

Average-Case Polynomial-Time Computability of Hamiltonian Dynamics

Akitoshi Kawamura

Kyushu University
Fukuoka, Japan
kawamura@inf.kyushu-u.ac.jp

Holger Thies

University of Tokyo
Tokyo, Japan
info@holgerthies.com

Martin Ziegler

KAIST
Daejeon, Republic of Korea
ziegler@cs.kaist.ac.kr
 <https://orcid.org/0000-0001-6734-7875>

Abstract

We apply average-case complexity theory to physical problems modeled by continuous-time dynamical systems. The computational complexity when simulating such systems for a bounded time-frame mainly stems from trajectories coming close to complex singularities of the system. We show that if for most initial values the trajectories do not come close to singularities the simulation can be done in polynomial time on average. For Hamiltonian systems we relate this to the volume of “almost singularities” in phase space and give some general criteria to show that a Hamiltonian system can be simulated efficiently on average. As an application we show that the planar circular-restricted three-body problem is average-case polynomial-time computable.

2012 ACM Subject Classification Theory of computation → Computational complexity and cryptography

Keywords and phrases Computable Analysis, Real computation, Dynamical systems, Average-case complexity, Computation in physics

Digital Object Identifier 10.4230/LIPIcs.MFCS.2018.30

Acknowledgements This work was supported by the Japan Society for the Promotion of Science (JSPS), Core-to-Core Program (A. Advanced Research Networks), JSPS KAKENHI Grant Numbers JP18H03203 and JP18J10407, the Korean Ministry of Science and ICT grant NRF-2016K1A3A7A03950702. We thank Florian Steinberg for discussions on average-case complexity in analysis and the anonymous reviewers for many helpful comments.

1 Introduction

Many phenomena in nature can be modeled by continuous-time dynamical systems. Analyzing such phenomena is usually done by simulating the evolution of a system with digital computers. It is therefore crucial to better understand the computational properties of dynamical systems. One of the most basic questions one can ask about such a system is given the state of the system at some time t_0 what will be the state at some time $t > t_0$, i.e., to simulate the evolution for a finite time-frame.



© Akitoshi Kawamura, Holger Thies, and Martin Ziegler;
licensed under Creative Commons License CC-BY

43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018).

Editors: Igor Potapov, Paul Spirakis, and James Worrell; Article No. 30; pp. 30:1–30:17

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

The extended Church-Turing thesis is the statement that any physical computation device can be simulated efficiently with a Turing machine. While the thesis might be incompatible with quantum computation, at least for classical (non-quantum) physical computation the assumption seems reasonable. Thus, it should be possible to simulate trajectories of dynamical systems for problems in (classical) physics for a bounded time-frame efficiently as nature already provides an efficient “computation device”. However, accurate numerical simulation of dynamical systems can be very hard. For example, Miller pointed out the difficulties when integrating the famous gravitational N -body problem [16]. The N -body problem is the problem of predicting the motion of N point masses under their mutual gravitational attraction. For a moderate number N of point masses, the main difficulty arises because close encounters of particles cause instabilities. Indeed, two particles colliding leads to a singularity in the analytic function describing the dynamics. On the other hand, Saari could show that at least for $N \leq 4$ singularities are rare in the sense that the set of initial values leading to singularities has Lebesgue measure zero [21, 22] (for $N > 4$ this is an open problem). Thus, a possible resolution to the inconsistency with the extended Church-Turing thesis might be that hard instances are extremely rare in nature and therefore it is possible to do the simulation efficiently on typical inputs. For further discussion on the extended Church-Turing thesis in classical physics, see e.g. [26].

The theory of real number computation based on a realistic model of approximation is known as *computable analysis*. Using this model, computational complexity theory can be applied to real functions [7, 8]. Simulating a continuous-time dynamical system corresponds to solving an initial value problem (IVP) for systems of ordinary differential equations. There are already several results on the computational complexity of solving IVPs in computable analysis. In particular, if the right-hand side function f of the equation $\dot{y} = f(y)$ is polynomial-time computable and Lipschitz-continuous, the unique solution y can be computed in PSPACE and can be hard for this class [6, 4]. On the other hand, for analytic right-hand side function the solution is also a polynomial-time computable function [18, 9]. However, this formulation does not really capture the notion of what is usually understood by simulating a dynamical system, as it is assumed that the solution exists on the whole time interval and only takes values in a known compact set, and there are several hidden factors depending on the function and the initial value that heavily influence the efficiency in practice.

Instead of fixing the initial value, it is therefore more natural to study the complexity of the function mapping initial values and time to the corresponding solution. This, however, poses the problem that the worst-case complexity for most interesting systems is unbounded. Indeed, it is quite obvious that the simulation should take longer the closer a trajectory approaches a singularity of the system as then higher precision is required. Nonetheless, the system might behave well for most initial values in the sense that the trajectory does stay far away from any singularities. We would then expect efficient simulation to be possible on typical inputs.

In this paper we want to formalize this intuition. In classical complexity theory, the notion of being efficiently computable on typical inputs is usually captured by *average-case complexity theory*, which often provides a more significant measure of the performance of an algorithm than worst-case complexity when the hard instances are rare. However, finding the right notion of average-case complexity poses some subtle difficulties. A structural theory of average-case complexity for discrete problems was introduced by Levin [13]. Schröder, Steinberg and Ziegler recently extended Levin’s definition of average-case complexity to problems on real numbers [23].

In this paper we use their definition to show that many physical problems are indeed efficiently solvable on average. We first give a short introduction to the model of computation in Section 2 and define the most important notions that we need in the rest of the paper. In Section 3 we formalize a parameterized result for IVPs with analytic right-hand side. We show that restricted to a compact domain where the dynamics take place the solution can be computed in time polynomial in the output precision and a natural parameter depending on the function and the domain. In Section 4 we use this to show that if the “probability of trajectories to get close to complex singularities of the system” is small and if the right-hand side function can be evaluated efficiently on points not close to singularities, the simulation can be done in polynomial time on average. We then focus on a special case of dynamical systems that play an important role in classical physics, the Hamiltonian systems, and show that there is a simple way to bound the above probability in terms of the volume of singularities in phase-space. Finally, we apply our theorem to show that a special case of the three-body problem, the planar circular restricted three-body problem, can be simulated in polynomial time on average.

2 Computability and complexity in analysis

Computable Analysis extends classical computability theory to deal with uncountable quantities such as real numbers. The theory already dates back to Turing [24] with later important contributions for example by Grzegorzczuk [3] and Lacombe [12]. The rigorous study of complexity in this model was initiated by Ko and Friedman [8]. In this section we briefly summarize the most important definitions. For a more detailed overview the reader is referred, e.g., to Weihrauch’s monograph [25].

2.1 Computing real numbers and functions

Classically, computability is typically defined using Turing machines or an equivalent model of computation. Turing machines compute functions from finite strings to finite strings, that is, functions $F : \Sigma^* \rightarrow \Sigma^*$ for some fixed finite alphabet Σ . While computations over discrete objects like natural numbers, rational numbers or graphs can be defined by choosing an appropriate encoding for these objects as finite strings, the set of reals is uncountable and it is therefore impossible to find such an encoding. On the other hand, any real number can be approximated with arbitrarily small error by rational numbers. A real number x can therefore be encoded by a function that gives arbitrary exact approximations of x . More formally: A function $\varphi : \Sigma^* \rightarrow \Sigma^*$ is called a *name* for a real number $x \in \mathbb{R}$ if it maps strings of length n to the binary encoding of some integer z such that $|x - \frac{z}{2^n}| \leq 2^{-n}$. Any real number has infinitely many different names. We say x is *computable* if it has a computable name.

