A Domain Theoretic Account of Picard's Theorem

A. Edalat and D. Pattinson
Department of Computing, Imperial College, London, UK

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, respectively from below and above, 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 an implementation framework to solve initial value problems, which is sound and, in contrast to techniques based on interval analysis, also complete: the unique solution can be actually computed within any degree of required accuracy.

1 Introduction

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

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

where the vector field $v : O \to \mathbb{R}^n$ is continuous in a neighbourhood $O \subseteq \mathbb{R}^n$ of the origin, and we look for a differentiable function $y = (y_1, \ldots, y_n) : [-a, a] \to \mathbb{R}^n$, defined in a neighbourhood of $0 \in \mathbb{R}$, that satisfies (1). By a theorem of Peano there is always a solution [9, page 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 [13, 3, 8, 15, 20, 18, 6].

On the algorithmic and more practical side, standard numerical packages for solving IVP's 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 [14].

Interval analysis [17] 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 [19, 1]. In the interval analysis approach, arithmetic operations are performed on intervals, and outward rounding is applied if the resulting interval endpoints are not machine representable. While this strategy guarantees soundness, i.e. containment of the exact result in the computed interval, one has in general no control over the rounding, which can produce unduly large intervals. As a consequence, for an implementation of the framework for solving IVP's based on interval analysis, one cannot in general guarantee completeness, that is, actual convergence to the solution. For the same reason, one has no control over the speed of convergence.

Domain theory [4] 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, respectively from below and above, 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. As a consequence, 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 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 [10], which used a domain for C^1 functions [11] and, at each stage of iteration, required a new approximation of 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.

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 coefficients. Using piecewise linear functions, we avoid the computation of rectangular enclosures of the solution, and we therefore reduce the wrapping effect, a well known phenomenon in interval analysis. 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 [2] 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 [1]. Of course we have to bear in mind that 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

$$v = (v_1, \ldots, v_n) : [-K, K]^n \to [-M, M]^n$$

which is defined in a compact rectangle containing the origin and consider the IVP given by Equation (1). Note that any continuous function on a compact rectangle is bounded, hence we can assume, without loss of generality, that v takes values in $[-M, M]^n$.

We construct solutions $y : [-a, a] \to \mathbb{R}^n$ of Equation (1) where $a > 0$ satisfies $aM \leq K$. This will guarantee that the expression $v(y)$ is well defined, since M is a bound for the derivative of y. We consider the n-dimensional Euclidean space $\mathbb{R}^n$ equipped with the maximum norm $\|x\| =$

$\max\{|x_1|, \ldots, |x_n|\}$, as this simplifies dealing with the Lipschitz conditions, which we introduce later. Approximations of real numbers live in the interval domain

$$\mathbb{IR} = (\{[a,b] \mid a,b \in \mathbb{R}, a \leq b\} \cup \{\mathbb{R}\}, \sqsubseteq) \text{ with } \alpha \sqsubseteq \beta \Leftrightarrow \beta \subseteq \alpha$$

ordered by reverse inclusion; the way below relation is given by $\alpha \ll \beta$ iff $\beta \subseteq \alpha^o$, where $(\cdot)^o$ denotes the interior of a set. For $n \geq 1$, the domain $\mathbb{IR}^n$ is isomorphic to the domain of n-dimensional rectangles $\{\alpha_1 \times \cdots \times \alpha_n \mid \alpha_i \in \mathbb{IR} \text{ for all } 1 \leq i \leq n\}$, and we do not distinguish between these two presentations. For a rectangle $R \subseteq \mathbb{R}^n$, the subset $\{S \in \mathbb{IR}^n \mid S \subseteq R\}$ of rectangles contained in R is a sub-domain of $\mathbb{IR}^n$, which is denoted by $\mathbf{I}R$. The powers $\mathbb{IR}^n$ of the interval domain and the sub-domain $\mathbf{I}R$, for a rectangle R, 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 $[\alpha_1^-, \alpha_1^+] \times \cdots \times [\alpha_n^-, \alpha_n^+]$. Similarly, if $f : X \to \mathbb{IR}^n$ is a function, we write $f = [f^-, f^+]$ if $f(x) = [f^-(x), f^+(x)]$ for all $x \in X$.

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

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

for all $x \in R$. Note that every continuous function $f : R \to \mathbb{R}^n$ has a canonical maximal extension $\mathbf{I}f$ defined by

$$\mathbf{I}f = (\mathbf{I}f_1, \ldots, \mathbf{I}f_n) : \mathbf{I}R \to \mathbb{IR}^n \text{ with } \mathbf{I}f_i(S) = [\inf_{x \in S} f_i(x), \sup_{x \in S} f_i(x)],$$

where $S \in \mathbf{I}R$ is a rectangle.This extension is maximal in the set of interval valued functions extending f. It is easy to see that $\mathbf{I}f$ is continuous w.r.t. the Scott topology on $\mathbf{I}R$ and $\mathbb{IR}^n$ if f is continuous w.r.t. the Euclidean topology.

We consider the following spaces for approximating the vector field and the solutions to the IVP:

- $\mathcal{V} = \mathbf{I}[-K,K]^n \to \mathbf{I}[-M,M]^n$, the set of continuous functions w.r.t. the Scott topology on $\mathbf{I}[-K,K]^n$ and $\mathbf{I}[-M,M]^n$.
- $\mathcal{S} = [-a,a] \to \mathbf{I}[-K,K]^n$, the set of continuous functions w.r.t. the Euclidean topology on $[-a,a]$ and the Scott topology on $\mathbf{I}[-K,K]^n$

That is, we will be concerned with spaces of functions both with interval and real input. For later reference, we include the following lemma, which links both presentations.

Lemma 2.1. *Suppose $P \in \mathrm{I\!R}^k, Q \in \mathrm{I\!R}^l$ and consider the following operations:*

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

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

For the proof, see [12]. 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, where X is an arbitrary set:

- For a rectangle $\alpha = [\alpha^-, \alpha^+]$, $w(\alpha) = \|\alpha^+ - \alpha^-\|$ denotes the *width* of α. If $f : X \to \mathrm{I\!R}^n$ is a function, $w(f) = \sup_{x\in X} w(f(x))$ is the *width* of f. In the special case where $X \subseteq \mathbb{R}$, we let $w_\alpha(f) = \sup_{x\in X} e^{-\alpha|x|} w(f(x))$ and call $w_\alpha(f)$ the *weighted width* of f w.r.t. the weight α; this gives $w(f) = w_0(f)$.
- Given two rectangles $\alpha = [\alpha^+, \alpha^-]$ and $\beta = [\beta^-, \beta^+]$, their *distance* is given by $d(\alpha, \beta) = \max_{1\le i\le n} |\alpha_i^+ - \beta_i^+| + |\alpha_i^- - \beta_i^-|$ if $\alpha^+ = (\alpha_1^+, \dots, \alpha_n^+)$ and similarly for $\alpha^-, \beta^+, \beta^-$. If $f, g : X \to \mathrm{I\!R}^n$ are functions, $d(f, g) = \sup_{x\in X} d(f(x), g(x))$ is the *distance* between f and g.

In the above, the weight α will enable us to show that the domain theoretic Picard operator is a contraction. Considering g as approximation to f, we understand the distance $d(f, g)$ as measure of the quality of the approximation. We mention two simple lemmas linking distance, width and weighted width.

Lemma 2.2. *Letf $: [-a, a] \to \mathrm{I\!R}^n$. Then $w_\alpha(f) \le w(f) \le e^{a\alpha} w_\alpha(f)$ for all $\alpha \ge 0$.*

Lemma 2.3. *For a rectangle $\gamma = [\gamma^-, \gamma^+]$ denote the midpoint of γ by $m(\gamma) = \frac{1}{2}(\gamma_1^+ + \gamma_1^-, \dots, \gamma_n^+, \gamma_n^-)$ and suppose $\alpha, \beta \in \mathrm{I\!R}^n$ are compact.*

(i) $\|m(\alpha) - m(\beta)\| \le \frac{1}{2} d(\alpha, \beta)$

(ii) $0 \le w(\beta) - w(\alpha) \le d(\alpha, \beta)$ *in case* $\beta \sqsubseteq \alpha$.

