

## A DOMAIN-THEORETIC ACCOUNT OF PICARD'S THEOREM

ABBAS EDALAT AND DIRK PATTINSON

*Abstract*

We present a domain-theoretic version of Picard's theorem for solving classical initial value problems in  $\mathbb{R}^n$ . For the case of vector fields that satisfy a Lipschitz condition, we construct an iterative algorithm that gives two sequences of piecewise linear maps with rational coefficients, which converge, from below and above respectively, exponentially fast, to the unique solution of the initial value problem. We provide a detailed analysis of the speed of convergence and the complexity of computing the iterates. The algorithm uses proper data types based on rational arithmetic, where no rounding of real numbers is required. Thus we obtain a sound implementation framework to solve initial value problems. In particular, the use of rational arithmetic guarantees that implementations of our technique will adhere to the bounds on convergence speed and algebraic complexity.

1. *Introduction*

We consider the initial value problem (IVP) given by the system of differential equations

$$\dot{y}_i(x) = v_i(y_1, \dots, y_n), \quad y_i(0) = 0 \quad (i = 1, \dots, n), \quad (1)$$

where the vector field  $v : O \rightarrow \mathbb{R}^n$  is continuous in a neighbourhood  $O \subseteq \mathbb{R}^n$  of the origin, and we look for a differentiable function  $y = (y_1, \dots, y_n) : [-a, a] \rightarrow \mathbb{R}^n$ , defined in a neighbourhood of  $0 \in \mathbb{R}$ , that satisfies (1). By a theorem of Peano there is always a solution [6, p. 19]. Uniqueness of the solution is guaranteed, by Picard's theorem, if  $v$  satisfies a Lipschitz condition. The question of computability and the complexity of the initial value problem has been studied in different contexts in computable analysis [12, 1, 5, 14, 20, 17, 4].

On the algorithmic and more practical side, standard numerical packages for solving IVPs try to compute an approximation to a solution with a specified degree of accuracy. Although these packages are usually robust, their methods are not guaranteed to be correct, and it is easy to find examples where they output inaccurate results [13].

Interval analysis [16] provides a method to give upper and lower bounds for the unique solution in the Lipschitz case with a prescribed tolerance, and has been developed and implemented for analytic vector fields [18, 3]. These approaches are concerned with the correctness of the computed values, and deliver interval values that are guaranteed to contain the true solution of the problem. Typically,

---

Received 16 November 2005, revised 18 September 2006; *published* 23 March 2007.

2000 Mathematics Subject Classification 03F60 (primary), 03B70 (secondary).

© 2007, Abbas Edalat and Dirk Pattinson

implementations of interval analysis techniques represent real numbers as floating-point intervals, and outward rounding is applied if the resulting interval endpoints are not machine-representable.

While this strategy guarantees soundness (that is, containment of the exact result in the computed interval), one has in general no control over the rounding, which can produce unduly large intervals, depending on the accuracy of the underlying floating-point numbers. While it is intuitively clear that more precise floating-point numbers give more accurate results, there is no general guarantee of actual convergence to the solution. For the same reason, one has no control over the speed of convergence.

Domain theory [2] presents an alternative technique, based on proper data types, to produce a provably correct solution with any given degree of accuracy. Using the domain of Scott continuous interval-valued functions on a compact interval, we define a domain-theoretic Picard operator, whose least fixed point contains any solution of the IVP. When the vector field is Lipschitz, the solution is unique, and we construct an iterative algorithm that gives two sequences of piecewise linear maps with rational coefficients, which converge, from below and above respectively, exponentially fast, to the unique solution of the initial value problem. Since the data types for representing the piecewise linear maps with rational coefficients are directly representable on a digital computer, no rounding of real numbers is required. The implementation of the domain-theoretic approach is also complete; that is, we can guarantee the convergence of the approximating iterates to the solution of the IVP also for the implementation. This property is not present in any other approach to validated solutions of differential equations. Furthermore, as a result of the data types that we use, we can give estimates for the speed of convergence of the approximating iterates, which are still valid for an actual implementation of our algorithm.

This simplifies the earlier treatment, discussed in detail in [7], which used a domain for  $C^1$  functions [8, 9]. That approach requires, at each stage of the iteration, a new approximation to the derivative of the solution. The new treatment is much more similar to the classical theorem in that it gives rise, in the Lipschitz case, to fast convergence of the approximations to the solution. As regards the question of computability of the solution of the IVP in the Lipschitz case, the two domain-theoretic techniques lead to the same result as those in computable analysis [12, 20, 19, 21]: if the vector field is computable then the unique solution of the IVP is also computable [7, Corollary 6.3].

We discuss two different bases to represent approximations to the solutions of the IVP, namely the piecewise linear and the piecewise constant functions with rational (or dyadic) coefficients. Using piecewise linear functions, we avoid the computation of rectangular enclosures of the solution, which gives tighter bounds on the solution. This comes at the expense of an increase in the size of the representation of the approximations to the solution. Using the base consisting of piecewise constant functions, we show that the order of the speed of convergence to the solution remains unchanged, while the time and space complexity for the representation of the iterates is much reduced.

Our approach relies on approximating the vector field with a sequence of (interval-valued) step functions which converge exponentially fast to an interval extension of the vector field. We discuss two techniques for obtaining such sequences. First, we

show how to compose two sequences of approximations such that the composition of the approximations still converges exponentially fast. Our second technique is based on a function which computes the values of the vector field to an arbitrary degree of accuracy, and we show how this gives rise to step functions with the desired properties.

A prototypical implementation using the GNU multi-precision library [11] shows that the resulting algorithms are actually feasible in practice, and we plan to refine the implementation and compare it in scope and performance with existing interval analysis packages like AWA [3], bearing in mind that the floating-point arithmetic used by interval software is executed on highly optimised processors, whereas the rational arithmetic needed for our implementation is performed by software.

## 2. Preliminaries and notation

For the remainder of the paper, we fix a continuous vector field

$$v = (v_1, \dots, v_n) : O \longrightarrow \mathbb{R}^n$$

where  $O \subseteq \mathbb{R}^n$  is a subset of  $\mathbb{R}^n$  with  $0 \in O$ , and we consider the IVP given by equation (1). Our aim is to approximate solutions  $y : [-a, a] \rightarrow \mathbb{R}^n$  of the initial value problem (1).

We use basic notions from domain theory; see, for example, [2, 10]. Our work is based on the interval domain  $\mathbb{IR} = \{[a^-, a^+] \mid a^- \leq a^+, a^-, a^+ \in \mathbb{R}\} \cup \{\mathbb{R}\}$ , ordered by reverse inclusion; that is,  $\alpha \sqsubseteq \beta$  if and only if  $\beta \subseteq \alpha$ . We write  $\perp = \mathbb{R}$  for the least element of  $\mathbb{IR}$ . The ‘way below’ relation on  $\mathbb{IR}$  is given by  $\alpha \ll \beta$  if and only if  $\beta \subseteq \alpha^\circ$ , where  $(\cdot)^\circ$  denotes the interior of a set. For  $n \geq 1$ , the domain  $\mathbb{IR}^n$  is isomorphic to the domain of  $n$ -dimensional rectangles  $\{\alpha_1 \times \dots \times \alpha_n \mid \alpha_i \in \mathbb{IR} \text{ for all } 1 \leq i \leq n\}$ , and we do not distinguish between these two presentations. For a rectangle  $A \in \mathbb{IR}^n$ , the subset  $\{S \in \mathbb{IR}^n \mid S \subseteq A\}$  of rectangles contained in  $A$  is a sub-domain of  $\mathbb{IR}^n$ , which is denoted by  $\mathbf{IA}$ .

We consider the  $n$ -dimensional Euclidean space  $\mathbb{R}^n$  equipped with the maximum norm  $\|x\| = \max\{|x_1|, \dots, |x_n|\}$ , as this simplifies dealing with the Lipschitz conditions, which we introduce later.

The powers  $\mathbb{IR}^n$  of the interval domain and the sub-domain  $\mathbf{IA}$ , for a rectangle  $A \in \mathbb{IR}^n$ , are continuous Scott domains. If  $\alpha^-, \alpha^+ \in \mathbb{R}^n$  with  $\alpha_i^- \leq \alpha_i^+$  for all  $1 \leq i \leq n$ , we write  $[\alpha^-, \alpha^+]$  for the rectangle  $[\alpha_1^-, \alpha_1^+] \times \dots \times [\alpha_n^-, \alpha_n^+]$ . Similarly, if  $f : X \rightarrow \mathbb{IR}^n$  is a function, we write  $f = [f^-, f^+]$  if  $f(x) = [f^-(x), f^+(x)]$  for all  $x \in X$ .

The link between the ordinary and the interval-valued function is provided by the notion of *extension*. If  $A \in \mathbb{IR}^n$  is a rectangle, we say that  $g : \mathbf{IA} \rightarrow \mathbb{IR}^n$  is an extension of  $f : A \rightarrow \mathbb{R}^n$  if

$$g(\{x_1\}, \dots, \{x_n\}) = \{f(x_1, \dots, x_n)\}$$

for all  $x \in A$ . Note that every continuous function  $f : A \rightarrow \mathbb{R}^n$  has a canonical maximal extension  $\mathbf{I}f$  defined by  $\mathbf{I}f = (\mathbf{I}f_1, \dots, \mathbf{I}f_n) : \mathbf{IA} \rightarrow \mathbb{IR}^n$  where

$$\mathbf{I}f_i(S) = \begin{cases} f_i(S), & \text{in cases where } f_i(S) \text{ is bounded,} \\ \perp, & \text{otherwise,} \end{cases}$$

for a rectangle  $S \in \mathbf{IA}$  ( $f_i(S)$  denotes direct image). This extension is maximal

in the set of interval-valued functions extending  $f$ . It is easy to see that  $\mathbf{I}f$  is continuous with respect to the Scott topology on  $\mathbf{I}A$  and  $\mathbf{I}\mathbb{R}^n$  if  $f$  is continuous with respect to the Euclidean topology.

If  $P \in \mathbf{I}\mathbb{R}^n$  and  $Q \in \mathbf{I}\mathbb{R}^m$ , we write  $\mathbf{I}P \Rightarrow \mathbf{I}Q$  for the set of continuous functions with respect to the Scott topology on  $\mathbf{I}P$  and the Scott topology on  $\mathbf{I}Q$ ; likewise,  $P \Rightarrow Q$  denotes the set of continuous functions with respect to the Euclidean topology on  $P$  and the Scott topology on  $\mathbf{I}Q$ . The following operations mediate between  $\mathbf{I}P \Rightarrow \mathbf{I}Q$  and  $P \Rightarrow \mathbf{I}Q$ ; this is as in [8].

LEMMA 2.1. *Suppose that  $P \in \mathbf{I}\mathbb{R}^k$  and  $Q \in \mathbf{I}\mathbb{R}^l$ , and consider the following operations:*

$$\begin{aligned} \mathcal{E} : (P \Rightarrow \mathbf{I}R) \ni f &\mapsto \lambda\alpha. \prod_{x \in \alpha} f(x) \in (\mathbf{I}P \Rightarrow \mathbf{I}R); \\ \mathcal{I} : (\mathbf{I}P \Rightarrow \mathbf{I}R) \ni f &\mapsto \lambda x. f(\{x\}) \in (P \Rightarrow \mathbf{I}R). \end{aligned}$$

*Then both  $\mathcal{I}$  and  $\mathcal{E}$  are continuous,  $\mathcal{I} \circ \mathcal{E} = \text{id}$  and  $\text{id} \sqsubseteq \mathcal{E} \circ \mathcal{I}$ .*

For the proof, see [10, II-3.9]. In order to measure the speed of convergence, as well as for technical convenience in the formulation of some of our results, we introduce the following notation.

The *width* of a compact interval  $[a, b]$  is given as  $w([a, b]) = b - a$  and its midpoint is  $m([a, b]) = (a + b)/2$ . We put  $w(\perp) = \infty$ . For  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbf{I}\mathbb{R}^n$  we let  $w(\alpha) = \max\{w(\alpha_i) \mid 1 \leq i \leq n\}$  and  $m(\alpha) = (m(\alpha_1), \dots, m(\alpha_n))$ . If  $X$  is a set and  $f : X \rightarrow \mathbf{I}\mathbb{R}^n$  is a function, the width of  $f$  is given as  $w(f) = \sup\{w(f(x)) \mid x \in X\}$ . In the special case where  $X \subseteq \mathbb{R}$ , we let  $w_\alpha(f) = \sup\{e^{-\alpha|x|}w(f(x)) \mid x \in X\}$ , and we call  $w_\alpha(f)$  the *weighted width* of  $f$  with respect to the weight  $\alpha$ ; this gives  $w(f) = w_0(f)$ . We will use the weighted width to show that the domain-theoretic Picard operator is a contraction.

Given two intervals  $\alpha = [a^-, a^+]$  and  $\beta = [b^-, b^+] \in \mathbf{I}\mathbb{R}$ , their *Hausdorff distance* is  $d(\alpha, \beta) = \max\{|a^+ - b^+|, |a^- - b^-|\}$ . Similarly, for  $\alpha = (\alpha_1, \dots, \alpha_n)$  and  $\beta = (\beta_1, \dots, \beta_n) \in \mathbf{I}\mathbb{R}^n$ , we let  $d(\alpha, \beta) = \max\{d(\alpha_i, \beta_i) \mid 1 \leq i \leq n\}$ , and we define the distance of two functions  $f, g : X \rightarrow \mathbf{I}\mathbb{R}^n$  as  $d(f, g) = \sup\{d(f(x), g(x)) \mid x \in X\}$ .

Considering  $g$  as an approximation to  $f$ , we view the distance  $d(f, g)$  as a measure of the quality of the approximation. We mention two simple lemmas linking distance, width and weighted width.

LEMMA 2.2. *Let  $f : [-a, a] \rightarrow \mathbf{I}\mathbb{R}^n$ . Then  $w_\alpha(f) \leq w(f) \leq e^{\alpha a} w_\alpha(f)$  for all  $\alpha \geq 0$ .*

For the next lemma, recall that  $m(\cdot)$  denotes the midpoint of a rectangle.

LEMMA 2.3. *Suppose that  $\alpha, \beta \in \mathbf{I}\mathbb{R}^n$  are compact. Then*

- (i)  $\|m(\alpha) - m(\beta)\| \leq d(\alpha, \beta)$ ;
- (ii)  $0 \leq w(\beta) - w(\alpha) \leq 2d(\alpha, \beta)$  in the case  $\beta \sqsubseteq \alpha$ .

The proof of both lemmas is a straightforward calculation, and is therefore omitted here.

Finally, if  $x \leq y$  are real numbers, a *partition* of  $[x, y]$  is a finite sequence  $(q_0, \dots, q_k)$  of real numbers such that  $x = q_0 < \dots < q_k = y$ , and the set of partitions of  $[x, y]$  is denoted by  $\mathcal{P}[x, y]$ . The *norm* of a partition  $Q = (q_0, \dots, q_k)$  is denoted by  $|Q| = \max\{q_i - q_{i-1} \mid 1 \leq i \leq k\}$ .

3. *Picard operator in domain theory*

In the classical proof of Picard's theorem on the existence and uniqueness of the solution of the initial value problem (1), one defines an integral operator on  $C^0[-a, a]$  by

$$y \mapsto \lambda x. \int_0^x v(y(t)) dt$$

(with the integral understood componentwise), which can be shown to be a contraction for sufficiently small  $a$  provided that  $v$  satisfies a Lipschitz condition [15]. An application of Banach's theorem then yields a solution of the initial value problem. We now define the domain-theoretic Picard operator for arbitrary Scott continuous vector fields  $u : \mathbf{IA} \rightarrow \mathbf{IB}$ , for  $A, B \in \mathbf{IR}^n$ , and focus on the special case where  $u$  is an extension of a classical function later. As in the classical proof, the Picard operator is an integral operator, and we therefore introduce the integral of interval-valued functions.

DEFINITION 3.1. Suppose that  $f = [f^-, f^+] : [-a, a] \rightarrow \mathbf{IR}$  is Scott continuous. For  $x \in [-a, a]$  we let

$$\int_0^x f(t) dt = \left[ \int_0^x f^{-\sigma(x)}(t) dt, \int_0^x f^{\sigma(x)}(t) dt \right],$$

where  $\sigma(x) = \text{sgn}(x)$  is the sign of  $x$  and  $f^0(t) = 1$ . If  $f = (f_1, \dots, f_n) : [-a, a] \rightarrow \mathbf{IR}^n$ , we let

$$\int_0^x f(t) dt = \left( \int_0^x f_1(t) dt, \dots, \int_0^x f_n(t) dt \right).$$

Note that if we integrate in the positive  $x$ -direction, then  $f^-$  contributes to the lower function associated with the integral of  $f$ , and  $f^+$  contributes to the upper function. If we integrate in the negative  $x$ -direction, the roles of  $f^-$  and  $f^+$  are swapped to ensure that the lower value of the integral is indeed smaller than the upper value. The following lemma shows that our definition is meaningful.