Similarly, a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is computable if there is a computable function mapping names of x to names of $f(x)$, i.e., an algorithm that computes approximations of the output $f(x)$ with arbitrary precision while having access to arbitrary exact approximations of the input x . Such computation on names can be formally defined using oracle machines. Oracle machines can make queries to an oracle, a function $\varphi : \Sigma^* \rightarrow \Sigma^*$, during the computation. When making such a query the string w on a special tape, the so-called oracle tape, is replaced by the value of $\varphi(w)$ in a single time-step. We denote the oracle machine M with oracle φ by M^φ . M^φ again computes a function $\Sigma^* \rightarrow \Sigma^*$. A computable real function can then be defined as follows: A partial function $f : \subseteq \mathbb{R} \rightarrow \mathbb{R}$ is computable if there is an oracle machine M such that whenever φ is a name for $x \in \text{dom } f$ the machine M^φ computes a name

for $f(x)$. The machine can behave arbitrarily on elements outside the domain of the function. The definition can be easily generalized to multidimensional functions $f : \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ by allowing multiple oracle and output tapes.

The running time $T_M(\varphi, w)$ for an oracle machine M with oracle $\varphi : \Sigma^* \rightarrow \Sigma^*$ and input $w \in \Sigma^*$ is defined as the number of steps the machine makes. If the machine M computes a real function $f : \mathbb{R} \rightarrow \mathbb{R}$ we can thus define the (worst-case) running time $T_M(x, n) \in \mathbb{N} \cup \{\infty\}$ of M for an argument $x \in \mathbb{R}$ and output precision $n \in \mathbb{N}$ as an upper bound of $T_M(\varphi, 0^n)$ over all names φ for x . We further say that a computable function $f : \mathbb{R} \rightarrow \mathbb{R}$ has (worst-case) time-complexity $T : \mathbb{N} \rightarrow \mathbb{N}$ if $T_M(x, n) \leq T(n)$ for all $x \in \text{dom } f$ and that f is polynomial-time computable if T is a polynomial. Thus, the time-complexity of a real function is given in terms of the number of steps necessary to approximate the solution up to precision 2^{-n} depending on the output precision $n \in \mathbb{N}$.

Most work in real complexity theory restricts $\text{dom } f$ to be a fixed compact set as this makes sure that a (finite) complexity bound exists. A more general theory taking into account the “size” of an oracle can be found in [5].

In this paper we sometimes use the notation that f is computable on some maximal domain $\text{dom } f$ and *uniformly* allows some time-bound on a restricted subset of the domain. By this we mean that there is a single algorithm to compute f on its domain and the running time of this algorithm can be bounded on this subset.

2.2 Average-case complexity for real functions

Worst-case complexity as defined above has the disadvantage that the cost can be dominated by a small number of instances. A problem that can be solved efficiently on most inputs might thus still be hard from the perspective of computational complexity. *Average-case complexity* studies the average running time of an algorithm over all inputs and often gives a more realistic measure for the efficiency of an algorithm. Similarly to the class \mathcal{P} in worst-case complexity we would like to have a class of average-case polynomial-time problems that we consider as “efficient on average”. This notion should fulfill some basic robustness criteria. For example, composition of an average-case polynomial-time computable function with a worst-case polynomial-time computable function should still be computable in polynomial time on average. However, already in the discrete case this robustness property is not fulfilled by the most intuitive definitions. A definition that resolves this problem was given by Levin [13]. Schröder, Steinberg and Ziegler [23] recently extended Levin’s notion to real number computations:

► **Definition 1.** Let (X, Σ, μ) denote a probability space and $T : X \times \mathbb{N} \rightarrow [0, \infty]$ a measurable function. We say that T is polynomial-time on average if there is some $\varepsilon > 0$ such that the function

$$n \mapsto \frac{1}{n} \int_X (T(x, n))^\varepsilon d\mu \quad (1)$$

is bounded by some constant $c > 0$.

If an algorithm runs in polynomial-time on average, there is a high probability that it runs in polynomial-time for a randomly selected input as by Markov’s inequality for any $k > 0$ it is

$$\Pr[T(x, n) \geq kn^{\frac{1}{\varepsilon}}] \leq \frac{c}{k^\varepsilon}.$$

In this paper we only consider the case where $X \subseteq \mathbb{R}^d$, Σ is the Borel algebra over \mathbb{R}^d and $\mu(A) = \frac{\lambda(A)}{\lambda(X)}$ where λ denotes the Lebesgue measure on \mathbb{R}^d .

3 Complexity of simulating dynamical systems

Dynamical systems are used to describe the evolution of a system over time. A dynamical system on the reals can be formally defined as follows:

► **Definition 2.** A dynamical system is a triple (X, T, Φ) of a non-empty set $X \subset \mathbb{R}^d$ called the *phase space*, a time set $T \subseteq \mathbb{R}$ and a (partial) function $\Phi : \subseteq X \times T \rightarrow X$ called *evolution operator* satisfying

1. $\Phi(x, 0) = x$, and
 2. $\Phi(\Phi(x, t_1), t_2) = \Phi(x, t_1 + t_2)$
- for $x \in X$ and all $t_1, t_2 \in T$ such that $(x, t_1), (\Phi(x, t_1), t_2) \in \text{dom } \Phi$.

A point $x \in X$ is also called a *state* of the system and $\Phi(x_0, t)$ the state at time t (w.r.t. the initial value x_0). For a fixed initial value x_0 the function $\Phi(x_0, \cdot)$ is called *trajectory* through x_0 . In this paper we only consider continuous-time systems where the time set is some real interval and the evolution operator is continuously differentiable with respect to time. The dynamics of such a system can be described by the solution of a system of autonomous first-order ODEs

$$\dot{y} = f(y), \quad y(t_0) = y_0 \tag{2}$$

for some function $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $y_0 \in \mathbb{R}^d$.

We call Equation (2) an *initial value problem* (IVP) with initial value y_0 . The corresponding solution function $y : \mathbb{R} \rightarrow \mathbb{R}^d$ of the IVP gives a trajectory of the system. We sometimes also write $y(t; t_0, y_0)$ for the solution function with initial value $y(t_0) = y_0$ to make the dependency on the initial value explicit.

In the following, we use multi-index notation for tuples of non-negative integers $\beta = (\beta_1, \dots, \beta_d)$, i.e., for $\beta \in \mathbb{N}^d$ and $x \in \mathbb{R}^d$ it is $|\beta| = \beta_1 + \dots + \beta_d$, $\beta! = \beta_1! \dots \beta_d!$, $x^\beta = x_1^{\beta_1} \dots x_d^{\beta_d}$ and $D^\beta f = D_1^{\beta_1} \dots D_d^{\beta_d} f$.

A function $f : U \rightarrow \mathbb{R}$ for $U \subseteq \mathbb{R}^d$ open is called *analytic* if for each $x_0 \in U$ there is an open neighborhood $V \subseteq U$ of x_0 such that for all $x \in V$ the power series expansion

$$\sum_{\beta \in \mathbb{N}^d} a_\beta (x - x_0)^\beta$$

converges to $f(x)$. A function $f : D \rightarrow \mathbb{R}$ on a (possibly) non-open domain D is called analytic if it can be extended to an analytic function on some open domain $U \supseteq D$.