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

3 The 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 v satisfies a Lipschitz condition [16]. 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 $u : \mathbf{I}[-K,K]^n \to \mathbf{I}[-M,M]^n$ 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 $f = [f^-, f^+] : [-a,a] \to \mathbf{I}[-M,M]$ is Scott continuous. For $x \in [-a,a]$ we let

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

where $\sigma = \mathrm{sgn}(x)$ is the sign of x and $f^0(t) = 0$. If $f = (f_1, \dots, f_n) : [-a,a] \to \mathrm{I\!R}^n$, we let $\int_0^x f(t)dt = (\int_0^x f_1(t)dt, \dots, \int_0^x f_n(t)dt)$.

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 integral is indeed smaller than the upper value. As we are going to assume that the vector field which defines the IVP is bounded, it suffices to consider bounded interval valued functions only. The following shows that our definition is meaningful:

Lemma 3.2. *Suppose $f : [-a,a] \to \mathbf{I}[-M,M]$ is Scott continuous.*

(i) f^- and f^+ are Lebesgue integrable

(ii) $\int_0^x f(t)dt \in \mathrm{I\!R}$ for all $x \in [-a,a]$.

Proof. For Scott continuous f, the functions f^-, f^+ are lower (resp. upper) semi continuous, hence measurable, and integrability follows from boundedness. If $\sigma = \mathrm{sgn}(x)$, then $\sigma f^{-\sigma} \le \sigma f^{\sigma}$ and $\int_0^x f^{-\sigma}(t)dt \le \int_0^x f^{\sigma}(t)dt$ follows from the definition of the ordinary integral. □

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

Lemma 3.3. *Let $f : [-a, a] \to \mathbb{IR}^n$.*

(i) The function $\lambda x. \int_0^x f(t)dt$ is Scott continuous.

(ii) The function $\int : f \mapsto \lambda x. \int_0^x f(t)dt$ is Scott continuous.

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

The domain theoretic Picard operator can now be defined as follows:

Definition 3.4. Suppose $u \in \mathcal{V}$. The *domain theoretic Picard operator* $P_u : \mathcal{S} \to \mathcal{S}$ 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. That $P_u(y) \in \mathcal{S}$ follows from our assumption $aM \leq K$. Lemma 3.3 shows that $P_u(y)$, for $y \in \mathcal{S}$, and P_u itself are continuous. $\square$

In the classical proof of Picard's theorem, one constructs solutions of IVP's as fixpoint of the (classical) Picard operator. The domain theoretic proof replaces Banach's theorem with Klenee's 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 choose the function $y_0 = \lambda t.[-K, K]^n$ with the least possible information as initial approximation.

Theorem 3.6. *Let $u \in \mathcal{V}$ and $y_{k+1} = P_u(y_k)$. Then $y = \bigsqcup_{k \in \mathbb{N}} y_k$ satisfies $P_u(y) = y$.*

Proof. Follows immediately from Kleene's Theorem, see e.g. [4, Theorem 2.1.19]. $\square$

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

Proposition 3.7. *Suppose u is an extension of v and y is the least fixpoint of P_u.*

(i) If $f : [-a, a] \to [-K, K]^n$ solves (1) 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 the continuity of both. $\square$

The previous proposition can be read as a soundness result: Every solution of the IVP is contained in the least fixpoint of the domain theoretic Picard operator.

4 The Lipschitz Case

We can ensure the uniqueness of the solution of the IVP by requiring that the vector field satisfies an interval version of the Lipschitz property. Recall that for metric spaces (M, d) and (M', d'), a function $f : M \to M'$ 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 [17].

Definition 4.1 (Lipschitz Condition). Suppose $u : \mathbf{I}[-K, K]^n \to \mathbf{I}[-M, M]^n$. 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 $v : [-K, K]^n \to [-M, M]^n$, 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 $\alpha_1 \times \cdots \times \alpha_n$ 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 $u_i^-(R(x, y)) \leq v(x) \leq u_i^+(R(x, y))$ 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 u is interval Lipschitz. Then $w(u(\alpha)) = 0$ whenever $w(\alpha) = 0$, and the induced real valued function $\bar{u}$, given by $\bar{u}(x) = z$ iff $u(\{x\}) = \{z\}$ is continuous.*

We now turn to show that the least fixpoint of the Picard operator associated with u has indeed width zero, if u satisfies an interval Lipschitz condition, and 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 for showing that the least fixpoint of the domain theoretic Picard operator actually has width 0 on the whole of $[-a, a]$.

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

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

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

as we had to show. □

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

Proposition 4.5. *Let $y_{k+1} = P_u(y_k)$ for all $k \in \mathbb{N}$. Then $w_{2L}(y_k) \leq 2^{-n} w_{2L}(y_0)$. In particular, $w(y_k) \leq 2^{-k} \cdot e^{2aL} \cdot w(y_0)$ and $\bigsqcup_{k\in\mathbb{N}} y_k$ is real valued and a solution of (1).*

Proof. The first statement follows by induction from Lemma 4.4 (Pick $\alpha = 2L$). The second claim is immediate from Lemma 2.2 and Proposition 3.7. $\square$

In order to be able to compute the integrals, we now consider approximations to u; the basic idea is that every continuous vector field can be approximated by a sequence of step functions (i.e. functions taking only finitely many values), which allows us to compute the integrals involved in calculating the approximations to the solution effectively. The key property which enables us to use approximations also to the vector field is the continuity of the mapping $u \mapsto P_u$.

Lemma 4.6. *The map $P : \mathcal{V} \to \mathcal{S} \to \mathcal{S}$, $u \mapsto P_u$, is continuous.*

Proof. Follows from 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 $u = \bigsqcup_{k\in\mathbb{N}} u_k$ 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. Follows from Theorem 3.6 and continuity of $u \mapsto P_u$ by the interchange-of-suprema law (see e.g. [4, 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 (Proposition 4.5). In 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 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 $u' \sqsubseteq u$ and $y \in \mathcal{S}$. Then $w_\alpha(P_{u'}(y)) \leq \frac{L}{\alpha} w_\alpha(y) + \frac{1}{\alpha e} d(u, u')$.*

Proof. For the i-th component $P_{u'}(y)_i$ we calculate using Lemma 4.4

$$\begin{aligned}
w_\alpha(P_{u'}(y)) &= \sup_{t\in[0,1]} e^{-\alpha|t|} \int_0^t \sigma(t)\cdot u'^{\sigma(t)}(y(x)) - u'^{-\sigma(t)}(y(x))dx \\
&\leq \sup_{t\in[-a,a]} e^{-\alpha|t|} \int_0^t \sigma(t)\cdot (u^{\sigma(t)}(y(x)) - u^{-\sigma(t)}(y(x)) + d(u,u'))dx \\
&\leq \sup_{t\in[-a,a]} e^{-\alpha|t|} w(P_u(y)) + \sup_{t\in[-a,a]} e^{-\alpha|t|}\cdot |t| \cdot d(u,u') \\
&\leq \frac{L}{\alpha} w_\alpha(y) + \frac{1}{e\alpha} d(u,u')
\end{aligned}$$

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

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

Proposition 4.9. *Suppose* $u = \bigsqcup_{k\in\mathbb{N}} u_k$ *with* $d(u,u_k) \leq 2^{-n}\cdot L\cdot e\cdot w(y_0)$ *and* $y_{k+1} = P_{u_k}(y_k)$ *for all* $k \in \mathbb{N}$. *Then* $w_{4L}(y_k) \leq 2^{-k}\cdot w_{4L}(y_0)$*; in particular* $w(y_k) \leq 2^{-k}\cdot e^{4aL}\cdot w(y_0)$ *and* $\bigsqcup_{k\in\mathbb{N}} y_k$ *is real valued and solves (1).*

Proof. We just show that $w_{4L}(y_k) \leq 2^{-n}\cdot w_{4L}(y_0)$ by induction on k; the second claim then follows from Lemma 2.2. There is nothing to show for the case $k = 0$. For the general case, we invoke Lemma 4.8 and obtain

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

as required. ☐

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 Proposition 4.9 provides an estimate on the speed of convergence. 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.

5 An Implementation Framework for Solving IVP's

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.