LEMMA 3.2. *Suppose that  $f : [-a, a] \rightarrow \mathbf{IR}$  is Scott continuous. Then*

- (i)  $f^-$  and  $f^+$  are measurable;
- (ii)  $\int_0^x f(t) dt \in \mathbf{IR}$  for all  $x \in [-a, a]$ .

*Proof.* For Scott continuous  $f$ , the functions  $f^-$  and  $f^+$  are, respectively, lower and upper semi-continuous, and hence measurable. If  $\sigma(x) = \text{sgn}(x)$ , then  $\sigma(x)f^{-\sigma(x)} \leq \sigma(x)f^{\sigma(x)}$ , and  $\int_0^x f^{-\sigma(x)}(t) dt \leq \int_0^x f^{\sigma(x)}(t) dt$  follows from the definition of the ordinary integral. Finally, we have to show that  $\int_0^x f^+(t) dt = \infty$  if and only if  $\int_0^x f^-(t) dt = -\infty$ , but this is clear as  $f^+(t) = \infty$  if and only if  $f^-(t) = -\infty$ .  $\square$

The following lemma shows that integration is compatible with taking suprema.

LEMMA 3.3. *Let  $f : [-a, a] \rightarrow \mathbf{IR}^n$ .*

- (i) *The function  $\lambda x. \int_0^x f(t) dt$  is Scott continuous.*
- (ii) *The function  $\int : ([-a, a] \Rightarrow \mathbf{IR}^n) \rightarrow ([-a, a] \Rightarrow \mathbf{IR}^n)$ , defined by  $f \mapsto \lambda x. \int_0^x f(t) dt$ , is Scott continuous.*

*Proof.* We assume that  $n = 1$ , from which the general case follows. If  $g(x) = \int_0^x f(t) dt$ , then  $g^-$  and  $g^+$  are continuous, and hence  $g$  is Scott continuous. The second statement follows from the monotone convergence theorem.  $\square$

In the following, we are interested in solutions  $y : [-a, a] \rightarrow \mathbb{R}^n$  of the initial value problem (1), and we fix the domain of definition  $[-a, a]$  of a solution for the remainder of the paper. In order to define the domain-theoretic Picard operator, we assume that  $u : \mathbf{IA} \rightarrow \mathbf{IB}$ , where either

- $A = [-K, K]^n$ ,  $B = [-M, M]^n$  and  $a \geq 0$  satisfies  $aM \leq K$ , or
- $A = B = \mathbb{R}^n$  and  $a \geq 0$  is arbitrary.

The restriction on the lifetime  $a$  of a solution in the first case is due to the fact that, for a solution  $z : [-a, a] \rightarrow \mathbb{R}^n$  of the IVP (1), we have  $\dot{z} = v(z) \leq M$ ; that is,  $M$  is a bound on the derivative of  $z$ . As  $z(0) = 0$ , we can only guarantee that  $z(t) \leq Mt$ , which gives rise to the restriction  $aM \leq K$  for the expression  $v(z(t))$  to be well defined for all  $t \in [-a, a]$ . Clearly, this restriction is not necessary if  $u$  is defined on all of  $\mathbf{IR}^n$ .

**DEFINITION 3.4.** Suppose that  $u \in \mathbf{IA} \Rightarrow \mathbf{IB}$ . The *domain-theoretic Picard operator*  $P_u : ([-a, a] \Rightarrow \mathbf{IA}) \rightarrow ([-a, a] \Rightarrow \mathbf{IA})$  is defined by  $P_u(y) = \lambda x. \int_0^x u(y(t)) dt$ .

**LEMMA 3.5.**  $P_u$  is well defined and continuous.

*Proof.* In case  $A = [-K, K]^n$  and  $B = [-M, M]^n$ , the lemma follows from the assumption that  $aM \leq K$  that  $P_u(y) \in ([0, a] \Rightarrow \mathbf{IA})$  whenever  $y \in [0, a] \Rightarrow \mathbf{IA}$ . Lemma 3.3 shows that  $P_u(y)$ , for  $y \in [0, a] \Rightarrow \mathbf{IA}$ , and  $P_u$  itself are continuous.  $\square$

In the classical proof of Picard's theorem, one constructs solutions of IVPs as fixpoints of the (classical) Picard operator. The domain-theoretic proof replaces Banach's theorem with the domain-theoretic fixpoint theorem in the construction of a fixed point of the (domain-theoretic) Picard operator. Unlike the classical case, where one chooses an arbitrary initial approximation, we need to choose an initial (interval-valued) function  $y_0$  which is invariant under the Picard operator; that is,  $y_0 \sqsubseteq P_u(y_0)$ .

**THEOREM 3.6.** Suppose that  $y_0 : [-a, a] \rightarrow \mathbf{IA}$  satisfies  $y_0 \sqsubseteq P_u(y_0)$ , and let  $y_{k+1} = P_u(y_k)$ . Then  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  satisfies  $P_u(y) = y$ .

*Proof.* The theorem follows immediately from the domain-theoretic fixpoint theorem; see, for example, [2, Theorem 2.1.19], applied to the directed complete partial order  $\uparrow y_0 = \{f : [-a, a] \rightarrow \mathbf{IA} \mid y_0 \sqsubseteq f\}$ .  $\square$

For computing solutions of IVPs, we will take  $y_0 = \lambda t. [-K, K]^n$  in the case that  $u : \mathbf{I}[-K, K]^n \rightarrow \mathbf{I}[-M, M]^n$  is defined in a bounded neighbourhood of the origin, and the restriction  $aM \leq K$  will ensure that  $y_0 \sqsubseteq P_u(y_0)$ . This situation is discussed in detail in Section 4. In the unbounded case, we need to take  $y_0$  such that the (unique) solution  $z$  of problem (1) satisfies  $y_0 \sqsubseteq z$ ; see Section 5 for details.

The bridge between the solution of the domain-theoretic fixpoint equation and the classical initial value problem is established in the following proposition, where  $\mathbf{S}f : [-a, a] \rightarrow \mathbf{I}[-K, K]^n$  denotes the function  $\lambda x. \{f(x)\}$ , for  $f : [-a, a] \rightarrow [-K, K]^n$ .

PROPOSITION 3.7. *Suppose that  $y_0 : [-a, a] \rightarrow \mathbf{IA}$  satisfies  $y_0 \sqsubseteq P_u(y_0)$ , and that  $y \in \uparrow y_0$  is the least fixpoint of the restricted Picard operator  $P_u : (\uparrow y_0) \rightarrow (\uparrow y_0)$ .*

- (i) *If  $f : [-a, a] \rightarrow \mathbf{IA}$  solves (1) and  $y_0 \sqsubseteq \mathbf{S}f$ , then  $y \sqsubseteq \mathbf{S}f$ .*
- (ii) *If  $y$  has width 0, then  $y^- = y^+$  solves (1).*

*Proof.* For the first statement, note that  $\mathbf{S}f$  is a fixed point of  $P_u$  and  $y$  is the least such. The second statement follows from the fundamental theorem of calculus; note that  $y^- = y^+$  implies continuity from the left and from the right.  $\square$

The previous proposition can be read as a soundness result. Assuming that the problem (1) admits a unique solution  $z$  with  $y_0 \sqsubseteq z$ , then the least fixpoint  $y$  of the domain-theoretic Picard operator will satisfy  $y \sqsubseteq z$ .

#### 4. Picard operator for Lipschitz vector fields

We now specialise our discussion to the case where  $v : [-K, K]^n \rightarrow [-M, M]^n$  is defined in a compact neighbourhood of the origin, and we fix an interval extension  $u : \mathbf{I}[-K, K]^n \rightarrow \mathbf{I}[-M, M]^n$  of  $v$ ; the unbounded case will be discussed in Section 5.

This allows us to consider the following spaces for approximating the vector field and the solutions to the IVP, where  $a \geq 0$  is with  $aM \leq K$ :

- $\mathcal{V} = \mathbf{I}[-K, K]^n \Rightarrow \mathbf{I}[-M, M]^n$  for approximations of the vector field;
- $\mathcal{S} = [-a, a] \Rightarrow \mathbf{I}[-K, K]^n$  for approximate solutions.

It is clear that in this setup, with  $aM \leq K$ , the function  $y_0 = \lambda t. [-K, K]^n : [-a, a] \rightarrow \mathbf{I}[-K, K]^n$  satisfies  $y_0 \sqsubseteq P_u(y_0)$ .

We can ensure the uniqueness of the solution of the IVP by requiring that the vector field satisfy an interval version of the Lipschitz property. Recall that for metric spaces  $(M, d)$  and  $(M', d')$ , a function  $f : M \rightarrow M'$  is Lipschitz, if there is  $L \geq 0$  such that  $d'(f(x), f(z)) \leq L \cdot d(x, z)$  for all  $x, z \in M$ . The following definition translates this property into an interval setting; see also [16].

DEFINITION 4.1 (LIPSCHITZ CONDITION). Suppose that  $u : \mathbf{I}[-K, K]^n \rightarrow \mathbf{I}[-M, M]^n$ . Then  $u$  is interval Lipschitz if there is some  $L \geq 0$  such that  $w(u(\alpha)) \leq L \cdot w(\alpha)$  for all  $\alpha \in \mathbf{I}[-K, K]^n$ . In this case,  $L$  is called an *interval Lipschitz constant* for  $u$ .

The following proposition describes the relationship between the classical notion and its interval version.

PROPOSITION 4.2. *For  $v : [-K, K]^n \rightarrow [-M, M]^n$ , the following are equivalent.*

- (i)  *$v$  is Lipschitz.*
- (ii) *The canonical extension of  $v$  satisfies an interval Lipschitz condition.*
- (iii)  *$v$  has an interval Lipschitz extension.*

*Proof.* If  $v$  is Lipschitz, then the canonical extension of  $v$  satisfies an interval Lipschitz condition. Now assume that  $u$  is an extension of  $v$  which is interval Lipschitz, and let  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in [-K, K]^n$ . Let  $R(x, y)$  denote the rectangle  $\alpha_1 \times \dots \times \alpha_n$  where  $\alpha_i = [x_i, y_i]$  in case  $x_i \leq y_i$ , and  $\alpha_i = [y_i, x_i]$

otherwise. Since  $u$  extends  $v$ , we have  $v(x) \in u(\{x\}) \subseteq u(R(x, y))$  by monotonicity. Hence  $u_i^-(R(x, y)) \leq v_i(x) \leq u_i^+(R(x, y))$  for all  $1 \leq i \leq n$ . Similarly,  $u_i^-(R(x, y)) \leq v_i(y) \leq u_i^+(R(x, y))$ . Now

$$\begin{aligned} \|v(x) - v(y)\| &= \max_{1 \leq i \leq n} \|v_i(x) - v_i(y)\| \\ &\leq \max_{1 \leq i \leq n} u_i^+(R(x, y)) - u_i^-(R(x, y)) \\ &= \max_{1 \leq i \leq n} w(u_i(R(x, y))) \\ &\leq L \cdot w(R(x, y)) \\ &= L\|x - y\|, \end{aligned}$$

as required. □

Note that every interval Lipschitz function induces a total and continuous classical function.

**COROLLARY 4.3.** *Suppose that  $u$  is interval Lipschitz. Then  $w(u(\alpha)) = 0$  whenever  $w(\alpha) = 0$ , and the induced real-valued function  $\bar{u}$ , given by  $\bar{u}(x) = z$  if and only if  $u(\{x\}) = \{z\}$ , is continuous.*

We now show that the least fixpoint of the Picard operator associated with  $u$  has indeed width zero, if  $u$  satisfies an interval Lipschitz condition. We assume for the remainder of the paper that  $u$  is an extension of  $v$  that satisfies an interval Lipschitz condition with Lipschitz constant  $L$ . In order to show that the least fixpoint of  $P_u$  has width zero, we use the weighted width, introduced in Section 2. The following lemma is the essential step in showing that the least fixpoint of the domain-theoretic Picard operator actually has width 0 on the whole of  $[-a, a]$ . This is most conveniently expressed as a statement about the whole space  $\mathcal{S}$  of solutions introduced at the beginning of Section 4.

**LEMMA 4.4.** *Let  $y \in \mathcal{S}$ . Then  $w_\alpha(P_u(y)) \leq (L/\alpha)w_\alpha(y)$ .*

*Proof.* For the  $i$ th component  $P_u(y)_i$  of  $P_u(y)$ , we calculate that

$$\begin{aligned} w_\alpha(P_u(y)) &= \sup_{t \in [-a, a]} e^{-\alpha|t|} \int_0^t u_i^{\sigma(t)}(y(x)) - u_i^{-\sigma(t)}(y(x)) dx \\ &\leq \sup_{t \in [-a, a]} e^{-\alpha|t|} \int_0^t \sigma(t) \cdot L \cdot e^{-\alpha|x|} e^{\alpha|x|} \cdot w(y(x)) dx \\ &\leq L \cdot w_\alpha(y) \cdot \sup_{t \in [-a, a]} e^{-\alpha|t|} \int_0^{|t|} e^{\alpha x} dx \\ &\leq \frac{L}{\alpha} w_\alpha(y) \end{aligned}$$

as required. □

Recall that  $y_0(x) = [-K, K]^n$ , and hence  $w(y_0) = w_\alpha(y_0) = 2K$  for all  $\alpha \geq 0$ . This gives us the following theorem for the (not weighted) width of the iterates  $y_k$ .

**THEOREM 4.5.** *Let  $y_{k+1} = P_u(y_k)$  for all  $k \in \mathbb{N}$ . Then  $w(y_k) \in \mathcal{O}(2^{-k})$ . In particular,  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  is real-valued and solves (1).*

*Proof.* Put  $\alpha = 2L$  in Lemma 4.4. Then, by induction, one obtains  $w_{2L}(y_k) \leq 2^{-k}w_{2L}(y_0)$ , and hence  $w(y_k) \leq 2^{-k}e^{2L}w(y_0)$  by Lemma 2.2. It follows from Proposition 3.7 that  $y$  solves the initial value problem.  $\square$

The last result is a simplification, and at the same time a generalisation to higher dimensions, of [9, Theorem 7.2] and [7, Theorem 6.1], which give a domain-theoretic proof that the unique solution of the initial value problem for a time-dependent, computable scalar field is computable.

Although the above theorem tells us that the iterates  $y_k$  of the Picard operator will converge to the unique solution, we need to take a further step to actually be able to compute the iterates. In particular, we need a way to effectively compute the integrals involved in the definition of  $P_u$ . For this reason, we now consider approximations to  $u$ ; the basic idea is that every continuous vector field can be approximated by a sequence of step functions (that is, functions taking only finitely many values), which allows us to compute the integrals involved in calculating the approximations to the solution effectively (they reduce to a finite sum). The key property that enables us also to use approximations to the vector field is the continuity of the mapping  $u \mapsto P_u$ , which is similar to the approach of [7, Section 6].

LEMMA 4.6. *The map  $P : \mathcal{V} \rightarrow (\mathcal{S} \Rightarrow \mathcal{S})$ ,  $u \mapsto P_u$ , is continuous.*

*Proof.* This follows from the continuity of  $u$  and the monotone convergence theorem.  $\square$

This continuity property allows us to compute solutions to the classical initial value problem by means of a converging sequence of approximations of  $u$ .

PROPOSITION 4.7. *Suppose that  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  and  $y_{k+1} = P_{u_k}(y_k)$  for  $k \in \mathbb{N}$ . Then  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  satisfies  $y = P_u(y)$ .*

*Proof.* The proof follows from Theorem 3.6 and the continuity of  $u \mapsto P_u$  by the interchange-of-suprema law (see, for example, [2, Proposition 2.1.12]).  $\square$

We have seen that the Lipschitz condition on the vector field ensures that the approximations of the solution converge exponentially fast (Theorem 4.5). In the presence of approximations of the vector field, the speed of convergence will also depend on how fast the vector field is approximated. The following estimate allows us to describe the speed of convergence of the iterates if the vector field is approximated by an increasing chain of vector fields.

LEMMA 4.8. *Let  $r \sqsubseteq u$  and  $y \in \mathcal{S}$ . Then  $w_\alpha(P_r(y)) \leq (L/\alpha)w_\alpha(y) + (2/\alpha\epsilon)d(u, r)$ .*

*Proof.* For the  $i$ th component  $P_r(y)_i$  we calculate, using Lemma 4.4, that

$$\begin{aligned} w_\alpha(P_r(y))_i &= \sup_{t \in [-a, a]} e^{-\alpha|t|} \int_0^t r_i^{\sigma(t)}(y(x)) - r_i^{-\sigma(t)}(y(x)) dx \\ &\leq \sup_{t \in [-a, a]} e^{-\alpha|t|} \int_0^t u_i^{\sigma(t)}(y(x)) - u_i^{-\sigma(t)}(y(x)) + 2d(u, r) dx \\ &\leq \sup_{t \in [-a, a]} e^{-\alpha|t|} w(P_u(y))_i + \sup_{t \in [-a, a]} e^{-\alpha|t|} \cdot |t| \cdot 2d(u, r) \\ &\leq \frac{L}{\alpha} w_\alpha(y) + \frac{2}{\alpha\epsilon} d(u, r) \end{aligned}$$

where the estimate for the second term follows from  $f'(1/\alpha) = 0$ ,  $f''(1/\alpha) < 0$  for  $f(x) = x \cdot e^{-\alpha x}$ .  $\square$

Using this estimate, we can now prove fast convergence if the approximations of the vector field converge fast, too.