We further say a function $f : D \rightarrow \mathbb{R}^m$, $f(x) = (f_1(x), \dots, f_m(x))$ is analytic if each of the component functions $f_1, \dots, f_d : D \rightarrow \mathbb{R}$ are analytic. Any function analytic on some subset $D \subseteq \mathbb{R}^d$ can be extended to a complex analytic function on an open set $G \subseteq \mathbb{C}^d$. In this paper we only consider IVPs where the right-hand side function f is analytic. By the Cauchy-Kowalevski theorem the solution function y of such an IVP is again analytic.

3.1 Parameterized complexity for analytic initial value problems

Most results on IVPs in computable analysis assume the initial value to be fixed. In this work, however, we want to consider the average complexity over all initial values. We therefore first need to formalize how exactly the complexity of the solution depends on the chosen initial value. More formally, assume that we allow initial values from some set $A \subseteq \mathbb{R}^d$. We want to give a complexity bound for the solution function $Y : \subseteq A \times [0, 1] \rightarrow \mathbb{R}^d$, $(y_0, t) \mapsto y(t; 0, y_0)$ that maps initial values at time 0 and time $t \in [0, 1]$ to the solution of the initial value

problem at time t . Note that the restriction that the time is between 0 and 1 is somehow arbitrary. However, for the average-case analysis in the following chapter it is reasonable to assume that the time is bounded as there is no natural choice for a distribution on time.

The domain of Y are tuples $(y_0, t) \in A \times [0, 1]$ where the solution with initial value y_0 exists up to time t . As this domain is not necessarily compact, we can not expect to find a simple complexity bound on the whole domain. However, we can subdivide the domain into subsets depending on some natural parameter of the system and give a complexity not only in terms of the output precision n but also in terms of an additional integer parameter. That is, we want to say that there is a uniform algorithm that computes the function on its whole domain and if we restrict the domain to some subset this algorithm runs in time polynomial in the output precision n and some parameter depending on the subset. We formalize this in the following definition.

► **Definition 3.** Let $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^e$ be a computable real function and let $D = \bigcup_{i \in I} D_i$ be some (fixed) partition of its domain with index set I . We say f is C -polynomial-time computable for a function $C : I \rightarrow \mathbb{N}$ if there is an oracle machine M that

1. computes f on all inputs $x \in D$, and
2. there is a polynomial $p : \mathbb{N} \rightarrow \mathbb{N}$ such that on inputs $x \in D_i$ the machine outputs a 2^{-n} approximation of $f(x)$ after at most $p(C(i) + n)$ steps.

We will usually partition the domain in terms of some positive real parameter α and use the short-hand notation “ f is α^{-1} -polynomial-time computable on D_α ”. Formally, this means that f is C -polynomial-time computable w.r.t. the partition of the domain $D = \bigcup_{\alpha > 0} D_\alpha$ and the function $C : \mathbb{R}^+ \rightarrow \mathbb{N}$, $C(\alpha) = \lceil \alpha^{-1} \rceil$.

Assume we are given an initial value problem (2) and the right-hand side function f is analytic on some compact $K \subseteq \text{dom } f$. Then there is some integer C_K such that the bound

$$|D^\beta f(x)| \leq C_K^{|\beta|+1} \beta! \quad (3)$$

holds for all $\beta \in \mathbb{N}^d$ and $x \in K$ [10]. On the other hand, if for some $y_0 \in \text{dom } f$ the solution exists up to time T then $y([0, T])$ is a compact subset of the domain. Thus, restricting f to this set suffices to compute the solution.

We usually can not assume that f is polynomial-time computable on its whole domain since its values can be unbounded. On the other hand on any compact subset $K \subseteq \text{dom } f$ of the domain, C_K is an upper bound for the values of f . It is therefore a reasonable assumption that f is C -polynomial-time computable where C denotes a function that assigns each compact subset $K \subseteq \text{dom } f$ an integer C_K as in Equation (3).

For simplicity let us from now assume that $T = 1$, i.e., we only consider initial values for which the solution exists up to time 1. The following theorem characterizes the complexity of the solution in terms of the parameterization given by C .

► **Theorem 4.** Assume $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$ is analytic and computable. Let $D_0 \subseteq D$ be some subset of initial values $y_0 \in D$ such that the solution y to the IVP (2) with initial value y_0 exists for all $t \in [0, 1]$. For each $y_0 \in D_0$ let further $K(y_0) := y([0, 1]; 0, y_0)$ and $D' := \bigcup_{y_0 \in D_0} K(y_0)$. Assume there is a function $C : D_0 \rightarrow \mathbb{N}$ such that

- f restricted to D' is C -polynomial-time computable, and
- $C(y_0)$ is a derivative bound for f on $K(y_0)$ as in (3).

Then there exists a C -polynomial-time computable function $Z : \mathbb{R}^d \times [0, 1] \times \mathbb{N} \rightarrow \mathbb{R}^d$ such that $Z(y_0, t, C(y_0)) = Y(y_0, t)$ for all $y_0 \in D_0$ and $t \in [0, 1]$. That is, Z computes Y when given $C(y_0)$ as additional information.

Note that in general it is not possible to effectively get the parameter $C(y_0)$ from the function [2, 20]. Thus, in the above theorem we have to provide it as an additional input. For most natural systems, however, it is usually easy to get an upper bound for the parameter. We therefore assume that we can effectively get the parameter for the rest of the paper.

The main idea of the proof is that a simple power series based approach suffices to compute a local solution on some small time interval $[t_0, t_0 + \delta]$ in time polynomial in $n + C(y_0)$ (see e.g. [17]). To get a solution on a bigger interval this algorithm can be iterated several times. To show theorem 4 it therefore suffices to show that polynomially in $C(y_0)$ many iterations suffice and that it suffices to compute the intermediate values in each iteration with polynomial precision. The proof of the theorem is very similar to the proof of a recent result by Bournez, Graça and Pouly [1] for polynomial ODEs over unbounded time. We therefore chose to omit the details here. However, as both the statement and notation of our theorem are quite different from theirs, we included a proof in the appendix for completeness.

4 Average-case complexity for dynamical systems

While theorem 4 gives a very general characterization for the complexity of simulating initial value problems, it is usually not very useful as the complexity bound can differ for each initial value y_0 . In general there is no way around this as the trajectories for distinct initial values can be quite different. On the other hand, it might be the case that most trajectories behave nicely in the sense that they stay far away from singularities of the system. In that case, we would like to say that the simulation can “usually” be done efficiently. This notion can be made formal using average-case complexity.

Such an average-case analysis, however, requires that we can get a bound on the probability for a trajectory to come close to a singularity. Usually, there is no easy way to assign such a probability. On the other hand, if the system has the special property that the volume of subsets of phase-space is preserved over time then a small volume of singularities in phase space indicates that the set of initial values coming close to singularities is also small. A large class of dynamical systems that have this property are *Hamiltonian systems*.