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

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

Definition 5.1. Let $D \subseteq \mathbb{R}$ and assume that $-a = a_0 < \cdots < a_k = a$ with $a_0, \ldots, a_k \in D$, $\beta_0, \ldots, \beta_k \in \mathbf{I}[-K,K]^n_D$ and $\gamma_1, \ldots, \gamma_k \in \mathbf{I}[-M,M]^n_D$, where R_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 $\mathcal{S}^L_D$ of piecewise D-linear functions $[-a,a] \to \mathbf{I}[-K,K]^n$,

$$f = (a_0, \ldots, a_k) \searrow^L (\beta_0, \ldots, \beta_k)$$

where $f(x)^{\pm} = \beta^{\pm}_{j-1} + \frac{x - a_{j-1}}{a_j - a_{j-1}}(\beta^{\pm}_j - \beta^{\pm}_{j-1})$ for $x \in [a_{j-1}, a_j]$. Every component of a D-linear function is piecewise linear and takes values in D at $a_0, a_1 \ldots, a_k$.

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

$$f = (a_0, \ldots, a_k) \searrow^C (\beta_1, \ldots, \beta_n), x \mapsto \begin{cases} \beta_i & x \in [a_{i-1}, a_i]^o \\ \beta_{i-1} \bigsqcap \beta_i & x = a_i \text{ and } 1 < i < k \end{cases}$$

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 $a_0, a_1, \ldots, a_k$.

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

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

(iv) For any f as above, we put $\mathcal{N}(f) = k$ and call it the *complexity of representation* of f.

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 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 5.2. *Suppose $D \subseteq \mathbb{R}$ is dense and $-a, a \in D$.*

(i) $\mathcal{V}_D$ is a base of $\mathcal{V}$.

(ii) $\mathcal{S}_D^C$ and $\mathcal{S}_D^L$ are bases of $\mathcal{S}$.

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 $\mathcal{S}_D^L$.

Lemma 5.3. *Suppose $D \subseteq \mathbb{R}$ is a subfield, $u \in \mathcal{V}_D$ and $y \in \mathcal{S}_D^L$. Then there is $f \in \mathcal{S}_D^C$ 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 $u = \gamma \searrow \beta$ consists of a single step function and $y = (a_0, \dots, a_k) \searrow^L \beta$ with $\beta = (\beta_0, \dots, \beta_k)$. In every open interval $\alpha = (a_{j-1}, a_j)$, the components of y are linear, hence the interval can be partitioned as $\alpha = (a_{j-1}, l_j] \cup (l_j, h_j) \cup [h, a_j)$ such that $\beta \ll y(x)$ iff $x \in (l_j, h_j)$ on (a_{j-1}, a_j). Computing l_j and h_j can be done by computing, for each component y_i, the intersection of $y_i^{\pm}$ with the constant function $\lambda x.\beta_i^{\pm}$. Since $D \subseteq \mathbb{R}$ is a subfield, $l_j, h_j \in D$. On the interval (a_{j-1}, a_j) we obtain $u(y(x)) = (a_{j-1}, l_j, h_j, a_j) \searrow^C ([-K, K]^n, \gamma, [-K, K]^n)$ for all $x \in [a_{j-1}, a_j] \setminus \{a_{j-1}, l_j, h_h, a_j\}$. Since this computation can be done in constant time, we can compute $f \in \mathcal{S}_D^C$ such that $f(x) = u(y(x))$ for all but finitely many x in time $\mathcal{O}(\mathcal{N}(y))$. We obtain $\mathcal{N}(u \circ y) \leq 3\mathcal{N}(y)$, since every interval (a_{j-1}, a_j) is sub-divided into at most three parts.

In 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 $f_j(x) = (\gamma_j \searrow \delta_j)(y(x))$ 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 right ordering of points in every subinterval (a_{j-1}, a_j), $j = 1, \dots, k$. 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's 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.

Lemma 5.4. *Suppose $D \subseteq \mathbb{R}$ is a subring and let $g(x) = \int_0^x f(x)dx$ for $f \in \mathcal{S}^C_D$. Then $g \in \mathcal{S}^L_D$ and $\mathcal{N}(g) = \mathcal{N}(f)$. Furthermore, g can be computed in $\mathcal{O}(\mathcal{N}(f))$ steps.*

Proof. Let $f = (a_0, \dots, a_k) \searrow^C (\beta_1, \dots, \beta_k)$. First suppose $0 \in \{a_0, \dots, a_k\}$. 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 $f_i^\pm$, hence $g \in \mathcal{S}^L_D$ 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 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 5.5. *Suppose $D \subseteq \mathbb{R}$ is a subfield, $u \in \mathcal{V}_D$ and $y \in \mathcal{S}^L_D$.*

(i) $P_u(y) \in \mathcal{S}^L_D$

(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 5.4 provides us with $f = (a_0, \dots, a_k) \searrow^C (\beta_1, \dots, \beta_k)$ 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 5.4. $\square$

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

Theorem 5.6. *Suppose $D \subseteq \mathbb{R}$ 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) $y_k \in \mathcal{S}^L_D$ *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. 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. The construction of sequences $(u_k)_{k\in\mathbb{N}}$ that approximate u is discussed in Section 7.

Also, computing over the base of piecewise linear functions eliminates the need of computing rectangular enclosures at every step of the computation. This partially avoids the well-known wrapping effect of interval analysis, but it comes at the cost of a high complexity of the representation of the iterates. The next section presents an alternative, which uses piecewise constant functions only.

6 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 also quadratic in general. 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 the left diagram in Figure 1.

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 which is needed to show that the approximations so obtained still converge to the solution. Technically, this is achieved by making the partitions of the interval $[-a, a]$ explicit.

Definition 6.1 (Partitions). Suppose $x \leq y$ are real numbers.

(i) A *partition* of $[x, y]$ is a finite sequence $(q_0, \dots, q_k)$ of real numbers such that $x = q_0 < \dots < q_k = y$; the set of partitions of $[x, y]$ is denoted by

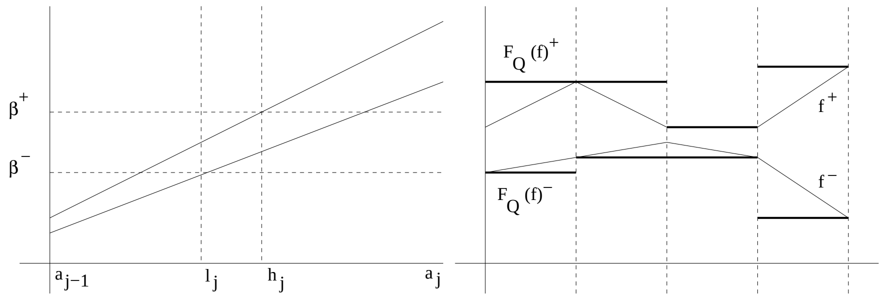

Figure 1: Subdivision of Intervals (left) and Flattening (right)

$\mathcal{P}[x, y]$. 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 *norm* $|Q|$ of a partition $Q = (q_0, \ldots, q_k)$ is given by $|Q| = \max_{1 \le i \le k} q_i - q_{i-1}$ and $\mathcal{N}(Q) = k$ is the *size* of Q.

(iii) A partition $Q = (q_0, \ldots, q_k)$ *refines* a partition $R = (r_0, \ldots, r_l)$ if $\{r_0, \ldots, r_l\} \subseteq \{q_0, \ldots, q_k\}$; this is denoted by $R \sqsubseteq Q$.

We are now ready for the definition of the flattening functional, which transforms piecewise linear functions to piecewise constant functions.

Definition 6.2. Suppose $Q \in \mathcal{P}[-a, a]$. The *flattening functional* $F_Q : \mathcal{S} \to \mathcal{S}$ associated with Q is defined by

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

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

Note that, geometrically speaking, F_Q computes an enclosure of semi continuous functions into rectangles, as illustrated by the right diagram in Figure 1.

Lemma 6.3. *F_Q is well defined and continuous.*

Proof. Since every basis element of $\mathcal{S}^C_{\mathbb{R}}$ is continuous, $F_Q(f)$ is continuous for every $f \in \mathcal{S}$, hence F_Q is well defined. Continuity follows from Lemma 2.1.

