Keijo Ruohonen

Undecidable event detection problems for ODEs of dimension one and two


<http://www.numdam.org/item?id=ITA_1997__31_1_67_0>
UNDECIDABLE EVENT DETECTION PROBLEMS FOR ODES OF DIMENSION ONE AND TWO (*)

Keijo Ruohonén (1)

Communicated by Xaffray

Abstract. – The ability of dynamical systems of various kinds to simulate Turing machines and thus manifest a universal computation power (and beyond) has gathered a lot of interest lately, see e.g. [16], [5], [6] and [4]. A similar line of investigation for ordinary differential equations was started in [11] and continued in [12] and [13]. In this context the minimum dimension required for universal computation is of interest. The dynamical systems in [5] and [6] are of small dimension and the topic of [4] is to find the smallest dimension for certain types of dynamical systems. The results in this paper show that for ODEs dimension two can be reached and, allowing somewhat complicated events, even dimension one.

1. INTRODUCTION

Many problems involving finitely given ordinary differential equations (ODEs) turn out to be algorithmically undecidable, something that probably has not been sufficiently appreciated. It is true that the classical noncomputability result of M.B. Pour-El and I. Richards [8] involves nonunique solutions and that these are not of much practical interest. There are however other undecidable problems, the event detection problems, which involve unique solutions and explicitly given ODEs. Indeed, event detection

(*) Received September 1995.
(1) Department of Mathematics, Tampere University of Technology, 33101 Tampère, Finland.

Informatique théorique et Applications/Theoretical Informatics and Applications
0988-3754/97/01/$ 7.00/© AFCET-Gauthier-Villars
K. RUOHONEN

is dynamically undecidable, i.e., the ODE simulates dynamically the steps of a Turing machine computation and its definition contains only the transition rules of the Turing machine and not results of whole computations, see [11].

The purpose of this paper is to extend the dynamical undecidability of event detection to ODEs of small dimension (that is, small number of dependent variables). This objective is of interest because the dimensions of the ODEs in [11],[12] and [13] are rather high. Although it is possible to get lower dimensions by using different types of machines in the simulation, to reach dimensions one and two requires separate constructs. These constructs are the subject of the present paper.

There has also been a recent interest in Turing machine simulation by low-dimensional dynamical systems, and the present paper may be considered as a contribution to this line of research, see e.g. [4],[5] and [6]. The conclusions reached here are similar to those in [4]: A rather natural two-dimensional dynamically undecidable event detection problem exists but to get to dimension one a much more complicated construct is needed. Somehow the smallest natural dimension for dynamical computation appears to be two, getting to dimension one strains naturality a lot.

Only some basic facts of computability, computable analysis and classical ODE theory are used. A good background is contained in [3],[9] and in [2] or [10].

2. PRELIMINAIRES

An event of an ODE $y' = f(y, t)$, with the initial value $y(0) = y_0$, occurs whenever at least one of the given equations

$$g_j(t, y(t), y'(t)) = 0 \ (j = 1, \ldots, k)$$

is satisfied for some $t$ in a given interval $I$. For aspects of numerical event detection see [15]. The event detection problem (EDP) is the problem of deciding for a given initial value problem and event on an interval $I$ whether or not the event occurs.

If a quite general approach is taken then it is not very difficult to obtain low-dimensional undecidable event detection problems. Indeed, take a universal Turing machine $M$ with nonnegative integer inputs and define the sequence $f_0, f_1, \ldots$ of rationals by

$$f_n = \begin{cases} 2^{-m} & \text{if } M \text{ stops in } m \text{ steps after receiving input } n \\ 0 & \text{if } M \text{ does not stop after receiving input } n. \end{cases}$$
Following the nomenclature of [9], the sequence \( f_0, f_1, \ldots \) is then not a computable sequence of rationals but it is a computable sequence of reals as it can be approximated by the computable double sequence \( f_{nk} \) \( (n = 0, 1, \ldots; k = 0, 1, \ldots) \) of rationals where

\[
f_{nk} = \begin{cases} 
2^{-m} & \text{if } \mathcal{M} \text{ stops in } m \leq k \text{ steps after receiving input } n \\
0 & \text{if } \mathcal{M} \text{ does not stop in } k \text{ steps after receiving input } n.
\end{cases}
\]