**THEOREM 4.9.** *Suppose that  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  with  $d(u, u_k) \in \mathcal{O}(2^{-k})$ . For  $k \geq 0$ , put  $y_{k+1} = P_{u_k}(y_k)$  and  $y = \bigsqcup_{k \in \mathbb{N}} y_k$ . Then  $w(y_k) \in \mathcal{O}(2^{-k})$  and  $y$  is real-valued and solves equation (1).*

*Proof.* We show that  $w_{4L}(y_k) \leq 2^{-k} \cdot w_{4L}(y_0)$  by induction on  $k$ ; this will imply that  $w(y_k) \in \mathcal{O}(2^{-k})$ . Without loss of generality, we assume that  $d(u, u_k) \leq 2^{-k-1} L e \cdot w(y_0)$ . There is nothing to show in the case  $k = 0$ . For the general case, we invoke Lemma 4.8 and obtain

$$\begin{aligned} w_{4L}(y_{k+1}) &\leq \frac{1}{4} w_{4L}(y_k) + \frac{2}{4Le} d(u, u_k) \\ &\leq \frac{1}{4} 2^{-k} w_{4L}(y_0) + \frac{1}{4} 2^{-k} w_{4L}(y_0) \\ &= 2^{-(k+1)} w_{4L}(y_0), \end{aligned}$$

as required. In combination with Lemma 2.2 this yields  $w(y_k) \leq 2^{-k} \cdot e^{4aL} w(y_0) \in \mathcal{O}(2^{-k})$ .  $\square$

Given a representation of  $u$  in terms of step functions, Proposition 4.7 gives rise to an algorithm for computing the solution of the initial value problem, and Theorem 4.9 provides an estimate on the speed of convergence.

We conclude with a brief remark on the maximal lifetime of solutions before we extend the method to vector fields defined on the whole of  $\mathbb{R}^n$ .

**REMARK 4.10.** Recall that we have assumed throughout the section that  $aM \leq K$ , where  $M$  is a bound on the absolute value of the vector field on  $[-K, K]^n$ . If we drop this restriction and use a modified Picard operator  $P'_u$  defined by  $P'_u(y) = \lambda t \cdot P_u(y)(t) \cap [-K, K]^n$ , the ensuing iterates  $y_k$  will still converge to a solution of the problem, provided that  $y_k(t) \in (-K, K)^n$  for all  $t \in [-a, a]$ .

In the next section, the method is extended to deal with vector fields defined on the whole of  $\mathbb{R}^n$ .

### 5. Picard's method for unbounded vector fields

In the previous section, we have shown how to construct domain-theoretic solutions of initial value problems given by a vector field defined on a rectangle  $[-K, K]$  containing the origin.

In practice, one often encounters the situation where  $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is defined on the whole of the  $n$ -dimensional Euclidean space, which renders the limitation of  $v$  being defined on some hyper-rectangle  $[-K, K]^n$  extremely restrictive: For the equation to be well-defined, one has to impose the restriction  $aM \leq K$  which poses an upper limit to the lifetime  $a$  of any solution. The next example illustrates this situation.

**EXAMPLE 5.1.** Consider the IVP  $\dot{y} = y + 1$  with initial condition  $y(0) = 0$ . This problem has the solution  $y(t) = e^t - 1$ , which is defined on the whole real line.

However, the requirement  $aM \leq K$  forces us to consider the vector field as being of type  $v : [-K, K] \rightarrow [-(K + 1), K + 1]$  (that is,  $M = K + 1$ ) and subsequently  $a \leq (K + 1)/K$ , which restricts the domain of definition of the constructed solution to the interval  $[-a, a]$  of width at most 2.

One situation where the global existence of solutions to IVPs is particularly important is the solution of linear boundary-value problems — that is, differential equations of the form

$$\dot{y} = Ay + g \quad \text{with boundary conditions involving } y(a) \text{ and } y(b),$$

where  $A$  is a (possibly time-dependent)  $n \times n$ -matrix. Clearly, we need to construct solutions in this case at least in the interval  $[a, b]$ .

In this section, we extend the construction of domain-theoretic solutions to vector fields defined over the whole of  $n$ -dimensional Euclidean space. Classically, this is achieved by constructing solutions locally, and then applying the extension theorem [6]. In the domain-theoretic setting, this is captured by amalgamating the local fixpoint iterations into one, albeit on a larger function space, so that a single iterative scheme yields the solution on the whole time interval.

For the remainder of the section, we fix a vector field  $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$  that satisfies the Lipschitz condition with Lipschitz constant  $L$ , and an extension  $u : \mathbb{IR}^n \rightarrow \mathbb{IR}^n$  that is interval Lipschitz with the same constant  $L$ . The problem that we face is the following: if we assume that the range of the classical vector field  $v : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is unbounded, we cannot start the iteration with the everywhere undefined function  $y_0 = \lambda x$ .  $\perp$ , since  $y_0$  is already a fixpoint of the Picard operator.

EXAMPLE 5.2. Suppose that  $v : \mathbb{R} \rightarrow \mathbb{R}$  is the identity function  $v(x) = x$  with extension  $u(\alpha) = \alpha$  for  $\alpha \in \mathbb{IR}$ . Then the function  $y = \lambda x$ .  $\perp$  is the least fixed point of  $P_u$ :

$$P_u(y)(t) = \int_0^t u(y(x)) \, dx = \int_0^t u(\perp) \, dx = \int_0^t \perp \, dx = \perp .$$

Note that the corresponding IVP  $\dot{y} = v(y)$ ,  $y(0) = 0$  has the unique solution  $y(t) = 0$ .

Therefore, a more sophisticated technique is called for. We now introduce local a priori bounds for solutions of IVPs. The idea is to fix a partition (cf. Section 2)  $Q = (q_0, \dots, q_k)$  of the interval  $[0, a]$  on whose symmetric expansion  $[-a, a]$  about 0 we want to construct a solution of the equation. We then define constants  $K_i$  such that the (unique) solution  $z : [-a, a] \rightarrow \mathbb{R}^n$  satisfies  $\|z(t)\| \leq K_i$  on every interval  $[-q_i, q_i]$  induced by the partition. This allows us to take  $y_0$  to be the least function satisfying  $[-K_i, K_i] \sqsubseteq y(t)$  for all  $t \in [-q_i, q_i]$  and all  $i = 0, \dots, k$  as starting-point of the iteration or, equivalently, to obtain a meaningful least fixpoint of the Picard operator in the space  $\uparrow y_0 = \{f : [-a, a] \rightarrow \mathbb{IR}^n \mid y_0 \sqsubseteq f\}$ .

DEFINITION 5.3. Suppose that  $Q = (q_0, \dots, q_k) \in \mathcal{P}[0, a]$  with  $|Q| < 1/2L$ . Define the constants

$$K_i^Q = \frac{q_i \|v(0)\|}{(1 - 2L|Q|)^i} \quad \text{for all } i = 0, \dots, k.$$

We drop the superscript  $Q$  if the partition  $Q$  is clear from the context. The constants  $K_i^Q$  are called the *local a priori bounds* induced by the partition  $Q$ , and we define the induced *global bound* by  $K^Q = K_k^Q$ .

We collect some straightforward properties, which will be used later.

LEMMA 5.4. *Suppose that  $Q = (q_0, \dots, q_k) \in \mathcal{P}[0, a]$  with  $|Q| < 1/2L$ .*

- (i)  $K_i^Q \geq K_{i-1}^Q + (q_i - q_{i-1})\|v(0)\| + 2L|Q|K_i^Q$  for all  $1 \leq i \leq k$ .
- (ii)  $K_{i-1}^Q \leq K_i^Q$  for all  $1 \leq i \leq k$ .

*Proof.* Throughout the proof, we drop the superscript  $Q$ . For the first item, we fix  $1 \leq i \leq n$  and calculate

$$\begin{aligned} K_i &= \frac{q_{i-1}\|v(0)\| + (q_i - q_{i-1})\|v(0)\|}{(1 - 2L|Q|)^i} \\ &\geq \frac{q_{i-1}\|v(0)\| + (1 - 2L|Q|)^{i-1}(q_i - q_{i-1})\|v(0)\|}{(1 - 2L|Q|)^i} \\ &= \frac{K_{i-1}}{1 - 2L|Q|} + \frac{(q_i - q_{i-1})\|v(0)\|}{1 - 2L|Q|}. \end{aligned}$$

Hence

$$K_i(1 - 2L|Q|) \geq K_{i-1} + (q_i - q_{i-1})\|v(0)\|;$$

that is,

$$K_i \geq K_{i-1} + (q_i - q_{i-1})\|v(0)\| + 2L|Q|K_i.$$

For the second claim, note that  $0 < |Q| < 1/2L$ ; hence  $1/(1 - 2L|Q|) > 1$ , and therefore

$$\begin{aligned} K_{i-1} &= \frac{q_{i-1}\|v(0)\|}{(1 - 2L|Q|)^{i-1}} \\ &\leq \frac{q_i\|v(0)\|}{(1 - 2L|Q|)^i} \\ &= K_i. \end{aligned} \quad \square$$

The following proposition justifies our choice of terminology. Note that we assume that  $v$  satisfies the Lipschitz condition with Lipschitz constant  $L$ .

PROPOSITION 5.5. *Suppose that  $Q = (q_0, \dots, q_k) \in \mathcal{P}[0, a]$  with  $|Q| \leq 1/2L$  and  $z : [-a, a] \rightarrow \mathbb{R}^n$  is the unique solution of the IVP (1). Then  $\|z(t)\| \leq K_i$  for all  $t \in [-q_i, q_i]$ .*

*Proof.* We show that  $\|z(t)\| \leq K_i$  for all  $t \in [-q_i, q_i]$  by induction on  $i$ . For  $i = 0$ , there is nothing to show. Now suppose that  $0 < i$ . If  $t \in [-q_{i-1}, q_{i-1}]$ , the result follows from the induction hypothesis in conjunction with Lemma 5.4. Let  $t \in [q_{i-1}, q_i]$ .

As  $z(t) = \int_0^t v(z(x))dx$ , we have the following estimate:

$$\begin{aligned}
 \|z(t)\| &\leq \left\| \int_0^{q_{i-1}} v(z(x)) dx \right\| + \left\| \int_{q_{i-1}}^t v(z(x)) dx \right\| \\
 &\leq \|z(q_{i-1})\| + \int_{q_{i-1}}^{q_i} \|v(z(x))\| dx \\
 &\leq \|z(q_{i-1})\| + \int_{q_{i-1}}^{q_i} \|v(z(x)) - v(z(0)) + v(z(0))\| dx \\
 &\leq \|z(q_{i-1})\| + \int_{q_{i-1}}^{q_i} L\|z(x) - z(0)\| + \|v(z(0))\| dx \\
 &\leq \|z(q_{i-1})\| + \int_{q_{i-1}}^{q_i} L\|z(x)\| + \|v(0)\| dx \\
 &\leq \|z(q_{i-1})\| + (q_i - q_{i-1})L \sup_{x \in [q_{i-1}, q_i]} \|z(x)\| + (q_i - q_{i-1})\|v(0)\|.
 \end{aligned}$$

Hence, as  $q_i - q_{i-1} \leq |Q| \leq 2|Q|$ , we find that

$$\sup_{x \in [q_{i-1}, q_i]} \|z(x)\| \leq \|z(q_{i-1})\| + 2|Q|L \sup_{x \in [q_{i-1}, q_i]} \|z(x)\| + (q_i - q_{i-1})\|v(0)\|.$$

By rearranging and using the induction hypothesis  $\|z(q_{i-1})\| \leq K_{i-1}$ , we see that

$$\sup_{x \in [q_{i-1}, q_i]} \|z(x)\|(1 - 2|Q|L) \leq K_{i-1} + (q_i - q_{i-1})\|v(0)\|,$$

and the result follows from Lemma 5.4.

Similarly, the claim is established for  $t \in [-q_i, -q_{i-1}]$ .  $\square$

Actually, one can prove the same statement with a sharper definition of  $K_i$  and show that

$$\|z(x)\| \leq \frac{q_i \|v(0)\|}{(1 - |Q|L)^i} \quad \text{on } [0, q_i].$$

However, as we shall see later, we need the a priori bounds of Definition 5.3 when we move to interval-valued functions.

For later reference, we include the following lemma, which will be used to show that the Picard operator, which we introduce in the next section, is well defined.

**LEMMA 5.6.** *Suppose that  $Q = (q_0, \dots, q_k) \in \mathcal{P}[0, a]$  with  $|Q| < 1/2L$ . Then*

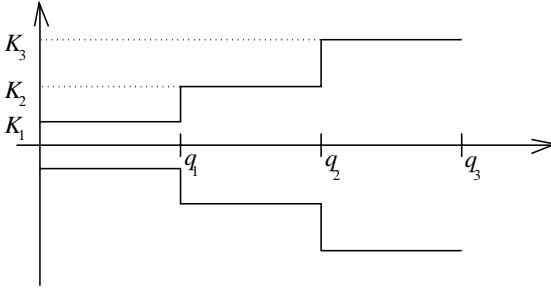
$$q_i \|v(0)\| + \sum_{j=1}^i 2LK_j |Q| \leq K_i$$

for all  $i = 0, \dots, k$ .

*Proof.* We proceed by induction on  $i$ , where there is nothing to show for the case  $i = 0$  (recall that  $K_0 = q_0 = 0$ ). Now suppose that  $i > 0$ . The induction hypothesis, together with Lemma 5.4, gives

$$\begin{aligned}
 q_i \|v(0)\| + \sum_{j=1}^i 2LK_j |Q| &= q_{i-1} \|v(0)\| + \sum_{j=1}^{i-1} 2LK_j |Q| + (q_i - q_{i-1}) \|v(0)\| + 2LK_i |Q| \\
 &\leq K_{i-1} + (q_i - q_{i-1}) \|v(0)\| + 2LK_i |Q| \leq K_i
 \end{aligned}$$

as required.  $\square$


 Figure 1: The set  $\mathcal{S}_Q$  for  $Q = (q_0, q_1, q_2, q_3)$ 

We now introduce the space on which we construct domain-theoretic solutions.

DEFINITION 5.7. Suppose that  $Q = (q_0, \dots, q_k)$  is a partition of  $[0, a]$  with  $|Q| < 1/2L$ , and take the a priori bounds  $K_i$  and the global bound  $K^Q$  as in Definition 5.3. We let

$$\mathcal{S}_Q = \{f : [-a, a] \rightarrow \mathbf{I}[-K^Q, K^Q]^n \mid f \upharpoonright (-q_i, q_i) \sqsupseteq \lambda t.[-K_i, K_i] \text{ for all } 1 \leq i \leq k\}$$

and we write

$$y_0^Q = \bigsqcup_{1 \leq i \leq k} [-q_i, q_i] \searrow [-K_i, K_i]$$

for the least element of  $\mathcal{S}_Q$ , where the step function  $\alpha \searrow \beta$  is defined by  $x \mapsto \beta$  if and only if  $x \in \alpha^o$ , and  $x \mapsto \perp$  otherwise. We call  $\mathcal{S}_Q$  the *solution space associated with  $Q$* , and we drop the sub/superscript  $Q$  if the partition is clear from the context.

Graphically, the set  $\mathcal{S}_Q$  is the set of functions whose interval values are bounded by a double staircase, illustrated in Figure 1. Using Lemma 5.6, we can now show that the Picard operator maps  $\mathcal{S}_Q$  to  $\mathcal{S}_Q$ .

LEMMA 5.8. *Let  $Q \in \mathcal{P}[0, a]$  with  $|Q| < 1/2L$ . Then  $P_u(y) \in \mathcal{S}_Q$  if  $y \in \mathcal{S}_Q$ .*

*Proof.* Suppose that  $Q = (q_0, \dots, q_k)$ . By monotonicity of  $P_u$  it suffices to show the statement for  $y = y_0$ , as  $y_0$  is the least element of  $\mathcal{S}$ .

We show, by induction on  $i$ , that  $[-K_i, K_i]^n \sqsubseteq P_u(y)(t)$  for all  $t \in [-q_i, q_i]$ . For  $i = 0$  there is nothing to show, so we suppose that  $i > 0$ .

As  $y_0 \sqsubseteq \lambda t.0$ , we see, by monotonicity of integration, that  $P_u(y_0) \sqsubseteq P_u(\lambda t.0)$ , and hence

$$\begin{aligned} P_u(y_0) &= \int_0^t u(y_0(x)) dx \\ &\sqsubseteq \int_0^t u(0) dx = \left\{ \int_0^t v(0) dt \right\} = \{t \cdot v(0)\}; \end{aligned}$$

that is,

$$t \cdot v(0) \in P_u(y_0) \tag{2}$$

for all  $t \in [0, a]$ .