□

In order to reduce the complexity of the representations of the iterates, we want to apply the flattening functional at every step of the computation. The following lemma is the stepping stone in 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 6.1.

Lemma 6.4. *Suppose $(Q_k)_{k\in\mathbb{N}}$ is an increasing sequence of partitions with $\lim_{k\to\infty}|Q_k| = 0$. Then $\bigsqcup_{k\in\mathbb{N}} F_{Q_k} = \mathrm{id}$.*

Proof. This follows from the fact that for every upper semi continuous function $f : [-a, a] \to \mathbb{R}$ and every decreasing chain $\alpha_0 \subseteq \alpha_1 \subseteq \dots$ of compact intervals containing x with $w(\alpha_k) \to 0$ as $k \to \infty$ 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 the 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 6.5. *Suppose $u = \bigsqcup_{k\in\mathbb{N}} u_k$, $(Q_k)_{k\in\mathbb{N}}$ is an increasing sequence of partitions with $\lim_{k\to\infty}|Q_k| = 0$ 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. Follows from the interchange-of-suprema law (see e.g. [4, 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:

Lemma 6.6. *Suppose $g = ([g_1^-, g_1^+], \dots, [g_n^-, g_n^+]) : [-a, a] \to \mathbb{IR}^n$ is Scott continuous and, for all $i \in \{1, \dots, n\}$, either g_i^+ or f_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 $x \in [-a, a]$ and choose two consecutive partition points q^-, q^+ of Q such that $x \in [q^-, q^+]$. Since upper (resp. lower) semi continuous functions attain their suprema (resp. 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 ith component of $F_Q(g)$. If we assume w.l.o.g.

that g_i^+ is Lipschitz continuous, we obtain 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. □

For the weighted width, we have the following corollary:

Corollary 6.7. *Under the hypothesis of the previous lemma, $w_\alpha(F_Q(g)) \leq w_\alpha(g) + N|Q|$.*

Proof. We have $w_\alpha(F_Q(g)) = \sup_{t\in[-a,a]} e^{-\alpha|t|} w(F_Q(g)) \leq \sup_{t\in[-a,a]} e^{-\alpha|t|}(w(g) + N|Q|) \leq w_\alpha(g) + 2N|Q|$. □

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

Lemma 6.8. *Let $u' \in \mathcal{V}$ with $u' \sqsubseteq u$, $y \in \mathcal{S}$ and $Q \in \mathcal{P}[-a,a]$. Then $w_\alpha(F_Q(P_u(y))) \leq \frac{L}{\alpha} w_\alpha(y) + \frac{1}{\alpha e} d(u,u') + \frac{K}{a}|Q|$.*

Proof. By definition, the upper and lower functions associated with the components of $g = P_{u'}(y)$ satisfy a Lipschitz condition with Lipschitz constant M; we have $\frac{K}{a} \leq M$ by assumption. The claim follows from Lemma 4.8 and Corollary 6.7. □

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 6.9. *Suppose $u = \bigsqcup_{k\in\mathbb{N}} u_k$ with $d(u,u_k) \leq 2^{-k} \cdot eLw(y_0)$, $(Q_k)_{k\in\mathbb{N}}$ is an increasing sequence in $\mathcal{P}[-a,a]$ with $|Q_k| \leq 2^{-k} \cdot \frac{a}{3} w(y_0)$ and $y_{k+1} = F_{Q_k}(P_{u_k}(y))$ for all $k \geq 0$. Then $w_{6L}(y_k) \leq 2^{-k} w_{6L}(y_0)$; in particular $w(y_k) \leq 2^{-k} \cdot e^{6aL} w(y_0)$ and $\bigsqcup_{k\in\mathbb{N}} y_k$ is real valued and solves (1).*

Proof. We just show the first statement; the second follows by Lemma 2.2. There is nothing to show for $k = 0$. For the inductive step we have by Lemma 6.8

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

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 6.10. *Suppose $D \subseteq \mathbb{R}$ is a subring and $Q \in \mathcal{P}_D[-a,a]$. Then F_Q restricts to a mapping $\mathcal{S}^L_D \to \mathcal{S}^C_D$.*

Proof. Suppose $f = (a_0, \ldots, a_l) \searrow^L (\beta_0, \ldots, \beta_l) \in \mathcal{S}^L_D$ and $Q = (q_0, \ldots, q_k) \in \mathcal{P}_D[-a,a]$. If $F_Q(f) = (q_0, \ldots, q_k) \searrow^C (\gamma_1, \ldots, \gamma_k)$, then the vertices of the γ_i are elements of the set $\bigcup_{1 \leq i \leq n} \{f_i^+(q_0), f_i^-(a_0), \ldots, f_i^+(q_k), f_i^-(a_k)\}$, which can be computed from the vertices of the β_j's without forming quotients. $\square$

The complexity of the algorithm underlying Theorem 6.5 over the bases $\mathcal{V}_D$ and $\mathcal{S}^C_D$ can now be summarised as follows; recall that $\mathcal{N}(Q) = k$ is the size of a partition $Q = (q_0, \ldots, q_k)$.

Proposition 6.11. *Suppose $D \subseteq \mathbb{R}$ is a subring, $y \in \mathcal{S}^C_D$ and $u \in \mathcal{V}_D$.*

(i) $F_Q(P_u(y)) \in \mathcal{S}^C_D$ 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 $y = (a_0, \ldots, a_k) \searrow^C (\beta_1, \ldots, \beta_k)$ and $u = \bigsqcup_{1 \leq j \leq l} \gamma_i \searrow \delta_i$. Then $u \circ y = (a_0, \ldots, a_k) \searrow^C (\beta'_1, \ldots, \beta'_k)$, where $\beta'_m = \bigsqcup\{\gamma_j \mid \beta_j \ll \gamma_j\}$. Clearly $u \circ y \in \mathcal{S}^C_D$. 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 β_m. By Lemma 5.4 we have that $P_u(y) \in \mathcal{S}^L_D$, and finally $F_Q(P_u(y)) \in \mathcal{S}^C_D$ by Lemma 6.10. 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 5.5 which is achieved since $P_u(f)$ does not changes its value in the subintervals $[a_i, a_{i+1}]$. We can now summarise our results concerning soundness and completeness of the algorithm with flattening as follows:

Theorem 6.12. *Suppose $D \subseteq \mathbb{R}$ is a subring and $u = \bigsqcup_{k \in \mathbb{N}} u_k$ with $u_k \in \mathcal{V}_D$. Furthermore, assume $(Q_k)_{k \in \mathbb{N}}$ is an increasing sequence of partitions with $\lim_{k \to \infty} |Q_k| = 0$ and $y_{k+1} = F_{Q_k}(P_{u_k})(y_k)$.*

(i) $y_k \in \mathcal{S}^C_D$ 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* $d(u, u_k)$ *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 guarantee both soundness and completeness also for an implementation of the domain theoretic approach where furthermore the representation complexity of the iterates are bounded above by the size of the partitions.

7 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 needs to satisfy an interval Lipschitz condition
3. The interval distance $d(u, u_k)$ needs to converge 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.

7.1 Composition of Approximations

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

We first treat the case where f is the maximal extension of a classical function, which needs some auxiliary lemmas. The first lemma is needed in the proof that a maximal extension of a classical Lipschitz continuous function satisfies Lipschitz condition of the form $d(\mathbf{I}f(\alpha), \mathbf{I}f(\beta)) \leq C \cdot d(\alpha, \beta)$.

Lemma 7.1. *Suppose* $\alpha, \beta \in \mathrm{I}\mathbb{R}^n$ *are compact with* $\alpha \sqsubseteq \beta$*. Then, for all* $y \in \alpha$ *there is some* $z \in \beta$ *with* $\|y - z\| \leq d(\alpha, \beta)$*.*

Proof. Suppose $\alpha = [a_1^-, a_1^+] \times \cdots \times [a_n^-, a_n^+]$ and $\beta = [b_1^-, b_1^+] \times \cdots \times [b_n^-, b_n^+]$. If $y = (y_1, \ldots, y_n) \in \alpha$ and $i \in \{1, \ldots, n\}$, we have one of the three cases:

$a_i^- \leq y_i \leq b_i^- \leq b_i^+ \leq a_i^+$**:** Putting $z_i = b_i^-$ ensures $|y_i - z_i| \leq b_1^- - a_i^- \leq d(\alpha, \beta)$.