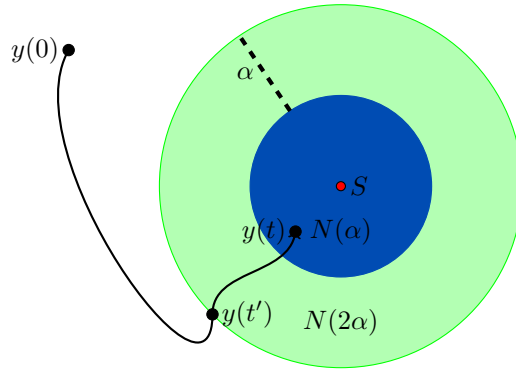
A d degree-of-freedom (time-independent) Hamiltonian system is a $2d$ -dimensional dynamical system $\dot{y} = f(y)$ where the state can be split into two variables $y(t) = (q(t), p(t))$ for smooth functions $q, p : \mathbb{R} \rightarrow \mathbb{R}^d$ that satisfy the system of $2d$ first order ordinary differential equations

$$\dot{q}(t) = \frac{\partial H(p, q)}{\partial p}, \quad \dot{p}(t) = -\frac{\partial H(p, q)}{\partial q} \quad (4)$$

for some smooth, real-valued function $H : \mathbb{R}^{2d} \rightarrow \mathbb{R}$. The functions q and p are called the *position* and *momentum* of the system and H is called the *Hamiltonian*. We only consider time-independent systems where the Hamiltonian is constant over time and therefore is sometimes also called the *energy* of the system.

4.1 Average-case complexity for Hamiltonian system

In this chapter we define some criteria under which a given Hamiltonian system can be simulated in polynomial time on average. Here, it is more natural to formalize the results in terms of distance to complex singularities of a complex analytic extension instead of a derivative bound. We first (independently of the system being Hamiltonian or not) formalize the notion of a point being close to a (complex) singularity of an analytic function.



■ **Figure 1** An α -singularity is a state of the system where the distance to a singularity is at most α . If the distance at time $t = 0$ is at least 2α and there is an α -singularity at time $t > 0$ there has to be some time t' when $y(t')$ is 2α close to the singularity. While the distance is between 2α and α the velocity is bounded which can be used to bound the minimum time the particle has to spend in the green region.

► **Definition 5.** Let $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$ analytic. For any $\alpha > 0$ and $A \subseteq D$ we define the following sets:

1. $G(\alpha)$ is the set of points $x \in D$ such that there is a complex analytic extension \tilde{f} of f that is analytic on all $z \in \mathbb{C}^d$ with $|x - z| \leq \alpha$,
2. $N(\alpha) := D \setminus G(\alpha)$,
3. $R_A \subseteq D$ is the set of points reachable from A in time $t \leq 1$, i.e., $x \in R_A$ if there is $y_0 \in A$ and $t \in [0, 1]$ such that $y(t; 0, y_0) = x$,
4. $G_A(\alpha) := G(\alpha) \cap R_A$ and $N_A(\alpha) := N(\alpha) \cap R_A$,
5. $A(\alpha) := \{y_0 \in A : y(t; 0, y_0) \in G(\alpha) \text{ for all } t \in [0, 1]\}$,
6. $B(\alpha) := A \setminus A(\alpha)$.

We call a point $x \in N(\alpha)$ an α -singularity, $x \in A(\alpha)$ an α -good initial value (α -bad for $x \in B(\alpha)$) and points in R_A reachable points.

If we restrict the initial values to be contained in a set A and the volume of $B(\alpha)$ is small in comparison to A and if the growth of f behaves “nicely” on $G(\alpha)$ then the simulation can be done in polynomial time on average:

► **Theorem 6.** Let $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$ analytic and computable. Consider the solution function Y to the IVP (2) restricted to a bounded subset of initial values $A \subseteq D$. Assume that there is some polynomial $m : \mathbb{N} \rightarrow \mathbb{N}$ such that $|\tilde{f}| \leq m(\alpha^{-1})$ holds for any point in $G_A(\frac{\alpha}{2})$ and that f on $G_A(\alpha)$ is α^{-1} -polynomial-time computable. Then

1. Y is α^{-1} -polynomial-time computable on initial values in $A(\alpha)$, and
2. If there is additionally some constant $\gamma > 0$ such that $\frac{\lambda(B(\alpha))}{\lambda(A)} \leq \alpha^\gamma$ then Y can be computed in polynomial time on average (w.r.t. the restriction of d -dimensional Lebesgue measure to A).

Proof. We show how to apply theorem 4 to prove the first part of the theorem. Let us first show that the polynomial m can be used to get a derivative bound on $G(\alpha)$.

Assume $|\tilde{f}(z)| \leq M(\alpha)$ for all $z \in G(\frac{\alpha}{2})$ for a function M with $M(\alpha) \geq \frac{2}{\alpha}$. For $z_0 \in G(\alpha)$, consider the polydisc $D = \prod_{i=1}^d D_i$ with $D_i := \{z \in \mathbb{C} : |z_{0,i} - z| \leq \frac{\alpha}{2}\}$. Now, \tilde{f} is analytic

on D by the definition of $G(\alpha)$. Furthermore $|\tilde{f}|$ is bounded on D by $M(\alpha)$. By Cauchy's integral formula it follows

$$\begin{aligned} |D^\beta f(z_0)| &= \left| \frac{\beta!}{(2\pi i)^d} \int_{\xi_1 \in \partial D_1} \cdots \int_{\xi_d \in \partial D_d} \frac{f(\xi_1, \dots, \xi_d)}{(\xi_1 - z_{0,1})^{\beta_1+1} \cdots (\xi_d - z_{0,d})^{\beta_d+1}} d\xi_1 \cdots d\xi_d \right| \\ &\leq \frac{\beta!}{(2\pi)^d} \left(2\pi \frac{\alpha}{2}\right)^d \frac{M(\alpha)}{\left(\frac{\alpha}{2}\right)^{\beta+d}} = \beta! M(\alpha) \left(\frac{2}{\alpha}\right)^\beta. \end{aligned}$$

Therefore, $|D^\beta f(z_0)| \leq \beta! M(\alpha)^{|\beta|+1}$ for all $z_0 \in G(\alpha)$. As we assume that the set of initial values is bounded and $t \in [0, 1]$, it further follows that the absolute value of points in $G_A(\alpha)$ is bounded by a linear function in $M(\alpha)$. Thus we can apply theorem 4 using the partition of the domain into sets $K_\alpha = G_A(\alpha)$ which concludes the proof of the first part of the theorem.

For the second part let μ be the restriction of the Lebesgue measure to A , i.e., $\mu(B) = \frac{\lambda(B)}{\lambda(A)}$ for all measurable subsets $B \subseteq A$. Let further $A_i = A(2^{-(i+1)}) \setminus A(2^{-i})$, i.e., A_i contains the initial values where the minimum distance to a complex singularity for any $t \in [0, 1]$ is between 2^{-i} and $2^{-(i+1)}$. Since $A_i \subseteq B(2^{-i})$ it holds $\lambda(A_i) \leq \lambda(B(2^{-i}))$ and by assumption $\mu(A_i) \leq 2^{-i\gamma}$. On the other hand A_i is contained in $A(2^{-(i+1)})$ thus by the first part of the theorem and the assumption on the time bound of f it follows that Y is computable on initial values in A_i in time $u(n + 2^i)$ for a polynomial u . Since $A = \bigcup_{i=0}^\infty A_i$ it holds

$$\int_A \frac{1}{n} T(x, n)^\varepsilon d\mu \leq \frac{1}{n} \sum_{i=0}^\infty \mu(A_i) T(A_i, n)^\varepsilon \leq \frac{1}{n} \sum_{i=0}^\infty 2^{-i\gamma} u(n + 2^i)^\varepsilon$$

Let m be the highest coefficient in the polynomial u then for $\varepsilon \leq \frac{\gamma}{2m}$ the sum converges and thus $\int_A \frac{1}{n} T(x, n)^\varepsilon d\mu$ is bounded. ◀

Theorem 6 holds even if the system is not Hamiltonian, but usually there is no simple way to get a bound for the measure for the second part of the theorem. For Hamiltonian systems, on the other hand, we can exploit the fact that they preserve phase-space volume over time (this is known as Liouville's theorem). Thus, there is a relation between the volume $\lambda(A(\alpha))$ of initial values leading to almost singularities and the volume $\lambda(N_A(\alpha))$ of almost singularities in phase-space.