Moreover, the width of  $P_u(y_0)$  can be estimated using the fact that  $u$  is interval Lipschitz as follows for  $t \in [q_{i-1}, q_i]$ :

$$\begin{aligned} w(P_u(y_0)(t)) &= \int_0^t w(u(y_0(x))) dx \\ &= \sum_{j=1}^{i-1} \int_{q_{j-1}}^{q_j} w(u(y_0(x))) dx + \int_{q_{i-1}}^t w(u(y_0(x))) dx \\ &\leq \sum_{j=1}^{i-1} 2LK_j(q_j - q_{j-1}) + 2LK_i(t - q_{i-1}) \end{aligned}$$

for all  $t \in [q_{i-1}, q_i]$ . Hence

$$w(P_u(y_0)(t)) \leq \sum_{j=1}^i 2LK_j(q_j - q_{j-1}) \leq \sum_{j=1}^i 2LK_j|Q|. \quad (3)$$

Putting  $W_i = \sum_{j=1}^i 2LK_j(q_j - q_{j-1})$ , equation (2) combined with equation (3) yields

$$\begin{aligned} [-q_i\|v(0)\|, q_i\|v(0)\|]^n + [-W_i, W_i]^n &\subseteq t\|v(0)\| + [-W_i, W_i]^n \\ &\subseteq P_u(y_0)(t) \end{aligned}$$

for all  $t \in [q_{i-1}, q_i]$ .

By Lemma 5.6 we have  $q_i\|v(0)\| + W_i \leq K_i$ , and therefore  $[-K_i, K_i]^n \subseteq P_u(y_0)(t)$  for all  $t \in [q_{i-1}, q_i]$ , as required. By an analogous argument, the same relation holds for  $t \in [-q_i, -q_{i-1}]$ . Note that for  $0 \leq t < q_i$  we find, by the induction hypothesis, that  $[-K_i, K_i]^n \subseteq [-K_{i-1}, K_{i-1}]^n \subseteq P_u(y_0)(t)$ , and hence the proof is complete.  $\square$

Since the estimate on convergence speed (Theorem 4.9) also remains valid in this extended setting, we have our next theorem.

**THEOREM 5.9.** *Suppose that  $Q \in \mathcal{P}[0, a]$  is a partition with  $|Q| \leq 1/2L$ , and let  $y_{k+1} = P_u(y_k)$ . Then  $w(y_k) \in \mathcal{O}(2^{-k})$  and  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  is real-valued and solves the IVP (1).*

*Proof.* The estimate on the convergence speed is identical to the proof of Theorem 4.9, and the result follows, since  $y$  is a real-valued fixpoint of  $P_u$ .  $\square$

We now discuss the case where the extension  $u$  of the classical vector field  $v$  is given in terms of an increasing sequence of approximations  $u = \bigsqcup_{k \in \mathbb{N}} u_k$ . For this, we need to reconsider the measure of convergence speed of the  $u_k$  to  $u$ .

**DEFINITION 5.10.** If  $r, u : \mathbf{IR}^n \rightarrow \mathbf{IR}^m$  and  $K \geq 0$ , the *restricted distance*  $d_K(r, u)$  is given by

$$d_K(r, u) = \sup\{d(r(\alpha), u(\alpha)) \mid \alpha \in \mathbf{I}[-K, K]^n\}.$$

If  $u = \bigsqcup_{k \geq 0} u_k$ , we say that  $d_c(u, u_k) \in \mathcal{O}(2^{-k})$  if, for all  $K \geq 0$ , we have  $d_K(u, u_k) \in \mathcal{O}(2^{-k})$ .

That is, we say that the sequence  $(u_k)$  converges exponentially fast to  $u$  if it converges exponentially fast on all compact sets; notationally, this is reflected by using the subscript 'c' in the statement  $d_c(u, u_k) \in \mathcal{O}(2^{-k})$ .

We now establish that working with approximations  $(u_k)$  of  $u$  does not destroy the convergence to a solution, and then we give an estimate of the convergence speed.

First, note that for  $u' \sqsubseteq u$ , it is no longer guaranteed that  $P_{u'}(y) \in \mathcal{S}_Q$  for all  $y \in \mathcal{S}_Q$ . This problem is addressed in the next lemma, where  $2\mathcal{S}_Q = \{2y \mid y \in \mathcal{S}_Q\}$ .

**LEMMA 5.11.** *Suppose that  $Q \in \mathcal{P}[0, a]$  with  $|Q| < 1/2L$ ,  $u' \sqsubseteq u$  with  $d_{2K^Q}(u, u') \leq (1/2)\|v(0)\|$ . Then  $P_{u'}(y) \in 2\mathcal{S}_Q$  for all  $y$  with  $y \in 2\mathcal{S}_Q$ .*

*Proof.* Similarly to the proof of Lemma 5.8, it suffices to show that the claim holds for

$$y = 2y_0 = \bigsqcup_{i=1, \dots, k} [-q_i, q_i] \setminus [-2K_i, 2K_i],$$

and to verify that  $tv(0) \in P_{u'}(y)(t)$  for all  $t \in [0, a]$ . We have, for  $t \in [q_{i-1}, q_i]$ :

$$\begin{aligned} w(P_{u'}(y)(t)) &= \int_0^t w(u'(y(x))) dx \\ &\leq \int_0^{q_i} w(u(y(x))) + 2d_{2K^Q}(u, u') dx \\ &\leq \sum_{j=1}^i (q_j - q_{j-1})(L \cdot 4K_i + 2d_{2K^Q}(u, u')) \\ &\leq \sum_{j=1}^i 4K_j L (q_j - q_{j-1}) + q_i \|v(0)\|. \end{aligned}$$

Again, if we let

$$W_i = \sum_{j=1}^i 4K_j L (q_j - q_{j-1}) + q_i \|v(0)\|,$$

we have

$$P_{u'}(y)(t) \supseteq [-q_i \|v(0)\|, q_i \|v(0)\|]^n + [-W_i, W_i]^n,$$

and as

$$q_i \|v(0)\| + W_i \leq 2q_i \|v(0)\| + \sum_{j=1}^i (q_j - q_{j-1}) 4LK_j \leq 2K_i$$

by Lemma 5.6, the result follows.  $\square$

We now show that the order of the convergence speed of the iterates to the solution remains unaffected if we compute the iterates using approximations of the vector field. This is similar to Lemma 4.8, except that we need an additional condition on the distance of  $u$  and  $u'$ .

**LEMMA 5.12.** *Suppose that  $Q \in \mathcal{P}[0, a]$  with  $|Q| \leq 1/2L$  and  $u' \sqsubseteq u$  with  $d_{2K^Q}(u, u') \leq (1/2)\|v(0)\|$  and  $y \in 2\mathcal{S}_Q$ . Then*

$$w_\alpha(P_{u'}(y)) \leq \frac{L}{\alpha} w_\alpha(y) + \frac{2}{\alpha e} d_{2K^Q}(u, u').$$

*Proof.* This is as for Lemma 4.8.  $\square$

Moving from weighted width to ordinary width, we obtain the main result of this section: fast convergence of the Picard iterates for unbounded vector fields.

**THEOREM 5.13.** *Suppose that  $u = \bigsqcup_k u_k$  with  $d_c(u, u_k) \in \mathcal{O}(2^{-k})$ . For  $k \geq 0$ , put  $y_{k+1} = P_{u_k}(y_k)$  and  $y = \bigsqcup_k y_k$ . Then  $P_u(y) = y$  and  $w(y_k) \in \mathcal{O}(2^{-k})$ .*

*Proof.* We proceed as in the proof of Theorem 4.9, but invoking Lemma 5.12 instead of Lemma 4.8.  $\square$

Our next goal is to show that this algorithm can be restricted to bases of the respective domains, showing that it can be implemented without loss of accuracy. We then give an estimate of the algebraic complexity of the algorithm.

## 6. An implementation framework for solving IVPs

We now show that the algorithm contained in Proposition 4.7 is indeed implementable, by showing that the computations can be carried out in the bases of the domains. In fact, we demonstrate that every increasing chain of (interval-valued) vector fields  $(u_k)_{k \in \mathbb{N}}$ , where each  $u_k$  is a base element of  $\mathcal{V}$ , gives rise to a sequence of base elements of  $\mathcal{S}$ , which approximate the solution and converge to it. Our discussion is restricted to the situation where the vector field  $v : [-K, K]^n \rightarrow [-M, M]^n$  is defined on a rectangle containing the origin, but all the results adapt to the unbounded setting outlined in Section 5. Recall that in this case we have to make the assumption that  $aM \leq K$  on the lifetime of a solution to guarantee well-definedness of the problem.

In view of the algorithm contained in Proposition 4.7, we consider simple step functions as the base of  $\mathcal{V}$ , and piecewise linear function as the base of  $\mathcal{S}$ . Note that in this setup, the domain-theoretic Picard operator computes integrals of piecewise constant functions, and hence produces piecewise linear functions.

We begin by introducing the bases with which we are going to work.

**DEFINITION 6.1.** Let  $D \subseteq \mathbb{R}$ , and assume that  $-a = a_0 < \dots < a_k = a$  with  $a_0, \dots, a_k \in D$ ,  $\beta_0, \dots, \beta_k \in \mathbf{I}[-K, K]_D^n$  and  $\gamma_1, \dots, \gamma_k \in \mathbf{I}[-M, M]_D^n$ , where  $\mathbf{IR}_D$  denotes the set of rectangles which are contained in  $R$  and whose endpoints lie in  $D$ . We consider the following classes of functions.

(i) The class  $\mathcal{S}_D^L$  of piecewise  $D$ -linear functions  $[-a, a] \rightarrow \mathbf{I}[-K, K]^n$ ,

$$f = (a_0, \dots, a_k) \searrow_{\mathcal{D}}^L (\beta_0, \dots, \beta_k)$$

where

$$f(x)^\pm = \beta_{j-1}^\pm + \frac{x - a_{j-1}}{a_j - a_{j-1}} (\beta_j^\pm - \beta_{j-1}^\pm) \quad \text{for } x \in [a_{j-1}, a_j].$$

Every component of a  $D$ -linear function is piecewise linear and at  $a_0, a_1, \dots, a_k$  takes values in  $D$ .

(ii) The set  $\mathcal{S}_D^C$  of piecewise  $D$ -constant functions  $[-a, a] \rightarrow \mathbf{I}[-K, K]^n$ ,

$$f = (a_0, \dots, a_k) \searrow_{\mathcal{D}}^C (\beta_1, \dots, \beta_n), \quad x \mapsto \begin{cases} \beta_i, & x \in [a_{i-1}, a_i]^\circ, \\ \beta_{i-1} \sqcap \beta_i, & x = a_i \text{ and } 1 < i < k, \end{cases}$$

where  $\sqcap$  denotes the greatest lower bound and  $(\cdot)^o$  is interior relative to the interval  $[-a, a]$ . The components of a  $D$ -constant function assume constant values in  $D$ , which only change at  $a_0, a_1, \dots, a_k$ .

- (iii) The set  $\mathcal{V}_D$  of finite suprema of consistent step functions  $\mathbf{I}[-K, K]^n \rightarrow \mathbf{I}[-M, M]^n$ ,

$$f = \bigsqcup_{1 \leq j \leq k} \beta_j \searrow \gamma_j : x \mapsto \bigsqcup \{\gamma_j \mid 1 \leq j \leq k, \beta_j \ll x\},$$

where  $(\beta_i \searrow \gamma_i)_{1 \leq i \leq k}$  are consistent if  $\beta_i^o \cap \beta_j^o \neq \emptyset \implies \gamma_i \cap \gamma_j \neq \emptyset$  for all  $1 \leq i, j \leq k$ .

- (iv) For any  $f$  as above, we put  $\mathcal{N}(f) = k$  and call it the *complexity of representation* of  $f$ . In more detail, we let the complexity of representation be given by

$$\begin{aligned} \mathcal{N}((a_0, \dots, a_k) \searrow^L (\beta_0, \dots, \beta_k)) &= \mathcal{N}((a_0, \dots, a_k) \searrow^C (\beta_1, \dots, \beta_k)) \\ &= \mathcal{N}\left(\bigsqcup_{1 \leq j \leq k} \beta_j \searrow \gamma_j\right) \\ &= k. \end{aligned}$$

The notation  $\mathcal{N}(\cdot)$ , which measures the size of representations, is as in [7]. Since we will not consider different representations for the same functions, we allow ourselves to blur the distinction between a function and its representation as a step function. The last section shows how to obtain a representation of  $u$  as a supremum of step functions.

If  $D$  is dense in  $\mathbb{R}$ , it is well known that the sets defined above are bases of their respective superspaces.

**PROPOSITION 6.2.** *Suppose that  $D \subseteq \mathbb{R}$  is dense and  $-a, a \in D$ . Then*

- (i)  $\mathcal{S}_D^L$  and  $\mathcal{S}_D^C$  are bases of  $\mathcal{S}$ ;
- (ii)  $\mathcal{V}_D$  is a base of  $\mathcal{V}$ .

We can now show that the Picard operator  $P_u$  associated with a simple step function  $u$  restricts to an endofunction on the set of basis elements of the space of linear step functions  $\mathcal{S}_D^L$ , and give estimates for the algebraic complexity of the procedure. The following lemma covers the operation of applying the vector field to an approximation of the solution, and simplifies [7, Algorithm 4.4].

**LEMMA 6.3.** *Suppose that  $D \subseteq \mathbb{R}$  is a subfield,  $u \in \mathcal{V}_D$  and  $y \in \mathcal{S}_D^L$ . Then we can effectively find  $f \in \mathcal{S}_D^C$  with  $\mathcal{N}(f) \leq 3\mathcal{N}(y)\mathcal{N}(u)$  and  $u \circ y(x) = f(x)$  for all but finitely many  $x \in [-a, a]$ . Moreover,  $f$  can be computed in time  $\mathcal{O}(\mathcal{N}(u)^2\mathcal{N}(y))$ .*

*Proof.* First suppose that  $u = \beta \searrow \gamma$  consists of a single step function, and that  $y = (a_0, \dots, a_k) \searrow^L \beta$  with  $\beta = (\beta_0, \dots, \beta_k)$ . In every open interval  $(a_{j-1}, a_j)$  we can find an open (possibly empty) subinterval  $\alpha_j \subseteq (a_{j-1}, a_j)$  such that  $y(x) \ll \gamma$  if and only if  $x \in \alpha_j$  for all  $x \in (a_{j-1}, a_j)$ : put  $\alpha_j = \bigcap_{1 \leq i \leq n} \{x \in (a_{j-1}, a_j) \mid y_i(x) \in \beta_i^o\}$ . As  $D \subseteq \mathbb{R}$  is a subfield, we have  $\alpha^\pm \in D$ . This gives

$$u(y(x)) = \begin{cases} \gamma, & x \in \alpha_1 \cup \dots \cup \alpha_k, \\ \perp, & \text{otherwise,} \end{cases}$$

for all but finitely many  $x \in [-a, a]$ .

Since  $\alpha_j$  can be computed in constant time (for every subinterval  $(a_{j-1}, a_j)$ ) and the  $\alpha_j$  are pairwise disjoint intervals satisfying  $\alpha_j^+ \leq \alpha_k^-$  for all  $j \leq k$ , we can compute  $f \in \mathcal{S}_D^C$  with  $f(x) = u(y(x))$  for all but finitely many  $x$  in time  $\mathcal{O}(\mathcal{N}(y))$ . Moreover, since  $\alpha_j$  splits every subinterval  $(a_{j-1}, a_j)$  into at most three parts, we can achieve  $\mathcal{N}(f) \leq 3\mathcal{N}(y)$ .

In the case  $u = \bigsqcup_{1 \leq j \leq l} \gamma_j \searrow \beta_j$ , we have

$$u \circ y = \bigsqcup_{1 \leq j \leq l} (\gamma_j \searrow \beta_j) \circ y.$$

For every  $j$ , we can compute  $f_j$  with  $f_j(x) = (\gamma_j \searrow \beta_j)(y(x))$  for all but finitely many  $y$  in  $\mathcal{O}(\mathcal{N}(y))$  algebraic steps. Hence  $\sup_j f_j$  can be performed in  $\mathcal{O}(\mathcal{N}(u)^2 \cdot \mathcal{N}(y))$  steps, taking into account that we need  $\mathcal{O}(\mathcal{N}(u)^2)$  steps to compute the support points where  $u \circ y$  changes its (piecewise constant) value. Note that the interval  $(-a, a)$  is subdivided into at most  $3\mathcal{N}(y)\mathcal{N}(u)$  parts. We have  $u \circ y(x) = f(x)$  for all but finitely many  $x$ , by construction.  $\square$

Now that we have a basis representation of  $u \circ y$ , it is easy to obtain a basis representation of  $P_u(y)$  by integration. Note that computing integrals can be performed over a base defined over a subring of  $\mathbb{R}$ ; we will make use of this fact later. We now consider integration of base functions, which plays the part of function updating [7, Algorithm 3.3].