$a_i^- \leq b_i^- \leq b_i^+ \leq y_i \leq a_i^+$**:** Similarly, for $z_i = b_i^+$ we have $|y_i - z_i| \leq a_1^+ - b_i^+ \leq d(\alpha, \beta)$.

$a_i^- \leq b_i^- \leq y_i \leq b_i^+ \leq a_i^+$**:** For $z_i = y_i$ we have $|y_i - z_i| = 0 \leq d(\alpha, \beta)$.

Hence for $z = (z_1, \ldots, z_n)$ we have $\|y - z\| = \max_{1 \leq i \leq n} |z_i - y_i| \leq d(\alpha, \beta)$ and $z \in \beta$. □

Using the lemma above, we can now give an estimate on the distance between the upper functions, evaluated at two intervals.

Lemma 7.2. *Suppose $R \subseteq \mathbb{R}^n$ is a rectangle and $f : R \to \mathbb{R}$ satisfies a Lipschitz condition with Lipschitz constant L. Then $(\mathbf{I}f)^+(\alpha) - (\mathbf{I}f)^+(\beta) \leq L \cdot d(\alpha, \beta)$ for all compact $\alpha, \beta \in \mathbf{I}R$ with $\alpha \sqsubseteq \beta$.*

Proof. By continuity of f and compactness of α, β, there are $y_\alpha \in \alpha$ and $y_\beta \in \beta$ with $(\mathbf{I}f)^+(\alpha) = f(y_\alpha)$ and $(\mathbf{I}f)^+(\beta) = f(y_\beta)$. Using the previous lemma with $y = y_\alpha$ we find $z \in \beta$ with $\|y - z\| \leq d(\alpha, \beta)$. Clearly $f(z) \leq f(y_\beta)$ and $f(y_\alpha) - f(y_\beta) \leq f(y_\alpha) - f(z) \leq L\|y_\alpha - z\| \leq L \cdot d(\alpha, \beta)$ by choice of z. □

As a corollary, we obtain a version of Lipschitz continuity for maximal extensions.

Corollary 7.3. *Suppose $R \subseteq \mathbb{R}^n$ is a rectangle and $f : R \to \mathbb{R}^m$ satisfies a Lipschitz condition with Lipschitz constant L. Then $d(\mathbf{I}f(\alpha), \mathbf{I}f(\beta)) \leq 2Ld(\alpha, \beta)$ for all compact $\alpha \sqsubseteq \beta \in \mathbf{I}R$.*

Proof. For the i-th component f_i of f we have, by the previous lemma and its dual, that $d(\mathbf{I}f_i(\alpha), \mathbf{I}f_i(\beta)) = (\mathbf{I}f)^+(\alpha) - (\mathbf{I}f)^+(\beta) + (\mathbf{I}f)^-(\beta) - (\mathbf{I}f)^-(\alpha) \leq 2L \cdot d(\alpha, \beta)$. □

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

Theorem 7.4. *Suppose $R \subseteq \mathbb{R}^m$ is a compact rectangle, $f_k : \mathbf{I}R \to \mathbb{IR}^l$ and $g_k : \mathbf{I}[-K, K]^n \to \mathbf{I}R$ are monotone sequences of Scott continuous functions satisfying the following requirements:*

1. *$f = \bigsqcup_k f_n$ and $g = \bigsqcup_k g_k$ are extensions of classical functions with f maximal*

2. *$d(f, f_k), d(g, g_k) \in \mathcal{O}(2^{-k})$*

3. *Both f and g are interval Lipschitz.*

Then $f \circ g$ is interval Lipschitz and the extension of a classical function; moreover $d(f_k \circ g_k, f \circ g) \in \mathcal{O}(2^{-k})$.

Proof. Only the statement on the convergence speed requires proof. We denote the interval Lipschitz constant of f by L and assume w.l.o.g. that $l = 1$; the general result then follows by taking the maximum over the components of f resp. f_k. Now the claim follows from the following calculation, where $\alpha \in \mathbf{I}[-K, K]^n$ is arbitrary:

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

where we have used Corollary 7.3 with α replaced by $g_k(\alpha)$ and $\beta = g(\alpha)$ in the last estimate. □

The proof of the theorem hinges on the fact that f is the maximal extension of a classical continuous function. The following example shows, that this hypothesis is necessary by showing that the assertion on the convergence speed may fail if f is not maximal.

Example 7.5. 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 $h : [0, 1] \to [0, 2]$ given by

$$h(x) = \begin{cases} 1 - \frac{1}{\mathrm{ld}(\frac{2}{1-x})} & x < 1 \\ 1 & x = 1 \end{cases}$$

where ld is the dyadic logarithm (logarithm w.r.t. base 2). Clearly h is differentiable in $[0, 1)$, and elementary analysis shows that $0 \leq h'(x) \leq \frac{1}{\ln 2} \leq 2$ for $x \in [0, 1)$, hence $h(x) \leq 2x$ for all $x \in [0, 1]$. Therefore the

Scott continuous function $f(\alpha) = [0, h(w(\alpha))]$ satisfies an interval Lipschitz condition $w(f(\alpha)) \leq 2w(\alpha)$. Putting $f_k = f$, we clearly have that $d(f, f_k) \leq 2^-k$. Note that f is a non-maximal interval extension of the constant zero function.

For $g(\alpha) = [0, w(\alpha)]$ and $g_k(\alpha) = [0, w(\alpha) + 2^{-k-1}]$ we also have 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 $\alpha_k = [0, 1 - 2^{-k-1}]$. Then $d(f_k \circ g_k, f \circ g) \geq d(f_k(g_k(\alpha_k)), f(g(\alpha_k))) = h(w(g_k(\alpha_k))) - h(w(g(\alpha_k))) = h(1) - h(1-2^{-k-1}) = \frac{1}{n}$, showing that function composition does not preserve exponential convergence speed.

The preceding example shows, that we need to work with maximal extensions if we want the composition of two approximating sequences to preserve the convergence speed. However, this imposes no limitation on our approach, since every chain of approximating functions can be converted to a chain approximating the maximal extension. This is demonstrated in the following lemmas. The first lemma shows how we need to modify the approximating functions in order to obtain the maximal extension in the limit. We restrict ourselves here to the case of functions with codomain $\mathbb{R}$; for the general case, our construction has to be repeated for each component.

Lemma 7.6. *Suppose* $f = \bigsqcup_{i \in I} : \mathrm{I}\mathbb{R}^n \to \mathrm{I}\mathbb{R}$ *is an extension of a function* $g : \mathbb{R}^n \to \mathbb{R}$. *If* $m(f_i) = \lambda\beta. \bigsqcap_{x \in \beta} f_i(\{x\})$ *then* $\bigsqcup_{i \in I} m(f_i)$ *is the maximal extension of* g.

Proof. By Lemma 2.1, we have $\bigsqcup_{i \in I} m(f_i) = m(\bigsqcup_{i \in I} f_i)$, hence for all $\beta \in \mathrm{I}\mathbb{R}^n$ we obtain $\bigsqcup_{i \in I} m(f_i)(\beta) = m(\bigsqcup_{i \in I} (f_i)(\beta)) = \bigsqcap_{x \in \beta} \bigsqcup_{i \in I} f_i(\{x\}) = \bigsqcap_{x \in \beta} g(x)$, which shows that $\bigsqcup_{i \in I} m(f_i)$ is the maximal extension of g. $\square$

The next lemma shows, how we can construct the functions $m(f)$ from f, if f is a step function.

Lemma 7.7. *Suppose* $f = \bigsqcup_{1 \leq i \leq k} \beta_i \searrow \gamma_i : \mathrm{I}\mathbb{R}^n \to \mathrm{I}\mathbb{R}$ *and let* $\mathrm{cov}(\beta) = \{J \subseteq \{1, \ldots, k\} \mid \beta \subseteq \bigcup_{j \in J} \beta_j^o\}$ *for* $\beta \in \mathrm{I}\mathbb{R}^n$. *Then* $\bigsqcap_{x \in \beta} f(\{x\}) = \bigsqcup_{J \in \mathrm{cov}(\beta)} \bigsqcap_{j \in J} \gamma_j$.