Saari uses a similar idea to show that for the three-body problem the set of initial values leading to collisions has measure 0. However, for our application we need a stronger quantification of how the measure correlates to the distance of the particles. The main difficulty is that almost singularities can occur at any time $t \in [0, 1]$. Thus, theoretically we would have to consider infinitely many "copies" of $N_A(\alpha)$, one for each possible time. Saari manages to replace this by only countably infinite many copies which suffices to show that the set of initial values leading to collisions has measure zero. However, this approach does not suffice to give a bound for α -singularities with $\alpha > 0$. We therefore need a slightly more complicated idea to show that finitely many copies suffice in our case. Note, however, that Saari's result holds for unbounded time while our idea is based on the time being bounded. In fact, a generalization of our result to unbounded time turns out to be false [27].

Let us now state our main theorem. As the property of preserving phase-space volume is in fact the only attribute we use about Hamiltonian systems, we can formulate the theorem in terms of the bigger class of volume-preserving (sometimes also called conservative) dynamical systems.

- **Theorem 7.** Let $\dot{y} = f(y)$ be a volume-preserving dynamical system with $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$ computable and analytic and let $A \subseteq D$ be a bounded subset of initial values. Assume that
- f is α^{-1} -polynomial-time computable on α -good initial values $y_0 \in G_A(\alpha)$,
 - for all $\alpha > 0$ the measure of α -bad subsets of phase space is bounded by $\lambda(N_A(\alpha)) \leq \alpha^\gamma$ for some $\gamma > 0$,
 - $|\tilde{f}|$ is bounded by $M_\alpha := M(\lceil \alpha^{-1} \rceil)$ for a polynomial $M : \mathbb{N} \rightarrow \mathbb{N}$ on $N(2\alpha) \setminus N(\alpha)$.

Then

1. $\lambda(B(\alpha)) \leq \left(\frac{2M_\alpha}{\alpha} + 1\right)\lambda(N_A(2\alpha))$
2. The solution function $Y : \subseteq A \times [0, 1] \rightarrow \mathbb{R}^d$ is polynomial-time computable on average.

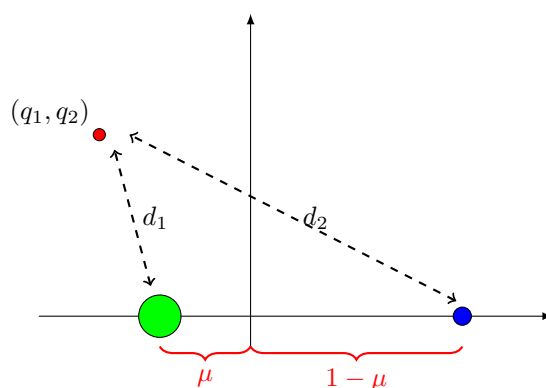
Proof. We only prove the first part. The second part then follows by applying theorem 6. For any $t \in [0, 1]$ let $B_t(\alpha) \subseteq A$ be the subset of initial values that lead to an α -singularity at time t , i.e., $y_0 \in B_t(\alpha)$ iff $y(t; 0, y_0) \in N(\alpha)$. As the system is volume-preserving it holds $\lambda(B_t(\alpha)) \leq \lambda(N_A(\alpha))$ for all t and α . Obviously, $B(\alpha) = \bigcup_{t \in [0, 1]} B_t(\alpha)$.

We now show that we can replace the infinite union by a union of finitely many slightly bigger sets. Let $y(0) \in B_t(\alpha)$ for some $t \in [0, 1]$, i.e., $y(0)$ is an initial value such that $y(t) \in N(\alpha)$. If $y(0) \in N(2\alpha)$ then it holds that $y(0) \in B_0(2\alpha)$ by definition. Otherwise there has to be some time $t' > 0$ where $y(t') \in N(2\alpha) \setminus N(\alpha)$ for the first time (see Fig. 1). As for any $z \in N(2\alpha) \setminus N(\alpha)$ by assumption $|\tilde{f}(z)| \leq M_\alpha$. It follows that for any $s \in [t', t' + \frac{\alpha}{M_\alpha}]$ the difference of the state at time t' and the state at time s can be at most α , i.e., $y(s) \in G(\alpha)$. In particular, if $y(0) \in G(2\alpha)$ and for some $t \in [0, 1]$ it holds that $y(t) \in N(\alpha)$ then there is some non-trivial time interval $[t_1, t_2]$ of length at least $\frac{\alpha}{M_\alpha}$ such that $y(s) \in N(2\alpha) \setminus N(\alpha)$ for all $s \in [t_1, t_2]$. Therefore, $[t_1, t_2]$ contains some multiple t^* of $\frac{\alpha}{2M_\alpha}$. In particular, $y(t^*) \in N(2\alpha)$ and thus by definition $y(0) \in B_{t^*}(2\alpha)$. This shows that for the finite subset $\{t_k\}_k \subset [0, 1]$ of multiples of $\frac{\alpha}{2M_\alpha}$ it holds that $B(\alpha) \subseteq \bigcup_{t_k} B_{t_k}(2\alpha)$. By applying the fact that the system is volume-preserving it holds $\lambda(B_{t_k}(2\alpha)) \leq \lambda(N_A(2\alpha))$ and thus the statement follows. ◀

4.2 Average-case complexity for the restricted three-body problem

As an application of theorem 7 we show that the solution of the restricted three-body problem can be computed in polynomial-time on average. The classical three-body problem is the problem of predicting the motion of three point masses under their mutual gravitational attraction. An important special case is that we assume that one particle P_3 has much smaller mass than the other two P_1 and P_2 . In that case P_3 does not influence the motion of P_1 and P_2 significantly. The problem can therefore be simplified by assuming that P_3 is massless. The motion of the heavy particles can then be seen as a two-body problem. A further simplification can be achieved by assuming that the heavy particles move on a circular orbit around their common center of mass and P_3 only moves in the plane defined by P_1 and P_2 . This problem is known as the *planar circular restricted three-body problem*. In spite of being much simpler than the general problem, the restricted problem shares many of the properties of the N -body problem that makes it interesting for our analysis. In particular, the restricted problem is a Hamiltonian dynamical system where the equation of motion is given by an analytic initial value problem. There is no general closed form solution and the motion can be chaotic [15]. The restricted three-body problem has been studied extensively by mathematicians and engineers.