**LEMMA 6.4.** *Suppose that  $D \subseteq \mathbb{R}$  is a subring and let  $g(x) = \int_0^x f(x) dx$  for  $f \in \mathcal{S}_D^C$ . Then  $g \in \mathcal{S}_D^L$  and  $\mathcal{N}(g) = \mathcal{N}(f)$ . Furthermore,  $g$  can be computed in  $\mathcal{O}(\mathcal{N}(f))$  steps.*

*Proof.* Let  $f = (a_0, \dots, a_k) \searrow_{\mathcal{D}}^C (\beta_1, \dots, \beta_k)$ . First suppose that  $0 \in \{a_0, \dots, a_k\}$ . Every component  $f_i = [f_i^-, f_i^+]$  consists of a pair of piecewise constant functions. On every interval  $[a_{j-1}, a_j]$ , for  $1 \leq j \leq k$ , the integral of  $f_i^\pm$  can be computed by multiplying the width of the interval by the value of  $f_i^\pm$ , and hence  $g \in \mathcal{S}_D^L$  since  $D \subseteq \mathbb{R}$  is a subring. This computation takes constant time; hence  $g$  can be computed in time  $\mathcal{O}(\mathcal{N}(f))$ , and clearly  $\mathcal{N}(g) = \mathcal{N}(f)$ . In the case  $0 \notin \{a_0, \dots, a_k\}$  we insert 0 as additional partition point and obtain  $\mathcal{N}(g) = \mathcal{N}(f) + 1$ , and  $g$  can be computed in  $\mathcal{O}(\mathcal{N}(f) + 1) = \mathcal{O}(\mathcal{N}(f))$  steps.  $\square$

Summing up, we have the following estimate on the algorithm induced by Proposition 4.7 if we compute over the base of piecewise linear functions.

**PROPOSITION 6.5.** *Suppose that  $D \subseteq \mathbb{R}$  is a subfield,  $u \in \mathcal{V}_D$  and  $y \in \mathcal{S}_D^L$ . Then*

- (i)  $P_u(y) \in \mathcal{S}_D^L$ ;
- (ii)  $P_u(y)$  can be computed in time  $\mathcal{O}(\mathcal{N}(u)^2 \mathcal{N}(y))$ ;
- (iii)  $\mathcal{N}(P_u(y)) \in \mathcal{O}(\mathcal{N}(u)\mathcal{N}(y))$ .

*Proof.* Lemma 6.4 provides us with  $f = (a_0, \dots, a_k) \searrow_{\mathcal{D}}^C (\beta_1, \dots, \beta_k)$  with  $\mathcal{N}(f) \in \mathcal{O}(\mathcal{N}(u) \cdot \mathcal{N}(y))$  such that  $u \circ y = f$  for all but finitely many arguments. Hence

$$P_u(y)(x) = \int_0^x (u \circ y)(t) dt = \int_0^x f(t) dt,$$

and the claims follow from Lemma 6.4.  $\square$

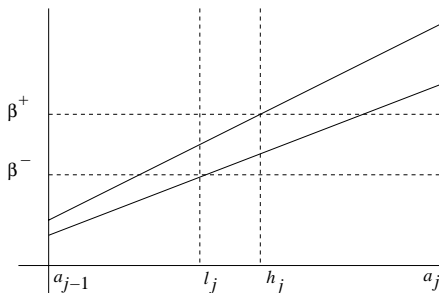


Figure 2: Subdivision of intervals.

We can now summarise our results for computing with piecewise linear functions as follows.

**THEOREM 6.6.** *Suppose that  $D \subseteq \mathbb{R}$  is a subfield and  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  with  $u_k \in \mathcal{V}_D$ . If  $y_{k+1} = P_{u_k}(y_k)$ , then*

- (i)  $y_k \in \mathcal{S}_D^L$  for all  $k \in \mathbb{N}$ ;
- (ii)  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  has width 0 and  $y^- = y^+$  solves the IVP (1);
- (iii)  $w(y_k) \in \mathcal{O}(2^{-k})$  if  $d(u, u_k) \in \mathcal{O}(2^{-k})$ .

Since the elements of  $\mathcal{S}_D^L$  for  $D = \mathbb{Q}$ , the set of rational numbers, can be represented faithfully on a digital computer, the theorem — together with Proposition 3.7 — guarantees soundness and completeness also for implementations of the domain-theoretic method, albeit at the expense of an exponential number of intermediate points. (This is addressed in the next section.) We also provide a guarantee on the speed of convergence, since the condition  $d(u, u_k) \in \mathcal{O}(2^{-k})$  can always be ensured by the library used to construct the sequence  $(u_k)$  of approximations to the vector field, which is discussed in Section 8.

Also, computing over the base of piecewise linear functions eliminates the need for computing rectangular enclosures at every step of the computation. This increase in accuracy comes at the expense of a high complexity of the representation of the iterates. The next section presents an alternative, which uses piecewise constant functions only.

### 7. Computing with piecewise constant functions

We have seen that the time needed to compute  $P_u(y)$  is quadratic in the complexity of the representation of  $u$  and linear in that of  $y$ . However, the complexity of the representation of  $P_u(y)$  is as in Proposition 6.5(iii). This implies that

$$\mathcal{N}(y_{k+1}) \in \mathcal{O}(\mathcal{N}(u_0) \dots \mathcal{N}(u_k)),$$

if  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  and  $y_{k+1} = P_{u_k}(y_k)$ .

The blow-up of the complexity of the representation of the iterates is due to the fact that each interval on which  $y$  is linear is subdivided when computing  $u \circ y$ , since we have to intersect linear functions associated with  $y$  with constant functions induced by  $u$ , as illustrated by Figure 2.

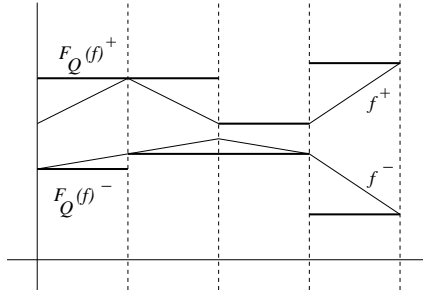


Figure 3: Flattening.

This can be avoided if we work with piecewise constant functions only. The key idea is to transform the linear step function  $P_u(y)$  into a simple step function before computing the next iterate: on every interval, replace the upper (linear) function by its maximum and the lower function by its minimum. We now develop the technical apparatus that is needed to show that the approximations so obtained still converge to the solution. Technically, this is achieved by using partitions of the interval  $[-a, a]$ , where we use the following terminology. (Recall that  $\mathcal{P}[x, y]$  denotes the set of partitions of  $[x, y]$ .)

DEFINITION 7.1 (PARTITIONS). Suppose that  $x \leq y$  are real numbers.

- (i) If  $D \subseteq \mathbb{R}$ , then  $\mathcal{P}_D[x, y] \subset \mathcal{P}[x, y]$  is the subset of partitions of  $[x, y]$  whose points lie in  $D$ .
- (ii) The *size* of a partition  $Q = (q_0, \dots, q_k)$  is given by  $\mathcal{N}(Q) = k$ .
- (iii) A partition  $Q = (q_0, \dots, q_k)$  *refines* a partition  $R = (r_0, \dots, r_l)$  if  $\{r_0, \dots, r_l\} \subseteq \{q_0, \dots, q_k\}$ ; this is denoted by  $R \sqsubseteq Q$ .
- (iv) The *range* of a partition  $Q = (q_0, \dots, q_k)$  is denoted by  $r(Q) = [q_0, q_k]$ .

We are now ready for the definition of the flattening functional, which transforms piecewise linear functions to piecewise constant functions. The following definition uses the notation introduced at the beginning of Section 4.

DEFINITION 7.2. Suppose that  $Q \in \mathcal{P}[-a, a]$ . The *flattening functional*  $F_Q : \mathcal{S} \rightarrow \mathcal{S}$  associated with  $Q$  is defined by

$$F_Q(f) = (q_0, \dots, q_k) \searrow^C (\gamma_1, \dots, \gamma_k),$$

where  $\gamma_i = \bigsqcap \{f(x) \mid x \in [q_{i-1}, q_i]\}$  for  $1 \leq i \leq k$ .

Note that, geometrically speaking,  $F_Q$  computes an enclosure of semi-continuous functions into rectangles, as illustrated by Figure 3.

LEMMA 7.3.  $F_Q$  is well defined and continuous.

*Proof.* Since  $F_Q(f)$  is a step function, we know that  $F_Q(f)$  is continuous, and  $F_Q$  is well defined. We now show that  $F_Q$  is continuous. Let  $f \in \mathcal{S}$ , and assume that

$f = \bigsqcup_{n \in \mathbb{N}} f_n$ . First suppose that  $x \in [q_i, q_{i+1}]^o$ . We find, by Lemma 2.1, that

$$\begin{aligned}
\bigsqcup_{n \in \mathbb{N}} F_Q(f_n)(x) &= \bigsqcup_{n \in \mathbb{N}} \prod \{f_n(x) \mid x \in [q_i, q_{i+1}]\} \\
&= \bigsqcup_{n \in \mathbb{N}} \mathcal{E}(f_n)([q_i, q_{i+1}]) \\
&= \mathcal{E}\left(\bigsqcup_{n \in \mathbb{N}} f_n\right)([q_i, q_{i+1}]) \\
&= \mathcal{E}(f)([q_i, q_{i+1}]) \\
&= \prod \{f(x) \mid x \in [q_i, q_{i+1}]\} \\
&= F_Q(f)(x).
\end{aligned}$$

For  $x \in \{q_0, \dots, q_k\}$ , the claim follows from the continuity of  $\prod : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ .  $\square$

In order to reduce the complexity of the representations of the iterates, we apply the flattening functional at every step of the computation. The following lemma is the stepping stone to proving that this does not affect convergence to the solution. In the context of partitions, we understand ‘increasing’ in terms of the refinement order  $\sqsubseteq$ , introduced in Definition 7.1.

**LEMMA 7.4.** *Suppose that  $(Q_k)_{k \in \mathbb{N}}$  is an increasing sequence of partitions of  $[-a, a]$  with  $\lim_{k \rightarrow \infty} |Q_k| = 0$ . Then  $\bigsqcup_{k \in \mathbb{N}} F_{Q_k} = \text{id}$ .*

*Proof.* This follows from the fact that, for every upper semi-continuous function  $f : [-a, a] \rightarrow \mathbb{R}$  and every decreasing chain  $\alpha_0 \supseteq \alpha_1 \supseteq \dots$  of compact intervals containing  $x$  with  $w(\alpha_k) \rightarrow 0$  as  $k \rightarrow \infty$ , one has  $f(x) = \inf_{k \in \mathbb{N}} \sup \{f(x) \mid x \in \alpha_k\}$ , and the dual statement for lower semi-continuous functions.  $\square$

The last lemma puts us in a position to show that the application of the flattening functional at every stage of the construction does not affect the convergence of the iterates to the solution.

**PROPOSITION 7.5.** *Suppose that  $u = \bigsqcup_{k \in \mathbb{N}} u_k$ ,  $(Q_k)_{k \in \mathbb{N}}$  is an increasing sequence of partitions with  $\lim_{k \rightarrow \infty} |Q_k| = 0$  and  $y_{k+1} = F_{Q_k}(P_{u_k}(y_k))$ . Then  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  satisfies  $y = P_u(y)$ .*

*Proof.* The proof follows from the interchange-of-suprema law (see, for example, [2, Proposition 2.1.12]), the previous lemma and Proposition 4.7.  $\square$

We now show that the speed of convergence is essentially unaffected if we apply the flattening functional at every stage of the computation. This result hinges on the following estimate on Lipschitz functions, whose Lipschitz constant  $N$  is not related to the Lipschitz constant of the vector field.

**LEMMA 7.6.** *Suppose that  $g = ([g_1^-, g_1^+], \dots, [g_n^-, g_n^+]) : [-a, a] \rightarrow \mathbb{R}^n$  is Scott continuous and, for all  $i \in \{1, \dots, n\}$ , either  $g_i^+$  or  $g_i^-$  satisfies a Lipschitz condition with Lipschitz constant  $N$ . If  $Q$  is a partition, then  $w(F_Q(g)) \leq w(g) + N|Q|$ .*

*Proof.* Fix  $1 \leq i \leq n$ , suppose that  $x \in [-a, a]$ , and choose two consecutive partition points  $q^-$  and  $q^+$  of  $Q$  such that  $x \in [q^-, q^+]$ . Since upper and lower semi-continuous functions attain, respectively, their suprema and their infima on compact intervals, there are  $x^-, x^+ \in [q^-, q^+]$  such that, for all  $x \in [q^-, q^+]$ , we have  $F_Q(g)_i^-(x) = g_i^-(x^-)$  and  $F_Q(g)_i^+(x) = g_i^+(x^+)$ , where  $F_Q(g)_i = [F_Q(g)_i^-, F_Q(g)_i^+]$  denotes the  $i$ th component of  $F_Q(g)$ . If we assume without loss of generality that  $g_i^+$  is Lipschitz continuous, we find for  $x \in [q^-, q^+]$  that

$$\begin{aligned} F_Q(g)_i^+(x) - F_Q(g)_i^-(x) &= |g_i^+(x^+) - g_i^-(x^-)| \\ &\leq |g_i^+(x^+) - g_i^+(x^-)| + |g_i^+(x^-) - g_i^-(x^-)| \\ &\leq N|x^+ - x^-| + w(g_i) \\ &\leq N|Q| + w(g), \end{aligned}$$

as required.  $\square$

For the weighted width, we have the following corollary.

**COROLLARY 7.7.** *Under the hypothesis of the previous lemma,*

$$w_\alpha(F_Q(g)) \leq w_\alpha(g) + N|Q|.$$

*Proof.* This follows from calculating

$$\begin{aligned} w_\alpha(F_Q(g)) &= \sup_{t \in [-a, a]} e^{-\alpha|t|} w(F_Q(g)(t)) \\ &\leq \sup_{t \in [-a, a]} e^{-\alpha|t|} (w(g) + N|Q|) \\ &\leq w_\alpha(g) + N|Q| \end{aligned}$$

as claimed.  $\square$

The last corollary allows us to estimate the width of an iterate computed after applying the flattening functional.

**LEMMA 7.8.** *Let  $u' \in \mathcal{V}$  with  $u' \sqsubseteq u$ ,  $y \in \mathcal{S}$  and  $Q \in \mathcal{P}[-a, a]$ . Then*

$$w_\alpha(F_Q(P_u(y))) \leq \frac{L}{\alpha} w_\alpha(y) + \frac{2}{\alpha e} d(u, u') + \frac{K}{\alpha} |Q|.$$

*Proof.* By definition, the upper and lower functions associated with the components of  $g = P_{u'}(y)$ , being integrals of functions bounded by  $M$ , satisfy a Lipschitz condition with Lipschitz constant  $M$ ; we have  $K/\alpha \leq M$ , by assumption. The claim follows from Lemma 4.8 and Corollary 7.7.  $\square$

We can now establish the main result of this section: applying the flattening functional at every step of the computation does not affect the order of the speed of convergence.

**PROPOSITION 7.9.** *Suppose that  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  with  $d(u, u_k) \in \mathcal{O}(2^{-k})$ , and that  $(Q_k)_{k \in \mathbb{N}}$  is an increasing sequence in  $\mathcal{P}[-a, a]$  with  $|Q_k| \in \mathcal{O}(2^{-k})$ . If  $y_{k+1} = F_{Q_k}(P_{u_k}(y_k))$ , then  $w(y_k) \in \mathcal{O}(2^{-k})$ . In particular,  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  is real-valued and solves the IVP (1).*

*Proof.* We can assume, by relabelling the indices of the sequence, that  $d(u, u_k) \leq 2^{-k} \cdot eLw(y_0)$  and  $|Q_k| \leq 2^{-k} \cdot (a/3)w(y_0)$ . We show that  $w_{6L}(y_k) \leq 2^{-k}w_{6L}(y_0)$ , which implies the claim with the help of Lemma 2.2. There is nothing to be shown for  $k = 0$ . For the inductive step we have, by Lemma 7.8,

$$\begin{aligned} w_{6L}(F_{Q_k}(P_{u_k}(y_k))) &\leq \frac{1}{6}w_{6L}(y_0) + \frac{1}{6}2^{-k}w_{6L}(y_0) + \frac{1}{6}K \\ &\leq \frac{1}{6}2^{-k}(w_{6L}(y_0) + w_{6L}(y_0) + w_{6L}(y_0)) \\ &= 2^{-(k+1)}w_{6L}(y_0), \end{aligned}$$

as required.  $\square$

We now show that the application of the flattening functional at every step avoids the blow-up of the size of the iterates. As a consequence, the algorithm with flattening can be implemented using a base of functions defined over a dense subring of  $\mathbb{R}$ , such as the dyadic numbers.