Proof. Let $\beta \in \mathrm{I}\mathbb{R}^n$ and pick, for every $x \in \beta$, $i^+(x) \in \{1, \ldots, k\}$ such that $\gamma^+_{i^+(x)} = \min\{\gamma_j^+ \mid j = 1, \ldots, k \text{ and } x \in \beta_j^o\}$. We write $i^+(\beta) = \{i^+(x) \mid x \in \beta$. Then $i^+(\beta)$ is finite and $\bigcup_{j \in i^+(x)} \beta_j$ covers β. Therefore

$$(\bigsqcap_{x \in \beta} \bigsqcup \{\gamma_i \mid x \in \beta_i^o\})^+ = \max\{\gamma_j^+ \mid j \in i^+(\beta)\} = (\bigsqcap \{\gamma_j \mid j \in i^+(\beta)\})^+.$$

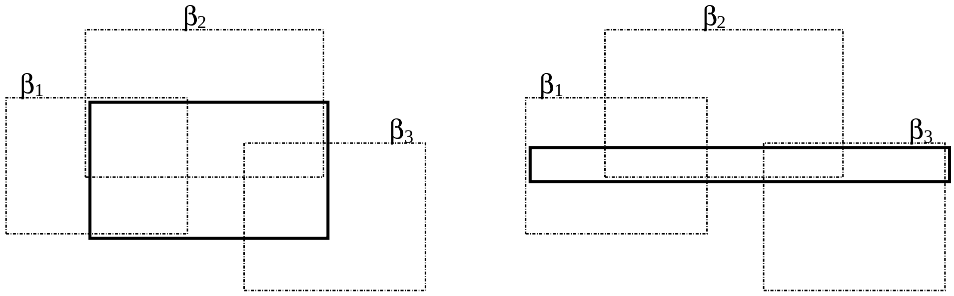

Figure 2: Two elements of $\mathcal{R}(f)$, indicated by solid lines

Similarly, if $1 \leq i^-(x) \leq k$ is such that $\gamma_{i^-(x)} = \max\{\gamma_j^- \mid x \in \beta_j^0\}$ and $i^-(\beta) = \{i^-(x) \mid x \in \beta\}$ we have that

$$(\bigsqcap_{x\in\beta} \bigsqcup\{\gamma_i \mid x \in \beta_i^o\})^- = \min\{\gamma_j^- \mid j \in i^-(\beta)\} = (\bigsqcap\{\gamma_j \mid j \in i^-(\beta)\})^-.$$

Combining the last two equations, we obtain

$$\bigsqcup_{J\in\mathrm{cov}(\beta)} \bigsqcap_{j\in J} b_j = \bigsqcap_{j\in j^-(\beta)} \gamma_j \sqcup \bigsqcap_{j\in i^+(\beta)} \gamma_j = \bigsqcap_{x\in\beta} \bigsqcup\{\gamma_i \mid x \in \beta_i^0\}.$$

that is, our claim. □

Finally, we show how the function $\gamma \mapsto \bigsqcup_{J\in\mathrm{cov}(\gamma)} \bigsqcap_{j\in J} b_j$ can be represented as a step function. This needs the following auxiliary notation: Assume $f = \bigsqcup_{i=1,\dots,k} \beta_i \searrow \gamma_i$ and suppose that $\beta_i = [(\beta_i)_1^-, (\beta_i)_1^+] \times \cdots \times (\beta_i)_n^-, (\beta_i)_n^+]$ for all $i = 1, \dots, k$.

We let $\mathcal{R}_i(f) = \{[b^-, b^+] \mid b^- \leq b^+$ and $b^-, b^+ \in \bigcup_{j=1,\dots,k}\{(\beta_j)_i^-, (\beta_j)_i^+\}$. In other words, the $\mathcal{R}_i$ are intervals whose endpoints are projections of the corners of the β_j's onto the i-th coordinate axes. Finally $\mathcal{R}(f) = \{\alpha_1 \times \cdots \times \alpha_n \mid \alpha_i \in \mathcal{R}_i(f)$ and $(\alpha_1 \times \cdots \times \alpha_n)^o \subseteq \bigcup_{j=1,\dots,k} \beta_j^o\}$. This means that $\mathcal{R}(f)$ is the set of rectangles contained in the domain of f, whose endpoints are intersections of edges of the β_j's.

Two elements of $\mathcal{R}(f)$ for $f = \bigsqcup_{i=1,2,3} \beta_i \searrow \gamma_i$ are shown below.

We are now ready to formulate the last lemma needed to construct approximations to the canonical extension of classical functions.

Lemma 7.8. *Let* $f = \bigsqcup_{i=1,\dots,k} \beta_i \searrow \gamma_i : \mathrm{I\!R}^n \to \mathrm{I\!R}$. *Then* $m(f) = \bigsqcup_{\alpha\in\mathcal{R}(f)} \alpha \searrow \bigsqcup_{J\in\mathrm{cov}(\alpha)} \bigsqcap_{j\in J} \gamma_j$ *where* cov *is as in Lemma 7.7.*

Proof. We split the proof into two parts corresponding to $\sqsubseteq$ and $\sqsupseteq$. Recall from Lemma 7.7, we knot that $m(f)(\beta) = \bigsqcup_{J \in \mathrm{cov}(\beta)} \bigsqcap_{j \in J} \gamma_j$.

Suppose now that $\beta \in \mathbb{IR}^n$ and $\alpha \ll \beta, \alpha \in \mathcal{R}(f)$. Then $\mathrm{cov}(\alpha) \subseteq \mathrm{cov}(\beta)$, hence $\bigsqcup_{J \in \mathrm{cov}(\beta)} \bigsqcap_{j \in J} \gamma_j \sqsubseteq \bigsqcup_{J \in \mathrm{cov}(\beta)} \bigsqcap_{j \in J} \gamma_j$. Therefore

$$\bigsqcup_{\mathcal{R}(f) \ni \alpha \ll \beta} \bigsqcup_{J \in \mathrm{cov}(\alpha)} \bigsqcap_{j \in J} \gamma_j \sqsubseteq \bigsqcup_{J \in \mathrm{cov}(\beta)} \bigsqcap_{j \in J} \gamma_j = m(f)(\beta).$$

We now establish the reverse relation. Let $\alpha_0 = \bigsqcup \{\alpha \in \mathcal{R}(f) \mid \alpha \ll \beta\}$. Then $\alpha_0 \in \mathcal{R}(f)$ and $\alpha_0 \ll \beta$ as $\mathcal{R}(f)$ is finite. Note that $\mathrm{cov}(\alpha_0) = \mathrm{cov}(\beta)$ by construction. Therefore

$$m(f)(\beta) = \bigsqcup_{J \in \mathrm{cov}(\beta)} \bigsqcap_{j \in J} b_j = \bigsqcup_{J \in \mathrm{cov}(\alpha_0)} \bigsqcap_{j \in J} b_j \sqsubseteq \bigsqcup_{\mathcal{R}(f) \ni \alpha \ll \beta} \bigsqcup_{J \in \mathrm{cov}(\alpha)} \bigsqcap_{j \in J} b_j$$

which concludes the proof. $\square$

Note that the left rectangle in Figure 7.1 cannot be covered by $\beta_1, \beta_2, \beta_3$, hence yields the value $\mathbb{R} \in \mathbb{IR}$.

7.2 Construction of Approximations

Now that we have seen how to compose approximations of interval vector fields compositionally, this section outlines a technique for constructing these approximations from scratch, given a function that computes the vector field up to an arbitrary degree of accuracy.

More precisely, we assume that $g = (g_1, \dots, g_n) : [-K, K]^n \cap \mathbb{Q}^n \times \mathbb{N} \to \mathbb{Q}^n$ is given such that $\|f(x) - g(x, k)\| \leq K \cdot 2^{-k}$. 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 v, and the results of this section show, that we obtain approximations by step functions for every *computable* vector field v.