By choosing appropriate units of measurement and a rotating coordinate system, the system can be brought in a simpler, normalized form only depending on a single parameter $\mu \in (0, 0.5]$. The masses of the particles P_1 and P_2 in the new units are given by μ and $1 - \mu$, respectively. The position of the heavy particles in the rotating coordinate system remains



■ **Figure 2** The planar circular restricted three-body problem in normalized form. P_1 and P_2 are fixed at position $(-\mu, 0)$ and $(1 - \mu, 0)$. They have masses $1 - \mu$ and μ and influence the motion of the massless particle P_3 at position (q_1, q_2) .

fixed at $(-\mu, 0)$ and $(1 - \mu, 0)$. We use $q = (q_1, q_2)$ to describe the coordinates of P_3 relative to that coordinate system (see Figure 2). A full derivation for the transformations as well as formulas to translate a solution for the normalized system to a solution for the non-modified system can e.g. be found in [11].

In Hamiltonian form the IVP can be written in terms of position $q \in \mathbb{R}^2$ and moment $p \in \mathbb{R}^2$ as

$$H(p, q) = \frac{1}{2} \|p\|^2 + q_2 p_1 - q_1 p_2 - \frac{\mu}{d_1} - \frac{1 - \mu}{d_2} \quad (5)$$

where $d_1 := \sqrt{(q_1 + \mu)^2 + q_2^2}$ and $d_2 := \sqrt{(q_1 + \mu - 1)^2 + q_2^2}$. This defines a dynamical system with phase space $\Gamma \subseteq \mathbb{R}^4$. We sometimes also consider the velocity of the particle given by $v(t) = (p_1(t) + q_2(t), p_2(t) - q_1(t))$ instead of the moment.

We assume that we start the simulation with initial values in the set A of points q_0, p_0 with $|q_0| \leq 1$ and $|p_0| \leq 1$. We first make the additional assumption that the energy is bounded by some constant $h > 0$, i.e., it holds $|H(q_0, p_0)| \leq h$. We denote the subset of initial values satisfying this bound by A_h .

In this problem, an α -singularity corresponds to the situation where P_3 gets close to either P_1 or P_2 . Note that the absolute value of the moments tend to infinity when approaching a singularity. A bound on the volume of α -singularities in phase-space is given by the following lemma.

► **Lemma 8.** *For any $\alpha \in [0, 0.5]$, the Lebesgue measure of $N_{A_h}(\alpha)$ is given by $\lambda(N_{A_h}(\alpha)) \leq 8\pi^2 h \alpha^2$.*

Proof. We show that for the set $N_1(\alpha)$ of points coming close to P_1 it holds that $\lambda(N_1(\alpha)) \leq \pi^2 h \alpha^2$. Then the claim follows as $N_{A_h}(\alpha) = N_1(\alpha) \cup N_2(\alpha)$. We first change to a new coordinate system (x, v) in terms of position and velocity of the particle such that P_1 is at the origin (note that this change preserves volume). The Hamiltonian of the problem in these coordinates can be written as

$$E(x, v) = \frac{1}{2} \|v\|^2 - \frac{1}{2} \|x\|^2 + x_1 \mu - \frac{\mu}{\|x\|} - \frac{1 - \mu}{d_2} - 0.5 \mu^2$$

Let $\Gamma \subseteq \mathbb{R}^4$ the phase-space of the problem. Γ can be parameterized in terms of the position

30:12 Average-Case Polynomial-Time Computability of Hamiltonian Dynamics

and the energy E by

$$\Phi : (E, r, \varphi, \psi) \mapsto (r \cos(\varphi), r \sin(\varphi), R(E, r, \varphi) \cos(\psi), R(E, r, \varphi) \sin(\psi)) \quad (6)$$

with

$$R(E, r, \varphi) := \sqrt{2E + \mu^2 - 2\mu r \cos(\varphi) + \frac{2\mu}{r} + \frac{2 - 2\mu}{d_2} + r^2} \quad (7)$$

It is $N_1(\alpha) \subseteq \Phi(G)$ with $G := [-h, h] \times [0, \alpha] \times [0, 2\pi] \times [0, 2\pi]$. Thus the volume can be bounded by

$$\lambda(N_1(h, \alpha)) \leq \int_{\Phi(G)} dq_1 dq_2 dv_1 dv_2 = \int_G |\det(D\Phi)| dr dh d\varphi d\psi = 4\pi^2 \cdot \alpha^2 \cdot h.$$

The last equality holds since $\det(D\Phi) = r$. ◀

It further holds that a bound on the energy implies that both the position and the velocity of the particle are bounded as long as the particle does not get close to one of the singularities.

► **Lemma 9.** *For any $q, v \in G(\alpha) \cap A_h$ it holds $\|q\| \leq 2h + 10$ and $\|v\| \leq \sqrt{(2h + 11)^2 + \frac{1}{\alpha}}$.*

Proof. Assume $d_1 \geq \alpha$ and $d_2 \geq \alpha$. The Hamiltonian in terms of the velocity is given by

$$H(v, q) = \frac{1}{2} \|v\|^2 - \frac{\mu}{d_1} - \frac{1 - \mu}{d_2} - \frac{1}{2} \|q\|^2.$$

Thus the bound on the Hamiltonian implies the bound

$$\|v\|^2 \leq 2h + |q|^2 + \frac{2}{\alpha} \quad (8)$$

on the velocity.

Now assume $\|q\| \geq 2$ then both $d_1 \geq 1$ and $d_2 \geq 2$ and thus $\|v\|^2 \leq 2h + |q|^2 + 2$ holds. Let $h' = \max(h, 2)$ then $\|v\| \leq h' + \|q\|$. Now consider for $n \geq 1$ the subsets U_n where $n - 1 \leq \|q\| \leq n$. For $n \geq 3$, $(q, v) \in U_n$ implies that $\|v\| \leq h' + n$. In particular the minimum time to get from U_n to U_{n+1} is $t_n := \frac{1}{h' + n}$. As for the initial values $\|q\| \leq 2$ the minimum time needed to reach a state where $\|q\| \geq N$ is bounded by

$$T_N := \sum_{i=3}^N \frac{1}{h' + i} = \mathcal{H}_{N+h'+3} - \mathcal{H}_{h'+3}$$

where $\mathcal{H}_N := \sum_{k=1}^N \frac{1}{k}$ denotes the N -th *harmonic number*. By the well known bound

$$\gamma + \log(N) \leq \mathcal{H}_N \leq \gamma + \log(N + 1)$$

it holds

$$T_N \geq \log\left(\frac{N + h' + 3}{h' + 4}\right).$$

As the time is bounded by 1 it follows that $\|q\| \leq 2h' + 6 \leq 2h + 10$. Inserting this back into (8) yields

$$\|v\| \leq \sqrt{2h + (2h + 10)^2 + \frac{2}{\alpha}}$$

from which the claim follows. ◀

From this it can be easily followed that the right-hand side function of the ODE system is computable in time polynomial in $\lceil \alpha^{-1} \rceil$ and h and that a polynomial bound in those terms holds for complex analytic extensions of the function on $N(2\alpha) \setminus N(\alpha)$.

As for initial values it holds that the magnitude of both position and moment are bounded by 1, it follows from the Equation for the Hamiltonian (5) that high energy can only be due to particles being already close at time 0. In particular for α -good initial values the following bound on the Hamiltonian holds.