**LEMMA 7.10.** *Suppose that  $D \subseteq \mathbb{R}$  is a subring and  $Q \in \mathcal{P}_D[-a, a]$ . Then  $F_Q$  restricts to a mapping  $\mathcal{S}_D^L \rightarrow \mathcal{S}_D^C$ .*

*Proof.* Suppose that  $f = (a_0, \dots, a_l) \searrow_{\mathcal{D}}^L (\beta_0, \dots, \beta_l) \in \mathcal{S}_D^L$  and  $Q = (q_0, \dots, q_k) \in \mathcal{P}_D[-a, a]$ . If  $F_Q(f) = (q_0, \dots, q_k) \searrow_{\mathcal{D}}^C (\gamma_1, \dots, \gamma_k)$ , then the vertices of the  $\gamma_i$  are elements of the set  $\bigcup_{1 \leq i \leq n} \{f_i^+(q_0), f_i^-(q_0), \dots, f_i^+(q_k), f_i^-(q_k)\}$ , which can be computed from the vertices of the  $\beta_j$  without forming quotients.  $\square$

The complexity of the algorithm underlying Theorem 7.5 over the bases  $\mathcal{V}_D$  and  $\mathcal{S}_D^C$  can now be summarised as follows; recall that  $\mathcal{N}(Q) = k$  is the size of a partition  $Q = (q_0, \dots, q_k)$ .

**PROPOSITION 7.11.** *Suppose that  $D \subseteq \mathbb{R}$  is a subring,  $y \in \mathcal{S}_D^C$  and  $u \in \mathcal{V}_D$ .*

- (i)  $F_Q(P_u(y)) \in \mathcal{S}_D^C$  and  $\mathcal{N}(F_Q(P_u(y))) = \mathcal{N}(Q)$ .
- (ii)  $F_Q(P_u(y))$  can be computed in time  $\mathcal{O}(\max(\mathcal{N}(u) \cdot \mathcal{N}(y), \mathcal{N}(Q)))$ .

*Proof.* For the first statement, assume that

$$y = (a_0, \dots, a_k) \searrow_{\mathcal{D}}^C (\beta_1, \dots, \beta_k) \quad \text{and} \quad u = \bigsqcup_{1 \leq j \leq l} \gamma_j \searrow \delta_j.$$

Then

$$u \circ y = (a_0, \dots, a_k) \searrow_{\mathcal{D}}^C (\beta'_1, \dots, \beta'_k), \quad \text{where } \beta'_m = \bigsqcup \{\delta_j \mid \beta_m \ll \gamma_j\}.$$

Clearly,  $u \circ y \in \mathcal{S}_D^C$ . Computing  $u \circ y$  takes  $\mathcal{O}(\mathcal{N}(u) \cdot \mathcal{N}(y))$  steps, since we have to match every step function in  $u$  against every  $\beta_m$ . By Lemma 6.4 we see that  $P_u(y) \in \mathcal{S}_D^L$ , and finally  $F_Q(P_u(y)) \in \mathcal{S}_D^C$  by Lemma 7.10. Computing  $P_u(y)$  from  $u \circ y$  takes time  $\mathcal{O}(\mathcal{N}(u) \cdot \mathcal{N}(y))$ , and  $F_Q(P_u(y))$  can be computed in  $\mathcal{O}(\max(\mathcal{N}(u) \cdot \mathcal{N}(y), \mathcal{N}(Q)))$  steps, hence the bound on the complexity.  $\square$

Note the complexity reduction compared to Proposition 6.5, which is achieved since  $P_u(f)$  does not change its value in the subintervals  $[a_i, a_{i+1}]$ . We can now summarise our results concerning the soundness and completeness of the algorithm with flattening as follows.

**THEOREM 7.12.** *Suppose that  $D \subseteq \mathbb{R}$  is a subring and  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  with  $u_k \in \mathcal{V}_D$ . Furthermore, assume that  $(Q_k)_{k \in \mathbb{N}}$  is an increasing sequence of partitions with  $\lim_{k \rightarrow \infty} |Q_k| = 0$  and  $y_{k+1} = F_{Q_k}(P_{u_k})(y_k)$ .*

- (i)  $y_k \in \mathcal{S}_D^C$  for all  $k \in \mathbb{N}$  and  $\mathcal{N}(y_k) = \mathcal{N}(Q_k)$ .
- (ii)  $y = \bigsqcup_{k \in \mathbb{N}} y_k$  has width 0 and  $y^- = y^+$  solves the IVP (1).
- (iii)  $w(y_k) \in \mathcal{O}(2^{-k})$ , if both  $d(u, u_k) \in \mathcal{O}(2^{-k})$  and  $|Q_k| \in \mathcal{O}(2^{-k})$ .

Note that, for a subring  $R \subseteq \mathbb{Q}$  of the rational numbers, the elements of  $\mathcal{V}_D$  and  $\mathcal{S}_D^C$  can be faithfully represented on a digital computer. Hence we can also guarantee both the soundness and the completeness for an implementation of the domain-theoretic approach where, furthermore, the representation complexity of the iterates is bounded above by the size of the partitions.

### 8. Approximating continuous functions

The theory outlined in the previous sections depends on an interval vector field  $u$ , given in terms of a supremum  $u = \bigsqcup_{k \in \mathbb{N}} u_k$  of step functions. In order to apply our theory, the following assumptions must be satisfied.

1.  $u$  is an extension of the classical vector field  $v$ .
2.  $u$  satisfies an interval Lipschitz condition.
3. The interval distance  $d(u, u_k)$  converges exponentially fast.

This section shows how to obtain a sequence  $(u_k)_{k \in \mathbb{N}}$  which satisfies the above assumptions. We discuss two techniques for constructing approximations of vector fields: first, we discuss compositions of approximations, and then we show how to construct interval-valued approximations from a function that computes the value of the vector field to an arbitrary degree of accuracy.

#### 8.1. Composition of approximations

In this section we assume that we have two functions  $g : \mathbb{I}\mathbb{R}^n \rightarrow \mathbb{I}\mathbb{R}^m$  and  $f : \mathbb{I}\mathbb{R}^m \rightarrow \mathbb{I}\mathbb{R}^k$ , approximated by sequences of step functions  $(g_n)$  and  $(f_n)$ , and we show how to use these approximations to compute approximations of  $f \circ g$ , subject to the conditions laid down at the beginning of the section.

We begin with an example showing that composition of approximations does not necessarily preserve the convergence speed.

**EXAMPLE 8.1.** This example shows that if  $f = \bigsqcup_k f_k$  and  $g = \bigsqcup_k g_k$ , and both  $(f_k)$  and  $(g_k)$  converge exponentially fast, then this is not necessarily true for the composition  $g \circ f$ , even if both  $f$  and  $g$  are interval Lipschitz.

Consider the continuous function  $h : [0, \infty) \rightarrow [0, 2]$  given by

$$h(x) = \begin{cases} 1 - \frac{1}{\log_2\left(\frac{2}{1-x}\right)}, & x < 1, \\ 1, & x \geq 1, \end{cases}$$

where  $\log_2$  is the dyadic logarithm (logarithm with respect to base 2). Clearly,  $h$  is differentiable in  $[0, 1)$ , and elementary analysis shows that  $0 \leq h'(x) \leq 1/\ln 2 \leq 2$  for  $x \in [0, 1)$ , and hence  $h(x) \leq 2x$  for all  $x \in \mathbb{R}$ . Therefore the Scott continuous

function  $f : \mathbb{I}\mathbb{R} \rightarrow \mathbb{I}\mathbb{R}$  with  $f(x) = [0, h(w(x))]$  satisfies the interval Lipschitz condition  $w(f(x)) \leq 2w(x)$ . Putting  $f_k = f$ , we clearly see that  $d(f, f_k) \leq 2^{-k}$ . Note that  $f$  is a non-maximal interval extension of the constant zero function.

For  $g(x) = [0, w(x)]$  and  $g_k(x) = [0, w(x) + 2^{-k-1}]$  we also find that  $g$  is interval Lipschitz and  $d(g, g_k) = 2^{-k-1} \leq 2^{-k}$ . We show that the composition  $f_k \circ g_k$  only converges linearly fast to  $f \circ g$ . Consider the interval  $x_k = [0, 1 - 2^{-k-1}]$ . Then

$$\begin{aligned} d(f_k \circ g_k, f \circ g) &\geq d(f_k(g_k(x_k)), f(g(x_k))) \\ &= h(w(g_k(x_k))) - h(w(g(x_k))) \\ &= h(1) - h(1 - 2^{-k-1}) \\ &= \frac{1}{k+2}, \end{aligned}$$

showing that function composition does not preserve exponential convergence speed.

As this example shows, we need extra conditions to ensure that composition of approximations preserves the speed of convergence. We propose to consider functions which are Hausdorff Lipschitz from below.

**DEFINITION 8.2.** Suppose that  $f : \mathbb{I}\mathbb{R}^n \rightarrow \mathbb{I}\mathbb{R}^m$ . Then  $f$  is *Hausdorff Lipschitz from below*, if and only if

$$d(f(\alpha), f(\beta)) \leq L \cdot d(\alpha, \beta)$$

for some  $L \geq 0$  and all  $\alpha \sqsubseteq \beta$ ,  $\alpha, \beta \in \mathbb{I}\mathbb{R}^n$ .

Note that we require the estimate to hold only if  $\alpha \sqsubseteq \beta$ ; hence being Hausdorff Lipschitz from below is a weaker condition than being Lipschitz with respect to the Hausdorff metric on  $\mathbb{I}\mathbb{R}^n$  or  $\mathbb{I}\mathbb{R}^m$ , respectively.

We briefly relate this condition to the interval Lipschitz condition that we have introduced already. Recall that  $f$  is interval Lipschitz, if  $w(f(\alpha)) \leq L \cdot w(\alpha)$  for some  $L \geq 0$  and all  $\alpha \in \text{dom}(f)$ ; that is,  $f$  increases the width of its argument only linearly.

**REMARK 8.3.** The notions ‘interval Lipschitz’ and ‘Hausdorff Lipschitz from below’ are unrelated, as shown by the following examples.

1. The function  $f$  in Example 8.1 is interval Lipschitz, but not Hausdorff Lipschitz from below.
2. The function  $\lambda x.[0, 1] : \mathbb{I}\mathbb{R} \rightarrow \mathbb{I}\mathbb{R}$  is Hausdorff Lipschitz from below, but not interval Lipschitz.

It is easy to see that the maximal extension of a classical Lipschitz function is also Hausdorff Lipschitz from below, but the converse is not true in general.

**PROPOSITION 8.4.** *Suppose that  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  satisfies a Lipschitz condition with Lipschitz constant  $L$ . Then  $d(\mathbf{I}f(\alpha), \mathbf{I}f(\beta)) \leq Ld(\alpha, \beta)$  for all compact  $\alpha \sqsubseteq \beta \in \mathbb{I}\mathbb{R}^n$ .*

*Proof.* Suppose that  $\alpha, \beta \in \mathbb{I}\mathbb{R}^n$ . Denote the minimal distance between a point  $x \in \mathbb{R}^n$  and a compact set  $c \subseteq \mathbb{R}^n$  by  $d_m(x, c)$ . By the definition of Hausdorff distance we find, for all  $x \in \alpha$ , an element  $y_x \in \beta$  such that  $\|x - y_x\| \leq d(\alpha, \beta)$ . Hence we find, for all  $x \in \alpha$ , that  $d_m(f(x), \mathbf{I}f(\beta)) \leq L\|x - y_x\| \leq Ld(\alpha, \beta)$ . Therefore  $d(\mathbf{I}f(\alpha), \mathbf{I}f(\beta)) \leq Ld(\alpha, \beta)$ .  $\square$

The next example shows that functions which are Hausdorff Lipschitz from below are not necessarily maximal.

EXAMPLE 8.5. Suppose that  $\mathbb{I}\mathbb{R} \times \mathbb{I}\mathbb{R} \rightarrow \mathbb{R}$  is the maximal extension of the subtraction function; that is,  $[a^-, a^+] - [b^-, b^+] = [a^- - b^+, a^+ - b^-]$ . Then the function  $f : \mathbb{I}\mathbb{R} \rightarrow \mathbb{I}\mathbb{R}$ ,  $x \mapsto x - x$ , is both interval Lipschitz and Hausdorff Lipschitz from below (actually it is even Lipschitz with respect to the Hausdorff distance), but not maximal, as the function  $\lambda x.0$  satisfies  $f \sqsubseteq \lambda x.0$ .

What makes functions that are Hausdorff Lipschitz from below attractive for our purposes is that the set of such functions is closed under composition, in contrast to maximal extensions.

LEMMA 8.6. *Suppose that  $f : \mathbb{I}\mathbb{R}^n \rightarrow \mathbb{I}\mathbb{R}^m$  and  $g : \mathbb{I}\mathbb{R}^m \rightarrow \mathbb{I}\mathbb{R}^k$  are Hausdorff Lipschitz from below. Then so is  $g \circ f$ .*

*Proof.* This follows immediately from the monotonicity of  $g$ . □

Proposition 8.4 and Example 8.5 lead us to think of functions that are Hausdorff Lipschitz from below as functions that are close to being maximal extensions, without actually being maximal. In particular, these functions are closed under composition, which makes them attractive for building libraries.

We are now in a position to prove the promised result on compositionality of approximations; in particular, we establish a guarantee of the convergence speed of composed approximations.

THEOREM 8.7. *Suppose that  $g_k : \mathbb{I}\mathbb{R}^n \rightarrow \mathbb{I}\mathbb{R}^m$  and  $f_k : \mathbb{I}\mathbb{R}^m \rightarrow \mathbb{I}\mathbb{R}^l$  are monotone sequences of Scott continuous functions with  $f = \bigsqcup_k f_k$  and  $g = \bigsqcup_k g_k$  that satisfy the following requirements.*

- (i) *Both  $f$  and  $g$  are interval Lipschitz and  $f$  is Hausdorff Lipschitz from below.*
- (ii)  *$d(f, f_k), d(g, g_k) \in \mathcal{O}(2^{-k})$ .*

*Then  $f \circ g$  is interval Lipschitz and the extension of a classical function and  $d(f_k \circ g_k, f \circ g) \in \mathcal{O}(2^{-k})$ . Moreover, if  $g$  is also Hausdorff Lipschitz from below, then so is  $f \circ g$ .*

*Proof.* Only the statement on the convergence speed requires proof. We denote the Hausdorff Lipschitz constant of  $f$  by  $L$  and assume without loss of generality that  $l = 1$ ; the general result then follows if we take the maximum over the components of  $f$  and  $f_k$ , respectively. Note that

$$d([a^-, a^+], [b^-, b^+]) \leq |a^- - b^-| + |a^+ - b^+| \leq 2d([a^-, a^+], [b^-, b^+]).$$

Using this fact, the claim follows from the following calculation:

$$\begin{aligned} d(f \circ g(\alpha), f_k \circ g_k(\alpha)) &= f_k^+(g_k(\alpha)) - f^+(g(\alpha)) + f^-(g(\alpha)) - f_k^-(g_k(\alpha)) \\ &= f_k^+(g_k(\alpha)) - f^+(g_k(\alpha)) + f^+(g_k(\alpha)) - f^+(g(\alpha)) \\ &\quad + f^-(g(\alpha)) - f^-(g_k(\alpha)) + f^-(g_k(\alpha)) - f_k^-(g_k(\alpha)) \\ &\leq 2d(f, f_k) + 2d(f(g(\alpha)), f(g_k(\alpha))) \\ &\leq 2d(f, f_k) + 2Ld(g, g_k) \in \mathcal{O}(2^{-k}), \end{aligned}$$

where  $\alpha \in \mathbb{I}[-K, K]^n$  was arbitrary. □

This theorem shows that the class of functions that are both interval Lipschitz and Hausdorff Lipschitz from below can be used to build a compositional library for fast-converging Lipschitz functions. In the next section, we address the task of actually constructing functions that fall into this class.

### 8.2. Construction of approximations

Having seen how to obtain approximations of interval vector fields compositionally, this section outlines a technique for constructing these approximations, given a function that computes the Lipschitz function  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  up to an arbitrary degree of accuracy.

More precisely, we assume that  $g : \mathbb{Q}^n \times \mathbb{N} \rightarrow \mathbb{Q}^m$  is given such that

$$\|f(x) - g(x, k)\| \leq 2^{-k}.$$

On a practical level, this allows us to compute approximations for a large class of functions. Moreover, the existence of a *computable* function  $g$  with the above property is equivalent to the computability of  $f$ , and the results of this section show that we obtain approximations by step functions for every *computable* Lipschitz vector field.

The idea of the construction is as follows: Given a rectangle  $\alpha \subseteq \mathbb{R}^n$ , we compute  $g(m(\alpha), k)$ , which gives the value of  $f$  at the midpoint  $m(\alpha)$  of  $\alpha$  up to an accuracy of  $2^{-k}$ . In order to accommodate for this inaccuracy, we extend this point value into a rectangle by extending it with  $2^{-k}$  into the direction of each coordinate axis. This rectangle is then subsequently extended using the Lipschitz constant of  $f$ , resulting in a rectangle that contains all values  $f(x)$  for  $x \in \alpha$ .

While it is straightforward to see that this method produces approximations of any Lipschitz functions up to an arbitrary degree of accuracy, more work is needed to show that these approximations are actually compositional — that is, Hausdorff Lipschitz from below in the sense of the previous section — and converge exponentially fast.