The idea of the construction is as follows: Given a rectangle $\alpha \subseteq [-K, K]^n$, we compute the value of $g(m(\alpha), k)$ of the midpoint $m(\alpha)$ of α up to an accuracy of $K \cdot 2^{-k}$. In order to accommodate for the this inaccuracy, we extend this point value into a rectangle by extending it with $K \cdot 2^- k$ 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 $x \in \alpha$. The formal definition is as follows, where we assume for the rest of the section, that $f : [-K, K]^n \to [-K, K]^n$ satisfies a Lipschitz condition with Lipschitz constant L and $g : [-K, K]^n \cap \mathbb{Q}^n \times \mathbb{N} \to [-K, K]^n \cap \mathbb{Q}^n$ is such that $\|g(x, k) - f(x)\| \leq K \cdot 2^- k$.

Definition 7.9. For a rectangle $\alpha = [a_1^-, a_1^+] \times \cdots \times [a_n^-, a_n^+]$, we denote the *midpoint* of α by $m(\alpha) = (\frac{a_1^+ - a_1^-}{2}, \ldots, \frac{a_n^+ - a_n^-}{2})$. For $\lambda \in \mathbb{R}$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the cube with centre x and width λ is given as

$$x \oplus \lambda = [x_1 - \lambda, x_1 + \lambda], \ldots, [x_n - \lambda, x_n + \lambda].$$

Given a partition $Q = (q_0, \ldots, q_k)$ of $[-K, K]$ we denote by

$$\mathcal{R}(Q) = \{[q_{i_1}, q_{j_1}] \times \cdots \times [q_{i_n}, q_{j_n}] \mid 0 \le i_l < j_l \le k, 1 \le l \le n\}$$

the set of rectangles with corners in Q. Finally, we define the family of functions f_Q^k by

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

We call the f_Q^k's 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 7.10. *Let $Q \in \mathcal{P}[-K, K]$ and $k \in \mathbb{N}$. Then $f_Q^k \sqsubseteq \mathbf{I}f$.*

Proof. We have to show $f(x) \in f_Q^k(\alpha)$ for all $\alpha \in \mathbf{I}[-K, K]^n$ and all $x \in \alpha$. This follows from $x \in g(m(\beta), k) \oplus (K \cdot 2^{-k} + L \cdot w(\beta))$ for all $\beta \ll \alpha, \beta \in \mathcal{R}(Q)$ by taking suprema.

So suppose $\beta \in \mathcal{R}(Q)$ and $\beta \ll \alpha$. To see that $f(x) \in \{g(m(\beta), k)\} \oplus (K \cdot 2^{-k} + L \cdot w(\beta))$ it suffices to show that $\|f(x) - g(m(\beta), k)\| \le K \cdot 2^{-k} + L \cdot w(\beta)$ for all $x \in \alpha$. But this follows from

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

where the estimate $\|x - m(\beta)\| \le w(\beta)$ follows from $\beta \ll \alpha$ and $x \in \alpha$. □

Before we give guarantees on the quality of 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 lemma, whose straightforward proof is omitted.

Lemma 7.11. *Suppose $R \sqsubseteq Q \in \mathcal{P}[-K, K]$ 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, i.e. that they converge to a function which is interval Lipschitz. We recall the order on partitions and their norm from Definition 6.1.

Lemma 7.12. *Suppose $(Q_k)_{k\in\mathbb{N}}$ is an increasing sequence of partitions with $\lim_{k\to\infty} |Q_k| = 0$. Then $\bigsqcup_{k\in\mathbb{N}} f_{Q_k}^k$ satisfies an interval Lipschitz condition with constant L.*

Proof. Pick $\alpha \in \mathbf{I}[-K, K]^n$. For any given $\epsilon > 0$, pick $k \geq 0$ s.t. $|Q_k| < \frac{\epsilon}{2}$ and $K \cdot 2^{-k} \leq \epsilon$. By choice of k, we find $\beta \in \mathcal{R}(Q_k)$ with $\beta \ll \alpha$ and $w(\beta) \leq w(\alpha)+\epsilon$. We now have $\beta \searrow \{g(m(\beta), k)\} \oplus (K \cdot 2^{-k} + L \cdot w(\beta)) \sqsubseteq f_{Q_k}^k$ and $\alpha \ll \beta$, whence

$$\begin{aligned} w(\bigsqcup_{k\in\mathbb{N}} f_{Q_k}^k(\alpha)) &\leq w(f_{Q_k}^k(\alpha)) \\ &\leq K \cdot 2^{-k} + L \cdot w(\beta) \\ &\leq K \cdot 2^{-k} + L \cdot (w(\alpha) + \epsilon) \\ &\leq (1 + L)\epsilon + L \cdot w(\alpha). \end{aligned}$$

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

As immediate corollary, we obtain the fact that $\bigsqcup_{k\in\mathbb{N}} f_{Q_k}^k$ is an extension of f.

Corollary 7.13. *The function $h = \bigsqcup_{k\in\mathbb{N}} f_{Q_k}^k$ is an extension of f.*

Proof. By Lemma 7.11, we have $f(x) \in h(\{x\})$ and Lemma 7.12 shows that $h(\{x\})$ is a singleton set. □

We have now 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 f. In the proof, we compare an upper approximation $u(\alpha) \sqsupseteq f(\alpha)$ with a lower approximation $l(\alpha) \sqsubseteq f_{Q_k}^k(\alpha)$, for a given rectangle α. The next lemma is a major stepping stone for establishing an upper approximation of f. If we recall the definition of $f_{Q_k}^k$, we see that the width of the right hand side of the step function $\alpha \searrow \{m(\alpha)\} \oplus (K \cdot 2^{-k} + L \cdot w(\alpha))$ only depends on the width of α. Hence given $\beta \in \mathbb{IR}$, it does not suffice to consider a minimal enclosure $\mathcal{R}(Q) \ni \alpha \ll \beta$ to find an upper bound for $f_{Q_k}^k(\beta)$. Instead we

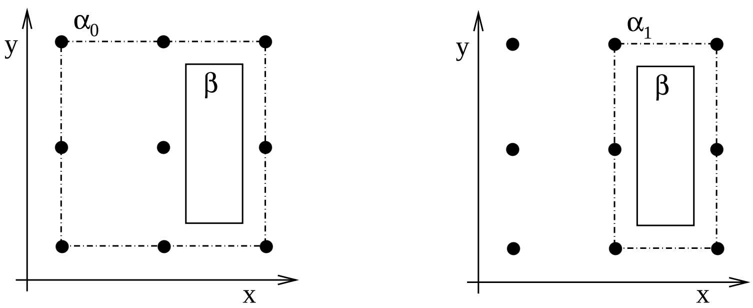

Figure 3: Approximations associated with $g(x, y, k) = (x, y)$.

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 7.2, where the dots indicate the grid points given by Q_k. Note that $f(^k_{Q_k})^+_1(\beta) = g(m(\alpha_0), k) + K \cdot 2^{-k} + \frac{L}{2} w(\alpha_0)$ despite the fact that α_1 is a better approximation of β.

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

Lemma 7.14. *Suppose $Q \in \mathcal{P}[-K, K]$ and $k \in \mathbb{N}$. Then, for all $i = 1, \dots, n$ and all $\alpha \in \mathbf{I}[-K, K]^n$,*

$$(f^k_Q)^+_i(\alpha) \geq \min\{f_i(m(\alpha')) \mid \alpha' \sqsubseteq \alpha, w(\alpha') = w(\alpha)\} + \frac{L}{2} w(\alpha)$$

where $(f^k_Q)^+_i$ is the upper function associated with the i-th component of f^k_Q.

`Check indices.`

Proof. Throughout the proof, we fix $1 \leq i \leq n$. First note that

$$\{\alpha' \sqsubseteq \alpha \mid w(\alpha') = w(\alpha)\} = \{\alpha + (\rho_1, \dots, \rho_n) \mid 0 \in \rho_i \text{ and } w(\alpha_i) + w(\rho_i) \leq w(\alpha)\}$$

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

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

We first show that

$$f^+_i(m(\beta)) + \frac{L}{2} w(\beta) \geq \min\{f_i(x) \mid x \in M(\alpha)\} + \frac{L}{2} w(\alpha)$$

for all $\beta \sqsubseteq \alpha$. Suppose β is given. In case $m(\beta) \in M(\alpha)$ there is nothing to show, and the claim follows from $w(\beta) \geq w(\alpha)$. Now assume $m(\beta) \notin M(\alpha)$.