► **Lemma 10.** *Assume $\|q\| \leq 1$, $\|p\| \leq 1$ and $q, p \notin N(\alpha)$ then $H(p, q) \leq 3 + \frac{2}{\alpha}$.*

Thus the above polynomial bounds can all be stated only in terms of α . In particular all conditions in theorem 7 are satisfied and thus the problem is polynomial-time computable on average:

► **Theorem 11.** *Simulating the restricted three-body problem for time $t \leq 1$ for initial values p_0, q_0 with $|p_0| \leq 1$, $|q_0| \leq 1$ is possible in polynomial time on average.*

5 Conclusion

We applied a recent definition of average-case complexity in analysis to problems in classical physics. We gave some general conditions which show that a continuous-time system can be computed efficiently on average. For the important special case of Hamiltonian systems, we showed that a simpler approach based on the volume of almost singularities in phase-space usually suffices. We applied our theory to the planar circular restricted three-body problem and showed that it indeed can be computed in polynomial time on average. The same can easily be done for some other simple dynamical systems.

However, our theory does not easily generalize to some other more complicated systems like the classical N -body problem as bounding the volume of singularities in phase space is more complicated for these systems. While we think that most systems in nature can indeed be simulated efficiently and that a similar result holds for, e.g., the general N -body problem it is unlikely that our approach can be easily adapted to this case as for $N > 4$ it is not even known if the set of initial conditions leading to singularities has measure zero.

Nonetheless, we hope that we can at least extend our theory to the general three-body problem and some other interesting problems in future work.

References

- 1 Olivier Bournez, Daniel S. Graça, and Amaury Pouly. On the complexity of solving initial value problems. In *ISSAC 2012-Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*, pages 115–121. ACM, New York, 2012. doi:10.1145/2442829.2442849.
- 2 Daniel S Graça, Ning Zhong, and Jorge Buescu. Computability, noncomputability and undecidability of maximal intervals of IVPs. *Transactions of the American Mathematical Society*, 361(6):2913–2927, 2009.
- 3 A. Grzegorzcyk. On the definitions of computable real continuous functions. *Fund. Math.*, 44:61–71, 1957.
- 4 Akitoshi Kawamura. Lipschitz Continuous Ordinary Differential Equations are Polynomial-Space Complete. *Computational Complexity*, 19(2):305–332, 2010.
- 5 Akitoshi Kawamura and Stephen Cook. Complexity theory for operators in analysis. *ACM Transactions on Computation Theory*, 4(2):Article 5, 2012.

- 6 Ker-I Ko. On the computational complexity of ordinary differential equations. *Information and Control*, 58(1-3):157–194, 1983.
- 7 Ker-I Ko. *Complexity theory of real functions*. Progress in Theoretical Computer Science. Birkhäuser Boston Inc., Boston, MA, 1991. doi:10.1007/978-1-4684-6802-1.
- 8 Ker-I Ko and Harvey Friedman. Computational complexity of real functions. *Theoretical Computer Science*, 20(3):323–352, 1982. doi:10.1016/S0304-3975(82)80003-0.
- 9 Ker-I Ko and Harvey Friedman. Computing power series in polynomial time. *Advances in Applied Mathematics*, 9(1):40–50, 1988.
- 10 Hikosaburo Komatsu. A characterization of real analytic functions. *Proceedings of the Japan Academy*, 36(3):90–93, 1960.
- 11 Wang Sang Koon, Martin W. Lo, Jerrold E. Marsden, and Shane D. Ross. *Dynamical systems, the three-body problem and space mission design*. World Scientific, 2000.
- 12 Daniel Lacombe. Sur les possibilités d’extension de la notion de fonction récursive aux fonctions d’une ou plusieurs variables réelles. In *Le raisonnement en mathématiques et en sciences expérimentales*, Colloques Internationaux du Centre National de la Recherche Scientifique, LXX, pages 67–75. Editions du Centre National de la Recherche Scientifique, Paris, 1958.
- 13 Leonid A. Levin. Average case complete problems. *SIAM Journal on Computing*, 15(1):285–286, 1986.
- 14 X. Mao. *Stochastic Differential Equations and Their Applications*. Ellis Horwood series in mathematics and its applications. Horwood Pub., 1997.
- 15 Andrea Milani. Chaos in the three body problem. In *Predictability, stability, and chaos in N-body dynamical systems*, pages 11–33. Springer, 1991.
- 16 R. H. Miller. Numerical difficulties with the gravitational n-body problem. In Dale G. Bettis, editor, *Proceedings of the Conference on the Numerical Solution of Ordinary Differential Equations*, pages 260–275, Berlin, Heidelberg, 1974. Springer Berlin Heidelberg.
- 17 Bernd Moiske and Norbert Th. Müller. Solving initial value problems in polynomial time. In *Proc. 22 JAIIO - PANEL '93, Part 2*, pages 283–293, Buenos Aires, 1993. URL: <http://www.uni-trier.de/mueller>.
- 18 Norbert Th. Müller. Uniform Computational Complexity of Taylor Series. In *Proc. 14th International Colloquium on Automata, Languages, and Programming*, volume 267 of *LNCS*, pages 435–444. Springer, 1987.
- 19 Norbert Th. Müller. Polynomial time computation of Taylor series. *Proc. 22 JAIIO-PANEL'93, Part 2, Buenos Aires*, 259:281, 1993.
- 20 Norbert Th. Müller. Constructive Aspects of Analytic Functions. In *Proc. Workshop on Computability and Complexity in Analysis*, volume 190 of *InformatikBerichte*, pages 105–114. FernUniversität Hagen, 1995.
- 21 Donald G. Saari. Improbability of Collisions in Newtonian Gravitational Systems. II. *Transactions of the American Mathematical Society*, 181:351–368, 1973. doi:10.2307/1996638.
- 22 Donald G. Saari. A global existence theorem for the four-body problem of Newtonian mechanics. *Journal of Differential Equations*, 26(1):80–111, 1977. doi:10.1016/0022-0396(77)90100-0.
- 23 Matthias Schröder, Florian Steinberg, and Martin Ziegler. Average-Case Bit-Complexity Theory of Real Functions. In *Mathematical Aspects of Computer and Information Sciences*, pages 505–519. Springer, Cham, 2015. doi:10.1007/978-3-319-32859-1_43.
- 24 A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(1):230–265, 1936. doi:10.1112/plms/s2-42.1.230.
- 25 Klaus Weihrauch. *Computable Analysis*. Springer, Berlin/Heidelberg, 2000.

- 26 Andrew Chi-Chih Yao. Classical physics and the Church–Turing Thesis. *Journal of the ACM (JACM)*, 50(1):100–105, 2003.
- 27 Lei Zhao. Quasi-Periodic Almost-Collision Orbits in the Spatial Three-Body Problem. *Communications on Pure and Applied Mathematics*, 68(12):2144–2176, 2015.

A Proof of Theorem 4

The proof consists of two parts. We first show that the local solution on some small interval with length polynomial in $C(y_0)^{-2}$ can be computed in $C(y_0)$ -polynomial time. We then show that iterating the algorithm for the local solution until reaching time $t = 1$ can still be done within this complexity bound.

A.1 Computing a local solution

We formulate the complexity result for the local solution as the following Lemma.