(A sequence of rationals is computable if there is an algorithm which on input \( n \) (resp. \( (n, k) \)) computes the denominator, the numerator and the sign of the \( n \)th (resp. the \( (n, k) \)th) term in the sequence. A sequence \( x_0, x_1, \ldots \) of reals is computable if it can be approximated by a computable double sequence \( r_{nk} \) \( (n = 0, 1, \ldots; k = 0, 1, \ldots) \) of rationals uniformly in \( n \), i.e., \( |x_n - r_{nk}| \leq 2^{-k} \) for all \( n \) and \( k \).) Now, detection of the event \( y(t) = 0 \) for the ODE \( y' = 0 \), given \( n \) and the initial value \( y(0) = f_n \), is undecidable on any interval containing 0, because \( f_n = 0 \) is undecidable. A further modification is obtained as follows. Define the smooth function

\[
g(x) = f_{[x+1/2]}e^{-\tan^2 \pi x}.
\]

It is easy to see that \( g \) is computable on \([0, \infty)\). Detection of the event \( y_1(t) = 0 \) for the ODE

\[
\begin{cases}
y_1' = g(y_2) - 1 \\
y_2' = 0
\end{cases}
\]

given an initial value \( y_1(0) = 1, y_2(0) = n \) where \( n \) is a nonnegative integer is then undecidable on \([0, 1]\). There is a similar construct giving undecidability of two-dimensional symbolical event detection described in [11].

EDP is dynamically undecidable for time intervals of the form \([0, T) \) and \([0, \infty) \), as was shown in [11] through dynamical simulation of a universal Turing machine by an explicitly given ODE. Indeed, in [11] initial values \( y_0 \) are \( n \)-tuples of nonnegative integers, \( f \) is a fixed explicitly given function and the event to be detected is of the simple form \( y_i(t) = 0.5 \). (No references to computability of reals or functions, or properties of symbolical expressions actually appear in [11].) Moreover, the solutions are computable. Even a smooth choice for \( f \) is possible. Extensions of the undecidability to parametric ODEs and to closed finite time intervals are given in [12] and [13].

The ODE used in [11] has a large dimension (that is, number of dependent variables). Reduction of the number of dependent variables depends heavily
on the internal structure of the ODE. For this purpose some characteristics of the construct in [11] are given.

The central idea of [11] is to simulate a counter machine $\mathcal{M}$ with $m$ counters and counter input (and no internal states) by a $2m + 1$-dimensional autonomous ODE

$$\frac{dq}{dt} = Q(q(t))$$

with an initial value at $t = 0$. As is well known, counter machines have universal computing power. The following properties of this simulation will be needed:

(A) The simulation of the $i$-th step of the computation of $\mathcal{M}$ takes place in two stages, the first stage in the time interval $2i - 2 < t < 2i - 1$ and the second in $2i - 1 < t < 2i$.

(B) Two copies of counters of $\mathcal{M}$ are kept, the first in $q_1, \ldots, q_m$ and the second in $q_{m+1}, \ldots, q_{2m}$, giving the counts of symbols in the counters. The state $q_{2m+1}$ is the time $t$ (whence $Q_{2m+1} = 1$).

(C) During the first stage of simulation counter transition of $\mathcal{M}$ is performed on $q_1, \ldots, q_m$ using $q_{m+1}, \ldots, q_{2m+1}$, and $q_{m+1}, \ldots, q_{2m}$ will remain unchanged. During the second stage the states $q_{m+1}, \ldots, q_{2m}$ are updated using $q_1, \ldots, q_m, q_{2m+1}$, and $q_1, \ldots, q_m$ remain unchanged.