For an arbitrary $x \in [-K, K]^n$, we write $\|M(\alpha) - x\| = \inf_{y \in M(\alpha)} \|y - x\|$ for the least distance between x and any point in the set $M(\alpha)$. Hence our assumption is $\|M(\alpha) - m(\beta)\| > 0$. Pick any $x \in M(\alpha)$ s.t. $\|M(\alpha) - m(\beta)\| = \|x - m(\beta)\|$ 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, owing to the description of the set $M(\alpha)$ above:
Case 1: $m(\beta)_j < x_j = m(\alpha)_j - \frac{1}{2}(w(\alpha) - w(\alpha_j))$
Case 2: $m(\alpha_j) + \frac{1}{2}(w(\alpha) - w(\alpha_j)) = x_j < m(\beta)_j$.

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

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

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

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 $f_Q^k = \bigsqcup_{\beta \in \mathcal{R}(Q)} g(m(\beta), k) \oplus (K \cdot 2^{-k} + \frac{L}{2}w(\beta))$ it suffices to show that

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

for all $\beta \ll \alpha$. But this now follows easily:

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

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

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

Corollary 7.15. *Suppose (Q_k) is an increasing sequence of partitions and $h = \bigsqcup_{k \in K} f^k_{Q_k}$. Then $h^+_i(\alpha) \geq \min\{f_i(m(\alpha')) \mid \alpha' \sqsubseteq \alpha, w(\alpha') = w(\alpha)\} + \frac{L}{2} w(\alpha)$ 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 7.16. *Suppose (Q_k) is an increasing sequence of partitions with $|Q_k| \leq \frac{K}{L} \cdot 2^{-k}$ and $h = \bigsqcup_k f^k_{Q_k}$. Then $d(h, f^k_{Q_k}) \leq 8K \cdot 2^{-k}$.*

Proof. We show that $d(h(\alpha), f^k_{Q_k}(\alpha)) \leq 8K \cdot 2^{-k}$ for all $\alpha \in \mathbf{I}[-K,K]^n$. So suppose $\alpha \in \mathbf{I}[-K,K]^n$ is given and $1 \leq i \leq n$. By compactness of α and continuity of f, we can find $\alpha' \sqsubseteq \alpha$ with $w(\alpha') = w(\alpha)$ s.t.

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

By Corollary 7.15 we have

$$h^+_i(\alpha) \geq f_i(m(\alpha')) + \frac{L}{2} w(\alpha') \tag{2}$$

(note $w(\alpha) = w(\alpha')$). As $|Q_k| \leq K \cdot 2^{-k}$, we can find $\beta \ll \alpha'$ with $d(\alpha', \beta) \leq 2|Q_k| = 2\frac{K}{L} \cdot 2^{-k}$. By definition of $f^k_{Q_k}$, we have

$$g(m(\beta),k) \oplus (K \cdot 2^{-k} + \frac{L}{2} w(\beta)) \sqsubseteq f^k_{Q_k}(\alpha)$$

hence

$$(f^k_{Q_k})^+_i(\alpha) \leq g_i(m(\beta),k) + K \cdot 2^{-k} + \frac{L}{2} w(\beta). \tag{3}$$

Combining equations (2) and (3) we obtain

$$
\begin{aligned}
(f^k_{Q_k})^+_i(\alpha) - h^+_i(\alpha) &\leq g_i(m(\beta),k) + K\cdot 2^{-k} + \frac{L}{2}w(\beta) - f_i(m(\alpha')) - \frac{L}{2}w(\alpha')\\
&\leq f_i(m(\beta)) + 2K\cdot 2^{-k} + \frac{L}{2}(w(\beta) - w(\alpha'))\\
&\leq L\cdot\|m(\beta) - m(\alpha')\| + \frac{L}{2}d(\alpha',\beta) + 2K\cdot 2^{-k}\\
&\leq Ld(\alpha',\beta) + 2K\cdot 2^{-k}\\
&\leq 2L|Q_k| + 2K\cdot 2^{-k}\\
&\leq 4K\cdot 2^{-k}
\end{aligned}
$$

where we have used Lemma 2.3 in line 3 and 4 of the estimate. Similarly one shows that $h^-_i(\alpha) - (f^k_{Q_k})^-_i \leq 4K\cdot 2^{-k}$, and we conclude that $d(h_i(\alpha),(f^k_{Q_k})_i(\alpha)) \leq 8K\cdot 2^{-k}$ 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 7.17. *Suppose (Q_k) is an increasing sequence of partitions with $|Q_k| \leq \frac{K}{L}\cdot 2^{-k}$ and let $h = \bigsqcup_{k\in\mathbb{N}} f^k_{Q_k}$. Then*

1. *h is an extension of f*
2. *h satisfies an interval Lipschitz condition with Lipschitz constant L*
3. *$d(u,u_k) \leq 8K\cdot 2^{-k}$*

for arbitrary $k\in\mathbb{N}$.

This shows, together with the results of Section 7.1, that we can build a library for approximating vector fields in the domain theoretic sense.

In conjunction with Theorem 5.6 and Theorem 6.12 we obtain a framework for solving initial value problems, which 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 for the first time also be guaranteed for implementations of our technique.

Acknowledgements

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

References

[1] AWA. A software package for validated solution of ordinary differential equations. `www.cs.utep.edu/interval-comp/intsoft.html`

[2] *The GNU multi precision library.* `www.swox.com/gmp/`.

[3] O. Aberth. Computable analysis and differential equations. In *Intuitionism and Proof Theory, Studies in Logic and the Foundations of Mathematics*, pages 47–52. North-Holland, 1970. Proc. of the Summer Conf. at Buffalo N.Y. 1968.

[4] S. Abramsky and A. Jung. Domain theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3. Clarendon Press, 1994.

[5] J. P. Aubin and A. Cellina. *Differential Inclusions.* Springer, 1984.

[6] V. Brattka. Computability of Banach space principles. Informatik Berichte 286, FernUniversität Hagen, Fachbereich Informatik, June 2001.

[7] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory.* Springer, 1998.

[8] J. P. Cleave. The primitive recursive analysis of ordinary differential equations and the complexity of their solutions. *Journal of Computer and Systems Sciences*, 3:447–455, 1969.

[9] E. A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations.* McGraw-Hill, 1955.

[10] A. Edalat, M. Krznarić, and A. Lieutier. Domain-theoretic solution of differential equations (scalar fields). In *Proceedings of MFPS XIX*, volume 83 of *Elect. Notes in Theoret. Comput. Sci.*, 2004. Full Paper in `www.doc.ic.ac.uk/~ae/papers/scalar.ps`.

[11] A. Edalat and A. Lieutier. Domain theory and differential calculus (Functions of one variable). In *Seventh Annual IEEE Symposium on Logic in Computer Science.* IEEE Computer Society Press, 2002. Full paper in `www.doc.ic.ac.uk/~ae/papers/diffcal.ps`.

[12] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, and D. Scott. *Continuous Lattices and Domains.* Cambridge University Press, 2003.

[13] A. Grzegorczyk. Computable functionals. *Fund. Math.*, 42:168–202, 1955.

[14] A. Iserles. *Numerical Analysis of Differential Equations.* Cambridge Texts in Applied Mathematics. CUP, 1996.

[15] Ker-I Ko. On the computational complexity of ordinary differential equations. *Inform. Contr.*, 58:157–194, 1983.

[16] A. N. Kolmogorov and S. V. Fomin. *Introductory Real Analysis.* Dover, 1975.

[17] R.E. Moore. *Interval Analysis.* Prentice-Hall, Englewood Cliffs, 1966.

[18] N. Th. Müller and B. Moiske. Solving initial value problems in polynomial time. In *In Proceedings of the 22th JAIIO - Panel'93*, pages 283–293, Buenos Aires, 1993.

[19] N. S. Nedialkov, K. R. Jackson, and G. F. Corliss. Validated solutions of initial value problems for ordinary differential equations. *Applied Mathematics and Computation*, 105:21–68, 1999.

[20] M. B. Pour-El and J. I. Richards. *Computability in Analysis and Physics.* Springer-Verlag, 1988.