solution for an initial value problem and therefore a condition for the flow in definition 1 to be well-defined.

We assume that f is locally Lipschitz continuous in its second argument: for every $(t, x) \in T \times X$ there exist ε-neighborhoods $U_\varepsilon(t)$ and $U_\varepsilon(t)$ around t and x, in which f is Lipschitz continuous w.r.t. the second argument (uniformly w.r.t. the first): the distance of function values is bounded by a constant times the distance of argument values:

Definition 26.

$$\begin{aligned}
&\textsf{\textit{local-lipschitz}}\ T\ X\ f :=\\
&\forall t \in T.\ \forall x \in X.\\
&\quad \exists \varepsilon > 0.\ \exists L.\\
&\qquad \forall t' \in U_\varepsilon(t).\ \forall x_1, x_2 \in U_\varepsilon(x).\ |f\ t'\ x_1 - f\ t'\ x_2| \le L \cdot |x_1 - x_2|
\end{aligned}$$

Now the only assumptions that we need to prove continuity of the flow are open sets for time and phase space and a locally Lipschitz continuous right-hand side f that is continuous in t:

Definition 27 (Conditions for unique solution).

1. *T is an open set*
2. *X is an open set*
3. *f is locally Lipschitz continuous on X: $\textsf{\textit{local-lipschitz}}\ T\ X\ f$*
4. *for every $x \in X$, $t \mapsto f(t, x)$ is continuous on T.*

These assumptions (the detailed proofs that these assumptions guarantee the existence of a unique solution for initial value problems has been presented in Theorem 3 of earlier work [6]).

7.2 The Frontier of the State Space

It is important to study the behavior of the flow at the frontier of the state space (e.g., as time or the solution tend to infinity). From this behavior, one can deduce conditions under which solutions can be continued. This yields techniques to gain more precise information on the existence interval *ex-ivl*.

If the solution only exists for finite time, it has to explode (i.e., leave every compact set):

Lemma 28 (Explosion for Finite Existence Interval).

$$\begin{aligned}
&\textsf{\textit{ex-ivl}}(x_0) =]\alpha, \beta[\implies \beta < \infty \implies \textsf{\textit{compact}}\ K \implies\\
&\exists t \ge 0.\ t \in \textsf{\textit{ex-ivl}}(x_0) \wedge \varphi(x_0, t) \notin K
\end{aligned}$$

This lemma can be used to prove a condition on the right-hand side f of the ODE, to certify that the solution exists for the whole time. Here the assumption guarantees that the solution stays in a compact set.

Lemma 29 (Global Existence of Solution).

$$\begin{aligned}
&(\forall s \in T.\ \forall u \in T.\ \exists L.\ \exists M.\ \forall t \in [s; u].\ \forall x \in X.\ |f\ t\ x| \le M + L \cdot |x|)\\
&\quad \implies \textsf{\textit{ex-ivl}}(x_0) = T
\end{aligned}$$

7.3 Continuity of the Flow

The following lemmas are all related to continuity of the flow. With the help of Grönwall's lemma 5, one can show that when two solutions (starting from different initial conditions x_0 and y_0) both exist for a time t and are restricted to some set Y on which the right-hand side f satisfies a (global) Lipschitz condition K, then the distance between the solutions grows at most exponentially with increasing time:

Lemma 30 (Exponential Initial Condition for Two Solutions).

$$
\begin{aligned}
& t \in \textit{ex-ivl}(x_0) \implies t \in \textit{ex-ivl}(y_0) \implies \\
& x_0 \in Y \implies y_0 \in Y \implies Y \subseteq X \implies \\
& \forall s \in [0;t].\ \varphi(x_0, s) \in Y \implies \\
& \forall s \in [0;t].\ \varphi(y_0, s) \in Y \implies \\
& \forall s \in [0;t].\ \textit{lipschitz}\ Y\ (f\ s)\ K \implies \\
& |\varphi(x_0, t) - \varphi(y_0, t)| \leq |x_0 - y_0| \cdot e^{K \cdot t}
\end{aligned}
$$

Note that it can be hard to establish the assumptions of this lemma, in particular the assumption that both solutions from x_0 and y_0 exist for the same time t. Consider figure 1: not all solutions (e.g., from x_3) do necessarily exist for the same time s. One can choose, however, a neighborhood of x_1, such that all solutions starting from within this neighborhood exist for at least the same time, and with the help of the previous lemma, one can show that the distance of these solutions increases at most exponentially:

Lemma 31 (Exponential Initial Condition of Close Solutions).

$$
\begin{aligned}
& a \in \textit{ex-ivl}(x_0) \implies b \in \textit{ex-ivl}(x_0) \implies a \leq b \\
& \exists \delta > 0.\ \exists K > 0.\ U_\delta(x_0) \subseteq X\ \wedge \\
& \quad (\forall y \in U_\delta(x_0).\ \forall t \in [a;b]. \\
& \qquad t \in \textit{ex-ivl}(y) \wedge |\varphi(x_0, t) - \varphi(y, t)| \leq |x_0 - y| \cdot e^{K \cdot |t|})
\end{aligned}
$$

Using this lemma is the key to showing continuity of the flow (theorem 6).