To formalise the construction, we now assume for the rest of the section that  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  satisfies a Lipschitz condition with Lipschitz constant  $L$  and  $g : \mathbb{Q}^n \times \mathbb{N} \rightarrow \mathbb{Q}^m$  is such that  $\|g(x, k) - f(x)\| \leq 2^{-k}$ .

**DEFINITION 8.8.** For a real vector  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  and  $\lambda \in [0, \infty)$ , we write  $x \oplus \lambda$  for the  $n$ -dimensional rectangle  $[x_1 - \lambda, x_1 + \lambda] \times \dots \times [x_n - \lambda, x_n + \lambda]$  with centre  $x$  and width  $2\lambda$ . Given a partition  $Q = (q_0, \dots, q_k)$ , we denote the set of  $n$ -dimensional rectangles with endpoints in  $Q$  by

$$\mathcal{R}(Q) = \{[q_{i_1}, q_{j_1}] \times \dots \times [q_{i_n}, q_{j_n}] \mid 0 \leq i_r < j_r \leq k \text{ for all } 1 \leq r \leq n\}.$$

Finally, we define the family of functions  $f_Q^k$  for  $k \in \mathbb{N}$  by

$$f_Q^k = \bigsqcup_{\alpha \in \mathcal{R}(Q)} \alpha \searrow g(m(\alpha), k) \oplus \left(2^{-k} + \frac{L}{2} \cdot w(\alpha)\right).$$

We call the  $f_Q^k$  the *approximation functions associated with  $Q$* .

It is easy to see that the approximation functions associated with a partition are sound in the sense that they give enclosures of the approximated functions.

LEMMA 8.9. *Let  $Q \in \mathcal{P}$  and  $k \in \mathbb{N}$ . Then  $f_Q^k \sqsubseteq \mathbf{I}f$ .*

*Proof.* We have to show that  $f(x) \in f_Q^k(\alpha)$  for all  $\alpha \in \mathbb{I}\mathbb{R}^n$  and all  $x \in \alpha$ . This follows from

$$f(x) \in g(m(\beta), k) \oplus \left(2^{-k} + \frac{L}{2} \cdot w(\beta)\right) \quad \text{for all } \beta \ll \alpha, \beta \in \mathcal{R}(Q)$$

by taking suprema.

So suppose that  $\beta \in \mathcal{R}(Q)$  and  $\beta \ll \alpha$ . To see that

$$f(x) \in g(m(\beta), k) \oplus (2^{-k} + (L/2) \cdot w(\beta)),$$

it suffices to show that

$$\|f(x) - g(m(\beta), k)\| \leq 2^{-k} + (L/2) \cdot w(\beta) \quad \text{for all } x \in \alpha.$$

But this follows from

$$\begin{aligned} \|f(x) - g(m(\beta), k)\| &\leq \|f(x) - f(m(\beta))\| + \|f(m(\beta)) - g(m(\beta), k)\| \\ &\leq L \cdot \|x - m(\beta)\| + 2^{-k} \\ &\leq L \cdot \frac{1}{2}w(\beta) + 2^{-k} \end{aligned}$$

where the estimate  $\|x - m(\beta)\| \leq (1/2)w(\beta)$  follows from  $\beta \ll \alpha$  and  $x \in \alpha$ .  $\square$

Before we give guarantees on the quality of the approximations constructed using this method, we need to check that the approximations constructed actually form an increasing chain. This is the content of the following easy lemma.

LEMMA 8.10. *Suppose that  $R \sqsubseteq Q \in \mathcal{P}$  and  $j \leq i$ . Then  $f_R^j \sqsubseteq f_Q^i$ .*

We now establish one of the criteria for approximations laid down at the beginning of the section: that is, that they converge to a function which is interval Lipschitz. Recall the order on partitions and their range from Definition 7.1.

LEMMA 8.11. *Suppose that  $(Q_k)_{k \in \mathbb{N}}$  is an increasing sequence of partitions with  $\lim_{k \rightarrow \infty} |Q_k| = 0$  and  $\bigcup_k r(Q_k) = \mathbb{R}$ . Then  $\bigsqcup_{k \in \mathbb{N}} f_{Q_k}^k$  satisfies the interval Lipschitz condition with constant  $L$ .*

*Proof.* Pick  $\alpha \in \mathbb{I}\mathbb{R}^n$ . For any given  $\varepsilon > 0$ , pick  $k \geq 0$  such that  $|Q_k| < \varepsilon/2$ ,  $2^{-k} \leq \varepsilon$  and  $r(Q_k)^n \ll \alpha$ . By the choice of  $k$ , we find  $\beta \in \mathcal{R}(Q_k)$  with  $\beta \ll \alpha$  and  $w(\beta) \leq w(\alpha) + \varepsilon$ . We now have  $\beta \searrow g(m(\beta), k) \oplus (2^{-k} + (L/2) \cdot w(\beta)) \sqsubseteq f_{Q_k}^k$  and  $\beta \ll \alpha$ , whence

$$\begin{aligned} w\left(\bigsqcup_{k \in \mathbb{N}} f_{Q_k}^k(\alpha)\right) &\leq w(f_{Q_k}^k(\alpha)) \\ &\leq w\left((\beta \searrow g(m(\beta), k) \oplus (2^{-k} + \frac{L}{2}w(\beta)))\right)(\alpha) \\ &\leq 2 \cdot 2^{-k} + L \cdot w(\beta) \\ &\leq 2 \cdot 2^{-k} + L \cdot (w(\alpha) + \varepsilon) \\ &\leq \varepsilon + Lw(\alpha) + L\varepsilon \\ &\leq (1 + L)\varepsilon + L \cdot w(\alpha). \end{aligned}$$

As  $\varepsilon > 0$  was arbitrary, we conclude that  $w(\bigsqcup_{k \in \mathbb{N}} f_{Q_k}^k(\alpha)) \leq L \cdot w(\alpha)$ .  $\square$

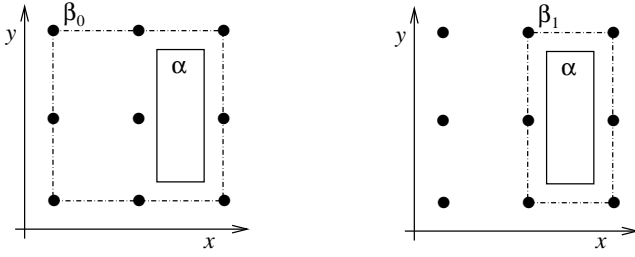


Figure 4: Approximations associated with  $g(x, y, k) = (x, y)$ .

**COROLLARY 8.12.** *The function  $h = \bigsqcup_{k \in \mathbb{N}} f_{Q_k}^k$  is an extension of  $f$ .*

*Proof.* By Lemma 8.9, we have  $f(x) \in h(\{x\})$ , and Lemma 8.11 shows that  $h(\{x\})$  is a singleton set.  $\square$

We have shown how to construct approximations which satisfy two of the three criteria needed to put our theory to work. We now turn to the last item and give an estimate on the convergence speed of the  $f_{Q_k}^k$  to  $h$ . In the proof, we compare an upper approximation of  $(f_{Q_k}^k)^+$  with a lower approximation of  $h^+$  for  $h = \bigsqcup_k f_{Q_k}^k$ . The next lemma is a major stepping-stone to establishing a lower approximation of  $h$ . If we recall the definition of  $f_{Q_k}^k$ , we see that the width of the interval  $m(\beta) \oplus (2^{-k} + (L/2) \cdot w(\beta))$  in the step function  $\beta \searrow m(\beta) \oplus (2^{-k} + (L/2) \cdot w(\beta))$  depends only on the width of  $\beta$ . Hence, given  $\alpha \in \mathbb{I}\mathbb{R}$ , it does not suffice to consider a minimal enclosure  $\mathcal{R}(Q) \ni \beta \ll \alpha$  to find an upper bound for  $f_{Q_k}^k(\alpha)$ . Instead, we need to consider all enclosures that have the same width as the minimal enclosure. This situation is illustrated for  $f(x, y) = g(x, y, k) = (x, y)$  in Figure 4, where the dots indicate the grid points given by  $Q_k$ . Note that

$$(f_{Q_k}^k)^+(\alpha) = g_1(m(\beta_0), k) + 2^{-k} + \frac{L}{2}w(\beta_0)$$

despite the fact that  $\beta_1$  is a better approximation of  $\alpha$ .

The next lemma accounts for this situation, and gives a lower bound for the upper function associated with  $f_{Q_k}^k$ .

**LEMMA 8.13.** *Suppose that  $Q \in \mathcal{P}$  with  $r(Q) \ll [-K, K]$  and  $k \in \mathbb{N}$ . Then, for all  $i = 1, \dots, n$  and all  $\alpha \in \mathbf{I}[-K, K]^n$ ,*

$$(f_{Q_k}^k)^+(\alpha) \geq \min\{f_i(m(\alpha')) \mid \alpha' \sqsubseteq \alpha, w(\alpha') = w(\alpha) + \frac{L}{2}w(\alpha)\},$$

where  $(f_{Q_k}^k)^+$  is the upper function associated with the  $i$ th component of  $f_{Q_k}^k$ .

*Proof.* Throughout the proof, we fix  $i$  with  $1 \leq i \leq n$  and  $\alpha = (\alpha_1, \dots, \alpha_n)$ . First note that

$$\begin{aligned} & \{\alpha' \sqsubseteq \alpha \mid w(\alpha') = w(\alpha)\} \\ &= \{\alpha + (\rho_1, \dots, \rho_n) \mid 0 \in \rho_j \in \mathbf{I}\mathbb{R} \text{ and } w(\alpha_j) + w(\rho_j) \leq w(\alpha)\}, \end{aligned} \quad (4)$$

whence the midpoint set  $M(\alpha) = \{m(\alpha') \mid \alpha' \sqsubseteq \alpha, w(\alpha') = w(\alpha)\}$  is

$$M(\alpha) = m(\alpha) + \frac{1}{2}([w(\alpha_1) - w(\alpha), w(\alpha) - w(\alpha_1)] \times \dots \times [w(\alpha_n) - w(\alpha), w(\alpha) - w(\alpha_n)]).$$

We first show that

$$f_i(m(\beta)) + \frac{L}{2}w(\beta) \geq \min\{f_i(x) \mid x \in M(\alpha)\} + \frac{L}{2}w(\alpha) \quad \text{for all } \beta \sqsubseteq \alpha.$$

Let  $\beta \sqsubseteq \alpha$  be given. In the case where  $m(\beta) \in M(\alpha)$ , there is nothing to show, and the claim follows from  $w(\beta) \geq w(\alpha)$ . Now assume that  $m(\beta) \notin M(\alpha)$ .

For an arbitrary  $x \in \mathbb{R}^n$ , we write  $\|M(\alpha) - x\| = \inf_{y \in M(\alpha)} \|y - x\|$  for the distance between  $x$  and the set  $M(\alpha)$ . Hence our assumption is that  $\|M(\alpha) - m(\beta)\| > 0$ . Pick any  $x \in M(\alpha)$  such that  $\|M(\alpha) - m(\beta)\| = \|x - m(\beta)\|$ , and let  $j \in \{1, \dots, n\}$  be such that  $\|M(\alpha) - m(\beta)\| = \|x - m(\beta)\| = |x_j - m(\beta)_j|$  (recall that  $\|\cdot\|$  denotes maximum norm).

We now claim that  $w(\beta_j) \geq w(\alpha) + 2\|M(\alpha) - m(\beta)\|$ . To see this, recall that  $x \in M(\alpha)$  and  $\|x - m(\beta)\|$  is minimal, and we have two cases, since  $x$  is a boundary point of  $M(\alpha)$ :

$$\text{Case 1: } m(\beta)_j < x_j = m(\alpha)_j - \frac{1}{2}(w(\alpha) - w(\alpha_j));$$

$$\text{Case 2: } m(\alpha_j) + \frac{1}{2}(w(\alpha) - w(\alpha_j)) = x_j < m(\beta)_j.$$

We treat only the first case, as the second is symmetric; so assume that  $m(\beta)_j < x_j$ . If

$$\beta = [b_1^-, b_1^+] \times \cdots \times [b_n^-, b_n^+],$$

we have

$$x_j - \frac{1}{2}(b_j^+ + b_j^-) = x_j - m(\beta)_j = \|M(\alpha) - m(\beta)\|,$$

whence

$$\begin{aligned} b_j^- &= 2x_j - 2\|M(\alpha) - m(\beta)\| - b_j^+ \\ &\leq 2m(\alpha)_j - (w(\alpha) - w(\alpha_j)) - 2\|M(\alpha) - m(\beta)\| - a_j^+ \\ &= a_j^+ + a_j^- - w(\alpha) + a_j^+ - a_j^- - a_j^+ - 2\|M(\alpha) - m(\beta)\| \\ &= a_j^+ - w(\alpha) - 2\|M(\alpha) - m(\beta)\|, \end{aligned}$$

where we have used  $\beta \sqsubseteq \alpha$  to obtain  $a_j^+ \leq b_j^+$  in the second line.

For the same reason, and using the last estimate, we now have

$$\begin{aligned} w(\beta_j) &= b_j^+ - b_j^- \\ &\geq a_j^+ - a_j^+ + w(\alpha) + 2\|M(\alpha) - m(\beta)\|, \end{aligned}$$

which implies our claim that  $w(\beta_j) \geq 2\|M(\alpha) - m(\beta)\| + w(\alpha)$ .

Using this fact, as a consequence of the choice of  $x$  we now have

$$\begin{aligned} f_i(m(\beta)) + \frac{L}{2}w(\beta) &= f_i(m(\beta)) - f_i(x) + f_i(x) + \frac{L}{2}w(\beta) \\ &\geq -L\|x - m(\beta)\| + f_i(x) + \frac{L}{2}(w(\alpha) + 2\|M(\alpha) - m(\beta)\|) \\ &= f_i(x) + \frac{L}{2}w(\alpha) \\ &\geq \min\{f_i(x) \mid x \in M(\alpha)\} + \frac{L}{2}w(\alpha), \end{aligned}$$

which concludes the proof of our first statement. We now show the lemma. As

$$f_Q^k = \bigsqcup_{\beta \in \mathcal{R}(Q)} \beta \searrow g(m(\beta), k) \oplus \left(2^{-k} + \frac{L}{2}w(\beta)\right),$$

it suffices to show that

$$g_i(m(\beta), k) + 2^{-k} + \frac{L}{2}w(\beta) \geq \min\{f_i(x) \mid x \in M(\alpha)\} + \frac{L}{2}w(\alpha) \quad \text{for all } \beta \ll \alpha.$$

But this now follows easily:

$$\begin{aligned} g_i(m(\beta), k) + 2^{-k} + \frac{L}{2}w(\beta) &\geq f_i(m(\beta)) + \frac{L}{2}w(\beta) \\ &\geq \min\{f_i(x) \mid x \in M(\alpha)\} + \frac{L}{2}w(\alpha), \end{aligned}$$

using our first result and the fact that  $\beta \sqsubseteq \alpha$ .  $\square$

We obtain the following immediate corollary, which we use in the estimate of the convergence speed to give an upper bound on  $h(\alpha)$ .

**COROLLARY 8.14.** *Suppose that  $(Q_k)$  is an increasing sequence of partitions and  $h = \bigsqcup_{k \in \mathbb{N}} f_{Q_k}^k$ . Then*

$$h_i^+(\alpha) \geq \min\{f_i(m(\alpha')) \mid \alpha' \sqsubseteq \alpha, w(\alpha') = w(\alpha)\} + \frac{L}{2}w(\alpha) \quad \text{for all } 1 \leq i \leq n.$$

Using the last corollary as an upper bound for the value of  $h$ , we can formulate and prove a statement on the convergence speed as follows.

**PROPOSITION 8.15.** *Suppose that  $(Q_k)$  is an increasing sequence of partitions with  $|Q_k| \in \mathcal{O}(2^{-k})$  and  $\bigcup_k r(Q_k) = \mathbb{R}$ . If  $h = \bigsqcup_k f_{Q_k}^k$ , then  $d(h, f_{Q_k}^k) \in \mathcal{O}(2^{-k})$ .*

*Proof.* Let  $K > 0$ . We show that  $d(h(\alpha), f_{Q_k}^k(\alpha)) \in \mathcal{O}(2^{-k})$  for all  $\alpha \in \mathbf{I}[-K, K]^n$ . Without loss of generality we can assume that  $r(Q_0) \ll [-K, K]$  and  $|Q_k| \leq (1/L) \cdot 2^{-k}$ .