► **Lemma 12.** *Let $D \subseteq \mathbb{R}^d$ analytic and computable. Then there is an algorithm that given an initial $y_0 \in \mathbb{R}^d$, time $t \in \mathbb{R}$ and an integer $C \in \mathbb{N}$ such that C is a derivative bound for f on the closed ball $\overline{B_C(y_0)}$ of radius $\frac{1}{C}$ around y_0 returns the solution $y(t)$ to the IVP (2) for any t with $|t - t_0| \leq \frac{1}{2(d+1)C^2}$. Assume f is C -polynomial-time computable on $\overline{B_C(y_0)}$ then so is y for all t that satisfy the above bound.*

Proof. Cauchy’s existence theorem guarantees that the solution exists and is analytic on this time-interval. The polynomial-time result has in principle already been shown, e.g., by Moiske and Müller [17]. The main idea is to first compute the coefficients of the power series of f around t_0 . Müller [19] showed that with the above bounds for any $\beta \in \mathbb{N}^d$ the coefficient a_β is computable in time polynomial in $n + C + |\beta|$. Further, if the power series of f around y_0 can be computed in time $T_f(n + |\beta|)$ then the k -th coefficient of the power series $(b_k)_{k \in \mathbb{N}}$ of any of the d components y_1, \dots, y_d of solution function y around t_0 can be computed from the power series of f in time $O(k^d T_f(n + k^d))$.

For any $i = 1, \dots, d$, the solution $y_i(t)$ for small enough t can then be approximated by summing up the truncated power series $\sum_{i=0}^N b_i(t - t_0)^i$ for some $N \in \mathbb{N}$. It can be shown that $|b_i| \leq (d + 1)^i C^{2i}$. Therefore for any t with $|t - t_0| \leq \frac{1}{2(d+1)C^2}$ it suffices to sum up $O(n)$ coefficients to make the truncation error less than $2^{-(n+1)}$. To additionally make the approximation error less than $2^{-(n+1)}$ and thereby the total error less than 2^{-n} it suffices to approximate each b_i with error less than $2^{-(2n+1)}$. ◀

A.2 Extending to a global solution

We now prove the theorem by iterating the local solution algorithm. Let us first summarize the assumptions:

1. The right-hand side function $f : D \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^d$ is analytic and computable,
2. The solution $y(t) := y(t; 0, y_0)$ exists for all time $t \in [0, 1]$,
3. Restricted to the set $K := y([0, 1]; 0, y_0)$ the integer C is a derivative-bound for f , i.e., $|D^\beta f(x)| \leq C^{|\beta|+1} \beta!$ for all $x \in K$,
4. The algorithm computing f gives a 2^{-n} approximation of $f(x)$ in time $\text{poly}(n + C)$ on any $x \in K$.

Note that each $x \in K$ is bounded by $|x| \leq C$ as it is in the image of y .

We want to show that iteratively using the local solution algorithm yields a polynomial-time algorithm mapping initial values $y_0 \in K$ and time $t \in [0, 1]$ to the solution $y(t)$ at time t .

Algorithm 1 Solving Initial Value Problems.

```

function SOLVE_IVP( $f, y_0, t, n, C$ )
     $t_{curr} \leftarrow 0$ 
     $y_{curr} \leftarrow \text{APPROX}(y_0, m)$ 
     $h \leftarrow \frac{1}{2(d+1)C^2}$ 

    while  $t_{curr} + h \leq t$  do
         $y_{curr} \leftarrow \text{LOCALSOLUTION}(y_{curr}, h, m, C)$ 
         $t_{curr} \leftarrow t_{curr} + h$ 

    return LOCALSOLUTION( $f, y_{curr}, t - t_{curr}, m, C$ )
    
```

For this it suffices to show that we can always reach time 1 in polynomially many iterations and that it suffices to approximate all intermediate values with polynomial precision. For simplicity we fix the index i of the solution function y_i and simply denote it by y . By lemma 12 we can assume that there is an algorithm `LocalSolution` that computes for inputs $y_0 \in K$, $t \in [0, 1]$ and $C, n \in \mathbb{N}$ with $t \leq \frac{1}{2(d+1)C^2}$ a rational approximation $q \in \mathbb{Q}$ such that $|y(t; t_0, y_0) - q| \leq 2^{-n}$. For a suitable choice of the intermediate precision parameter m Algorithm 1 computes an approximation q to the solution at any time $t \in [0, 1]$ such that $|y(t; t_0, y_0) - q| \leq 2^{-n}$. It remains to show that m can be chosen to be polynomial in $n + C$. The algorithm makes at most $2(d+1)C^2$ steps to reach any time $t \in [0, 1]$.

Let us first consider the error when computing the local solution y_{i+1} from y_i . There are two types of errors. The first one is due to the local solution algorithm only returning an approximation to the real solution with precision 2^{-m} . The second kind of error arises because of the accumulated error in y_i : Instead of solving the IVP problem with initial value y_i we use an approximation z_i of y_i as initial value. To bound the second kind of error we use the following fact:

► **Lemma 13.** *Assume $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ fulfills the conditions above. Then each component function $f_i : \mathbb{R}^d \rightarrow \mathbb{R}$ is Lipschitz-continuous on K with Lipschitz-constant $L = \sqrt{d}C^2$.*

Proof. This simply follows from the fact that all partial derivatives of f are bounded by C^2 . ◀

This can be used to get a bound on the local error.

► **Lemma 14.** *Assume f fulfills the conditions above. Let $y_0, z_0 \in K$ be initial condition with $\|y_0 - z_0\|_\infty \leq \varepsilon$. Then for all $t < \frac{1}{2(d+1)C^2}$ it is $|y(t; y_0) - y(t; z_0)| \leq 2\varepsilon$.*

Proof. It holds that $y(t; y_0) = y_0 + \int_0^t |f(y(\tau))|$. Thus

$$\begin{aligned}
 & |y(t; y_0) - y(t; z_0)| \\
 & \leq \|y_0 - z_0\|_\infty + \int_0^t |f(y(\tau; y_0)) - f(y(\tau; z_0))| d\tau \\
 & \leq \varepsilon + \int_{t_0}^t L |y(\tau; y_0) - y(\tau; z_0)| d\tau
 \end{aligned}$$

By Grönwall's Lemma [14, Chapter 1, Theorem 8.1] it follows that

$$|y(t; y_0) - y(t; z_0)| \leq \varepsilon e^{Lt}.$$

For $t \leq \frac{1}{2(d+1)C^2} \leq \frac{\sqrt{d}}{2(d+1)L}$ it then holds

$$|y(t; y_0) - y(t; z_0)| \leq \varepsilon e^{\frac{\sqrt{d}}{2(d+1)}} \leq 2\varepsilon. \quad \blacktriangleleft$$

It is now easy to compute the total error for the algorithm depending on the parameter m for the intermediate precision.

► **Lemma 15.** *Assume algorithm 1 needs N iterations to reach the final time t . The total error E is bounded by $E \leq 2^{N+1-m}$.*

Proof. Let E_k be the error of y_{curr} after k iterations of the while-loop in Algorithm 1. It is $E_0 \leq 2^{-m}$ and $E_{k+1} \leq 2E_k + 2^{-m}$. It follows by induction that $E_N \leq 2^{N+1-m} - 2^{-m}$. ◀

Since the maximum number of iterations is bounded by $N = 2(d+1)C^2$ to achieve precision 2^{-n} it suffices to choose $m \geq n + 2(d+1)C^2 + 1$ which proves the theorem.