A different kind of continuity is not with respect to the initial condition, but with respect to the right-hand side of the ODE.

Lemma 32 (Continuity with respect to ODE). *Assume two right-hand sides f, g defined on X and uniformly close $|f\ x - g\ x| < \varepsilon$. Furthermore, assume a global Lipschitz constant K for f on X. Then the deviation of the flows φ_f and φ_g can be bounded:*

$$|\varphi_f(x_0, t) - \varphi_g(x_0, t)| \leq \frac{\varepsilon}{K} \cdot e^{K \cdot t}$$

7.4 Differentiability of the Flow

The proof for the differentiability of the flow incorparates many of the tools that we have presented up to now, we will therefore go a bit more into the details of this proof.

Assumptions The assumptions in definition 27 are not strong enough to prove differentiability of the flow. However, a continuously differentiable right-hand side $f : \mathbb{R}^n \to \mathbb{R}^n$ suffices. To be more precise:

Definition 33 (Criterion for Continuous Differentiability of the Flow).

$$\exists f' : \mathbb{R}^n \to (\mathbb{R}^n \to_{\mathsf{bl}} \mathbb{R}^n).\ (\forall x \in X.\ \mathsf{D}f|_x = f'\ x) \wedge \mathsf{continuous\text{-}on}\ \ X\ \ f'$$

From now on, we denote the derivative along the flow from x_0 with $A_{x_0} : \mathbb{R} \to \mathbb{R}^n$:

Definition 34 (Derivative along the Flow). $A_{x_0}(t) := \mathsf{D}f|_{\varphi(x_0,t)}$

The derivative of the flow is the solution to the so-called variational equation, a non-autonomous linear ODE. The initial condition ξ is supposed to be a perturbation of the initial value (like v_x and v_y in figure 2) and in what follows we will prove that the solution to this ODE is a good (linear) approximation of the propagation of this perturbation.

$$\begin{cases} \dot{u}(t) = A_{x_0}(t) \cdot u(t) \\ u(0) = \xi \end{cases}, \tag{4}$$

We will write $u_{x_0}(\xi, t)$ for the flow of this ODE and omit the parameter x_0 and/or the initial value ξ if they can be inferred from the context.

As a prerequisite for the next proof, we begin by proving that $u_{x_0}(\xi, t)$ is linear in ξ, a property that holds because u is the solution of a linear ODE (this is often also called the "superposition principle").

Lemma 35 (Linearity of $u_{x_0}(\xi, t)$ in ξ).

$$\alpha \cdot u_{x_0,a}(t) + \beta \cdot u_{x_0,b}(t) = u_{x_0,\alpha\cdot a+\beta\cdot b}(t).$$

Because $\xi \mapsto u_{x_0}(\xi, t) : \mathbb{R}^n \to \mathbb{R}^n$ is linear on Euclidean space, it is also bounded linear, so we will identify this function with the corresponding element of type $\mathbb{R}^n \to_{\mathsf{bl}} \mathbb{R}^n$. The main efforts go into proving the following lemma, showing that the aforementioned function is the derivative of the flow $\varphi(x_0, t)$ in x_0.

Lemma 36 (Space Derivative of the Flow). *For* $t \in \mathsf{ex\text{-}ivl}(x_0)$,

$$(\mathsf{D}(x \to \varphi(x, t))|_{x_0}) \cdot \xi = u_{x_0}(\xi, t)$$

Proof. The proof starts out with the integral identities of the flow, the perturbed flow, and the linearized propagation of the perturbation:

$$\begin{aligned} \varphi(x_0, t) &= x_0 + \int_0^t f(\varphi(x_0, s))\,\mathrm{d}s \\ \varphi(x_0 + \xi, t) &= x_0 + \xi + \int_0^t f(\varphi(x_0 + \xi, s))\,\mathrm{d}s \\ u_{x_0}(\xi, t) &= \xi + \int_0^t A_{x_0}(s) \cdot u_{x_0}(\xi, s)\,\mathrm{d}s \\ &= \xi + \int_0^t f'(\varphi(x_0, s)) \cdot u_{x_0}(\xi, s)\,\mathrm{d}s \end{aligned}$$

Then, for any fixed ε, after a sequence of estimations (3 pages in the textbook proof) involving e.g., uniform convergence (section 5) of the first-order remainder term of the Taylor expansion of f, continuity of the flow (theorem 6), and linearity of u (lemma 35) one can prove the following inequality.

$$\frac{\|\varphi(x_0+\xi,t)-\varphi(x_0,t)-u_{x_0}(\xi,t)\|}{\|\xi\|} \leq \varepsilon$$