Suppose now that  $\alpha \in \mathbf{I}[-K, K]^n$  is given, and that  $1 \leq i \leq n$ . By the compactness of  $\alpha$  and continuity of  $f$ , we can find  $\alpha' \sqsubseteq \alpha$  with  $w(\alpha') = w(\alpha)$  such that

$$f_i(m(\alpha')) = \min\{f_i(m(\gamma)) \mid \gamma \sqsubseteq \alpha, w(\gamma) = w(\alpha)\}.$$

By Corollary 8.14 we have

$$h_i^+(\alpha) \geq f_i(m(\alpha')) + \frac{L}{2}w(\alpha') \quad (5)$$

(note that  $w(\alpha) = w(\alpha')$ ). As  $|Q_k| \leq (1/L)2^{-k}$ , we can find  $\beta \ll \alpha'$  with  $d(\alpha', \beta) \leq 2|Q_k| = 2(1/L) \cdot 2^{-k}$ . By the definition of  $f_{Q_k}^k$ , we have

$$g(m(\beta), k) \oplus \left(2^{-k} + \frac{L}{2}w(\beta)\right) \sqsubseteq f_{Q_k}^k(\alpha),$$

and hence

$$(f_{Q_k}^k)_i^+(\alpha) \leq g_i(m(\beta), k) + 2^{-k} + \frac{L}{2}w(\beta). \quad (6)$$

Combining equations (5) and (6) we obtain

$$\begin{aligned} (f_{Q_k}^k)_i^+(\alpha) - h_i^+(\alpha) &\leq g_i(m(\beta), k) + 2^{-k} + \frac{L}{2}w(\beta) - f_i(m(\alpha')) - \frac{L}{2}w(\alpha') \\ &\leq f_i(m(\beta)) + 2 \cdot 2^{-k} + \frac{L}{2}(w(\beta) - w(\alpha')) - f_i(m(\alpha')) \\ &\leq L \cdot \|m(\beta) - m(\alpha')\| + \frac{L}{2}d(\alpha', \beta) + 2 \cdot 2^{-k} \\ &\leq \frac{3}{2}Ld(\alpha', \beta) + 2 \cdot 2^{-k} \\ &\leq 3L|Q_k| + 2 \cdot 2^{-k} \\ &\leq 5 \cdot 2^{-k}, \end{aligned}$$

where we have used Lemma 2.3 in lines 3 and 4 of the estimate. Similarly, one shows that  $h_i^-(\alpha) - (f_{Q_k}^k)_i^-(\alpha) \leq 5 \cdot 2^{-k}$ , and we conclude that  $d(h_i(\alpha), (f_{Q_k}^k)_i(\alpha)) \leq 5 \cdot 2^{-k}$ , which implies the claim, as  $i$  was arbitrary.  $\square$

In summary, we have the following theorem, which shows that the approximations satisfy all the conditions discussed at the beginning of the section.

**THEOREM 8.16.** *Suppose that  $(Q_k)$  is an increasing sequence of partitions with  $|Q_k| \in \mathcal{O}(2^{-k})$  and  $\bigcup_{k \geq 0} r(Q_k) = \mathbb{R}$ , and let  $h = \bigsqcup_{k \in \mathbb{N}} f_{Q_k}^k$ . Then*

- (i)  $h$  is an extension of  $f$ ;
- (ii)  $h$  satisfies an interval Lipschitz condition with Lipschitz constant  $L$ ;
- (iii)  $d(h, f_k^{Q_k}) \in \mathcal{O}(2^{-k})$ .

### 8.3. Compositionality of approximations

We have now established conditions which allow us to compose function approximations in such a way that the order of magnitude of the convergence speed is preserved. On the other hand, we have described a method to construct fast-converging approximations from scratch. In this section, we show that the approximations  $f_{Q_k}^k$  are amenable to building a library for approximating Lipschitz functions, by showing that their suprema are Hausdorff Lipschitz from below, which entails that the composition of approximations preserves fast convergence (Theorem 8.7).

For the purposes of this section, we assume that  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a classical Lipschitz function,  $(Q_k)$  is an increasing sequence of partitions with  $|Q_k| \in \mathcal{O}(2^{-k})$ , and  $\bigcup_{k \geq 0} r(Q_k) = \mathbb{R}$ . Furthermore, we assume that  $f_{Q_k}^k$  is constructed as in Definition 8.8.

Our main result is to show that the functions  $h = \bigsqcup_k f_{Q_k}^k$  can be used to build a compositional library of fast-converging approximations to Lipschitz vector fields. In the light of Theorem 8.7, we therefore have to show that the function  $h = \bigsqcup_k f_{Q_k}^k$  is Hausdorff Lipschitz from below.

We fix the function  $h = \bigsqcup_k f_{Q_k}^k$ . The proof of the Hausdorff Lipschitz property is split into several lemmas.

**LEMMA 8.17.** *Suppose that  $\alpha' \sqsubseteq \alpha$  with  $w(\alpha) = w(\alpha')$ . Then there are  $(x_1, \dots, x_n) \in \mathbb{R}^n$  such that:*

- (i)  $|x_i| \leq (1/2)(w(\alpha) - w(\alpha_i))$  for all  $i = 1, \dots, n$ ;
- (ii)  $m(\alpha') = m(\alpha) + (x_1, \dots, x_n)$ .

*Proof.* Suppose that  $\alpha = [\alpha_1^-, \alpha_1^+] \times \dots \times [\alpha_n^-, \alpha_n^+]$ , and similarly for  $\alpha'$ . Then we see, for  $i = 1, \dots, n$ , that

$$\alpha'_i = [\alpha_i^- - a_i, \alpha_i^+ + a_i] \quad \text{where } 0 \leq a_i^-, a_i^+ \text{ and } a_i^- + a_i^+ \leq w(\alpha) - w(\alpha_i).$$

Putting  $x_i = (a_i^+ - a_i^-)/2$  satisfies the hypothesis of the lemma. □

**LEMMA 8.18.** *Let  $\alpha \in \mathbb{IR}^n$ . Then  $h_i^+(\alpha) \leq f_i(m(\alpha)) + (L/2)w(\alpha)$ .*

*Proof.* Let  $\varepsilon > 0$  and find  $k \geq 0$  such that  $2^{-k} \leq \varepsilon$ ,  $|Q_k| \leq \varepsilon$  and  $r(Q_k)^n \ll \alpha$ . By the choice of  $k$ , we can find  $\alpha_0 \in \mathcal{R}(Q_k)$  with  $\alpha_0 \ll \alpha$  and  $d(\alpha, \alpha_0) \leq \varepsilon$ .

We calculate that

$$\begin{aligned}
 h_i^+(\alpha) &\leq (f_k^{Q_k})_i^+(\alpha) \\
 &\leq g_i(m(\alpha_0), k) + 2^{-k} + \frac{L}{2}w(\alpha_0) \\
 &\leq f_i(m(\alpha_0)) + 2 \cdot 2^{-k} + \frac{L}{2}w(\alpha_0) \\
 &\leq f_i(m(\alpha)) + |f_i(m(\alpha_0) - m(\alpha))| + 2 \cdot 2^{-k} + \frac{L}{2}w(\alpha) + \frac{L}{2}|w(\alpha_0) - w(\alpha)| \\
 &\leq f_i(m(\alpha)) + \frac{L}{2}w(\alpha) + L\|m(\alpha) - m(\alpha_0)\| + \frac{L}{2}(w(\alpha_0) - w(\alpha)) + 2 \cdot 2^{-k} \\
 &\leq f_i(m(\alpha)) + \frac{L}{2}w(\alpha) + Ld(\alpha, \alpha_0) + Ld(\alpha, \alpha_0) + 2 \cdot 2^{-k} \\
 &\leq f_i(m(\alpha)) + \frac{L}{2}w(\alpha) + 2L\varepsilon + 2\varepsilon,
 \end{aligned}$$

which proves the claim, as  $\varepsilon$  was arbitrary.  $\square$

The next lemma gives the first half of the Hausdorff Lipschitz property.

LEMMA 8.19. *Let  $\alpha \sqsubseteq \beta \in \mathbb{IR}^n$ . Then  $h_i^+(\alpha) - h_i^+(\beta) \leq 3Ld(\alpha, \beta)$ .*

*Proof.* By the compactness of the midpoint set

$$\{m(\gamma) \mid \gamma \sqsubseteq \beta, w(\beta) = w(\gamma)\},$$

we find  $\beta' \sqsubseteq \beta$  with  $w(\beta') = w(\beta)$  such that

$$h_i^+(\beta) \geq \min\{f_i(m(\gamma)) \mid \gamma \sqsubseteq \beta, w(\gamma) = w(\beta)\} + \frac{L}{2}w(\beta) = f_i(m(\beta')) + \frac{L}{2}w(\beta),$$

where the first estimate is Corollary 8.14. By Lemma 8.17, there are  $x_1, \dots, x_n$  such that  $m(\beta) + (x_1, \dots, x_n) = m(\beta')$  and  $|x_i| \leq (1/2)(w(\beta) - w(\beta_i))$ . We put  $x = (x_1, \dots, x_n)$  and observe that, for  $i = 1, \dots, n$ ,

$$\begin{aligned}
 \frac{1}{2}(w(\beta) - w(\beta_i)) &\leq \frac{1}{2}(w(\alpha) - w(\beta_i) - w(\alpha_i) + w(\alpha_i)) \\
 &\leq \frac{1}{2}(w(\alpha) - w(\alpha_i)) + \frac{1}{2}(w(\alpha_i) - w(\beta_i)) \\
 &\leq \frac{1}{2}(w(\alpha) - w(\alpha_i)) + d(\alpha, \beta).
 \end{aligned}$$

Hence we find  $y_i$  such that  $|y_i| \leq (1/2)(w(\alpha) - w(\alpha_i))$  such that  $|x_i - y_i| \leq d(\alpha, \beta)$  for all  $i = 1, \dots, n$ . Now put  $\alpha' = \alpha'_1 \times \dots \times \alpha'_n$ , where

$$\alpha'_i = \begin{cases} [\alpha_i^-, \alpha_i^+ + 2y_i], & \text{if } y_i \geq 0, \\ [\alpha_i^-, \alpha_i^+], & \text{if } y_i \leq 0, \end{cases}$$

and let  $y = (y_1, \dots, y_n)$ . Then  $\alpha' \sqsubseteq \alpha$  and  $w(\alpha') = w(\alpha)$ . By the monotonicity of  $h$ , we have  $h_i^+(\alpha) \leq h_i^+(\alpha')$ . Using Corollary 8.14 and Lemma 8.18, this gives

$$\begin{aligned}
 h_i^+(\alpha) - h_i^+(\beta) &\leq f_i(m(\alpha')) + \frac{L}{2}w(\alpha) - \left(f(m(\beta')) + \frac{L}{2}w(\beta)\right) \\
 &\leq L\|m(\alpha') - m(\beta')\| + \frac{L}{2}(w(\alpha) - w(\beta)) \\
 &\leq \|m(\alpha) + x - m(\beta) - y\| + Ld(\alpha, \beta) \\
 &\leq L\|m(\alpha) - m(\beta)\| + L\|y - x\| + Ld(\alpha, \beta) \\
 &\leq Ld(\alpha, \beta) + Ld(\alpha, \beta) + Ld(\alpha, \beta),
 \end{aligned}$$

where  $h_i^+(\alpha) \geq h_i^+(\beta)$  follows from the monotonicity of  $h_i$ .  $\square$

As a corollary, we obtain a bound on the difference between the upper values of  $h$ .

COROLLARY 8.20. *Let  $\alpha \sqsubseteq \beta \in \mathbf{IR}^n$  and  $1 \leq i \leq n$ . Then  $|h_i^+(\alpha) - h_i^+(\beta)| \leq 3Ld(\alpha, \beta)$ .*

Similarly, one proves the dual statement that  $|h_i^-(\alpha) - h_i^-(\beta)| \leq 3d(\alpha, \beta)$ . These two results together show that  $h$ , as constructed, is Hausdorff Lipschitz from below.

THEOREM 8.21. *Let  $\alpha \sqsubseteq \beta \in \mathbf{IR}^n$ . Then  $d(h(\alpha), h(\beta)) \leq 3Ld(\alpha, \beta)$ . In particular,  $h$  is Hausdorff Lipschitz from below.*

*Proof.* By Corollary 8.20 and its dual, we see for  $1 \leq i \leq n$ , that

$$\begin{aligned} d(h(\alpha), h(\beta)) &= \max\{|h_i^+(\alpha) - h_i^+(\beta)|, |h_i^-(\alpha) - h_i^-(\beta)|\} \\ &\leq \max\{3Ld(\alpha, \beta), 3Ld(\alpha, \beta)\} \\ &= 3Ld(\alpha, \beta). \end{aligned}$$

Taking the maximum over  $i = 1, \dots, n$  establishes the claim.  $\square$

This shows, together with the results of Section 8.1, that we can build a compositional library for domain-theoretic approximations of Lipschitz vector fields.

In conjunction with Theorem 7.12 we obtain a framework for solving initial value problems that is based on proper data types, and can therefore be directly implemented on a digital computer. Moreover, working with rational or dyadic numbers, the speed of convergence can also be guaranteed for implementations of our technique.

*Acknowledgements* This work has been supported by the EPSRC in the UK, the EU project ‘APPSEM-II’ and a grant by the ‘Deutsche Forschungsgemeinschaft’.

### References

1. O. ABERTH, ‘Computable analysis and differential equations’, *Intuitionism and proof theory*, Proc. of the Summer Conf., Buffalo, NY, 1968, Studies in Logic and the Foundations of Mathematics (North-Holland, 1970) 47–52. 83
2. S. ABRAMSKY and A. JUNG, ‘Domain theory’, *Handbook of logic in computer science*, vol. 3 (ed. S. Abramsky, D. M. Gabbay and T. S. E. Maibaum; Clarendon Press, 1994). 84, 85, 88, 91, 104
3. AWA, a software package for validated solution of ordinary differential equations, <http://www.lsi.upc.es/~robert/mirror/interval-comp/intsoft.html>. 83, 85
4. V. BRATTKA, *Computability of Banach space principles*, Informatik Berichte 286 (FernUniversität Hagen, June 2001). 83
5. J. P. CLEAVE, ‘The primitive recursive analysis of ordinary differential equations and the complexity of their solutions’, *J. Comput. System Sci.* 3 (1969) 447–455. 83
6. E. A. CODDINGTON and N. LEVINSON, *Theory of ordinary differential equations* (McGraw-Hill, 1955). 83, 93
7. A. EDALAT, M. KRZNARIĆ and A. LIEUTIER, ‘Domain-theoretic solution of differential equations (scalar fields)’, *Proceedings of MFPS XIX*, Electron. Notes Theoret. Comput. Sci. 83 (Elsevier, Amsterdam, 2003), <http://www.doc.ic.ac.uk/~ae/papers/scalar.ps>. 84, 91, 100, 101

8. A. EDALAT and A. LIEUTIER, 'Domain theory and differential calculus (functions of one variable)', *Seventh Annual IEEE Symposium on Logic in Computer Science* (IEEE Computer Society Press, 2002), <http://www.doc.ic.ac.uk/~ae/papers/diffcal.ps>. 84, 86
9. A. EDALAT and A. LIEUTIER, 'Domain theory and differential calculus (functions of one variable)', *Math. Structures Comput. Sci.* 14 (2004) 771–802. 84, 91
10. G. GIERZ, K. H. HOFMANN, K. KEIMEL, J. D. LAWSON, M. MISLOVE and D. S. SCOTT, *Continuous lattices and domains* (Cambridge University Press, 2003). 85, 86
11. GNU: *The GNU multi precision library*, <http://www.swox.com/gmp/>. 85
12. A. GRZEGORCZYK, 'Computable functionals', *Fund. Math.* 42 (1955) 168–202. 83, 84
13. A. ISERLES, *Numerical analysis of differential equations*, Cambridge Texts in Applied Mathematics (Cambridge University Press, 1996). 83
14. KER-I KO, 'On the computational complexity of ordinary differential equations', *Inform. Contr.* 58 (1983) 157–194. 83
15. A. N. KOLMOGOROV and S. V. FOMIN, *Introductory real analysis* (Dover, 1975). 87
16. R. E. MOORE, *Interval analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1966). 83, 89
17. N. TH. MÜLLER and B. MOISKE, 'Solving initial value problems in polynomial time', *Proceedings of the 22nd JAIIO - Panel'93*, Buenos Aires, 1993, <http://www.informatik.uni-trier.de/~mueller/Forschung/ivp-panel93.pdf>. 83
18. N. S. NEDIALKOV, K. R. JACKSON and G. F. CORLISS, 'Validated solutions of initial value problems for ordinary differential equations', *Appl. Math. Comput.* 105 (1999) 21–68. 83
19. M. B. POUR-EL and J. I. RICHARDS, 'A computable ordinary differential equation which possesses no computable solution', *Ann. Math. Logic* 17 (1979) 61–90. 84
20. M. B. POUR-EL and J. I. RICHARDS, *Computability in analysis and physics* (Springer, 1988). 83, 84
21. K. WEIHRAUCH, *Computable analysis (an introduction)* (Springer, 2000). 84

Abbas Edalat [ae@doc.ic.ac.uk](mailto:ae@doc.ic.ac.uk)

Dirk Pattinson [dirk@doc.ic.ac.uk](mailto:dirk@doc.ic.ac.uk)

Department of Computing  
 Imperial College London  
 United Kingdom