This shows that $u_{x_0}(\xi,t)$ is indeed a good approximation for the propagation of the initial perturbation ξ and exactly the definition for the space derivative of the flow. ☐

Note that $u_{x_0}(\xi,t)$ yields the space derivative in direction of the vector ξ. The total space derivative of the flow is then the linear function $\xi \mapsto u_{x_0,\xi}(t)$. But this derivative can also be described as the solution of the following "matrix-valued" variational equation:

$$\begin{cases} \dot{W}_{x_0}(t) = A_{x_0}(t) \circ W_{x_0}(t) \\ W_{x_0}(0) = \mathsf{Id} \end{cases} \tag{5}$$

This initial value problem is defined for linear operators of type $\mathbb{R}^n \to_{bl} \mathbb{R}^n$. Thanks to lemma 29, one can show that it is defined on the same existence interval as the flow φ. The solution W_{x_0} is related to solutions of the variational IVP as follows:

$$u_{x_0}(\xi,t) = W_{x_0}(t) \cdot \xi$$

The derivative of the flow φ at (x_0,t) with respect to t is given directly by the ODE, namely $f(\varphi(x_0,t))$. Therefore and according to lemma 24 the total derivative of the flow is characterized as follows:

Theorem 37 (Derivative of the Flow).

$$\mathsf{D}\varphi|_{(x_0,t)} \cdot (\xi,\tau) = W_{x_0}(t) \cdot \xi + \tau \cdot f\ (\varphi(x_0,t))$$

7.5 Continuity of Derivative

Regarding the continuity of the derivative $\mathsf{D}\varphi|_{(x_0,t)} \cdot (\xi,\tau)$ with respect to (x_0,t): $\tau \cdot f\ (\varphi(x_0,t))$ is continuous because of definition 27 and theorem 6.

$W_{x_0}(t)$ is continuous with respect to t, so what remains to be shown is continuity of the space derivative regarding x_0. The proof of this statement relies on theorem 32, because for different values of x_0, W_{x_0} is the solution to ODEs with slightly different right-hand sides. A technical difficulty here is to establish the assumption of *global* Lipschitz continuity for theorem 32.

8 Conclusion

To conclude, our formalization contains essentially all lemmas and proofs of at least 22 pages (Chapter 17) of the textbook by Hirsch *et al.* [2] and additionally required some more general-purpose background to be formalized, in particular uniform limits and the Banach space of (bounded) linear functions. The separate type for bounded linear functions was a minor complication that was necessary because of the type class based library for analysis in Isabelle/HOL. We showed the concrete usability of our results by verifying the connection of the abstract formalization with a concrete rigorous numerical algorithm.

References

1. Haftmann, F.: Code Generation from Specifications in Higher-Order Logic. Dissertation, Technische Universität München, München (2009)
2. Hirsch, M.W., Smale, S., Devaney, R.L.: Differential Equations, Dynamical Systems, and an Introduction to Chaos. Elsevier Academic Print (2013)
3. Hölzl, J., Immler, F., Huffman, B.: Interactive Theorem Proving: 4th International Conference, ITP 2013, Rennes, France, July 22-26, 2013. Proceedings, chap. Type Classes and Filters for Mathematical Analysis in Isabelle/HOL, pp. 279–294. Springer Berlin Heidelberg, Berlin, Heidelberg (2013), `http://dx.doi.org/10.1007/978-3-642-39634-2_21`
4. Huffman, B., Kunčar, O.: Certified Programs and Proofs: Third International Conference, CPP 2013, Melbourne, VIC, Australia, December 11-13, 2013, Proceedings, chap. Lifting and Transfer: A Modular Design for Quotients in Isabelle/HOL, pp. 131–146. Springer International Publishing, Cham (2013), `http://dx.doi.org/10.1007/978-3-319-03545-1_9`
5. Immler, F.: Formally verified computation of enclosures of solutions of ordinary differential equations. In: Badger, J., Rozier, K. (eds.) NASA Formal Methods, LNCS, vol. 8430, pp. 113–127. Springer International Publishing (2014)
6. Immler, F., Hölzl, J.: Numerical analysis of ordinary differential equations in Isabelle/HOL. In: Beringer, L., Felty, A. (eds.) Interactive Theorem Proving, LNCS, vol. 7406, pp. 377–392. Springer (2012)
7. Lelay, C., Melquiond, G.: Différentiabilité et intégrabilité en Coq. Application à la formule de d'Alembert. In: JFLA - Journées Francophone des Langages Applicatifs - 2012. Carnac, France (Feb 2012), `https://hal.inria.fr/hal-00642206`
8. Nipkow, T., Paulson, L.C., Wenzel, M.: Isabelle/HOL: A proof assistant for higher-order logic. LNCS, Springer (2002)
9. Tucker, W.: A rigorous ODE solver and Smale's 14th problem. Foundations of Computational Mathematics 2(1), 53–117 (2002)