(D) $Q_j$ is of the form $Q_j(q) = P_j(q_{m+1}, \ldots, q_{2m})s(t)$ ($j = 1, \ldots, m$) where the value of $P_j$ is $-1$, $0$ or $1$,

$$\int_{2i-2}^{2i-1} s(t)dt = 1$$

and $s(t)$ is zero during the second stage. Denote for brevity

$$P = (P_1, \ldots, P_m).$$

$P$ and $s$ are smooth, and so is $q(t)$.

(E) If the input count of $\mathcal{M}$ is $n$ then the initial value is

$$e = (0, \ldots, 0, n, 0, \ldots, 0, n, 0)$$

where the $n$'s are at the $m$-th and the $2m$-th positions.

(F) When $\mathcal{M}$ halts at the $i$th step, then the value of $q_1$, which hitherto has been $0$, is raised to $1$ during $2i - 2 < t < 2i - 1$, and will stay there.
for $t \geq 2i - 1$. The ODE does not “halt.” Halting of $\mathcal{M}$ is thus signalled by the event $q_1(t) = 0.5$ which is undecidable.

Structural properties and explicit construction of $Q$ are given in [11] (Theorem 1 and its proof).

3. ODEs OF DIMENSION TWO

During the first stage the values of $q_{m+1}, \ldots, q_{2m}$ are nonnegative integers. These are coded in the value of the second dependent variable $z_2$. Similarly during the second stage the values of $q_1, \ldots, q_m$ are nonnegative integers and these are coded in the value of the first dependent variable $z_1$. In both cases the coding scheme is the same so only $z_2$ is treated in what follows. To describe the scheme take $m$ nonnegative integers $i_1, \ldots, i_m$. If $q_1 = i_1$ then the value of the variable $z_2$ is in the interval

$$\sum_{j=1}^{i_1} 2^{-j} \leq z_2 < \sum_{j=1}^{i_1+1} 2^{-j} \text{ i.e. } 1 - 2^{-i_1} \leq z_2 < 1 - 2^{-i_1-1}.$$  

Similarly, if in addition $q_2 = i_2$ then the value $z_2$ is in the interval

$$1 - 2^{-i_1} + 2^{-i_1-1}(1 - 2^{-i_2}) \leq z_2 < 1 - 2^{-i_1} + 2^{-i_1-1}(1 - 2^{-i_2-1}),$$

and, in general, if $q_1 = i_1, \ldots, q_l = i_l$ then

$$\sum_{j=1}^{i_1} 2^{-i_1-\cdots-i_j-j+1}(1 - 2^{-i_j}) \leq z_2 < \sum_{j=1}^{l-1} 2^{-i_1-\cdots-i_j-j+1}(1 - 2^{-i_j})$$

$$+ 2^{-i_1-\cdots-i_{l-1}-l+1}(1 - 2^{-i_{l-1}-1}) \ (l = 1, \ldots, m).$$

To put it in another way, if $q_1 = i_1, \ldots, q_l = i_l$ then the binary representation of $z_2$ is of the form

$$z_2 = 0.1\overbrace{\cdots101\cdots10\cdots01\cdots1}^{i_1 \text{ bits}} \overbrace{\cdots}^{i_2 \text{ bits}} \overbrace{\cdots}^{i_l \text{ bits}}.$$  

It is assumed that binary representations containing only finitely many 0’s are not allowed. Thus, in general, any number $z$ in the interval $[0,1)$ has the binary representation

$$z = \sum_{j=1}^{\infty} 2^{-i_1-\cdots-i_{j-1}-j+1}(1 - 2^{-i_j}) = 0.1\overbrace{\cdots101\cdots10\cdots01\cdots1}^{i_1 \text{ bits}} \overbrace{\cdots}^{i_2 \text{ bits}} \overbrace{\cdots}^{i_j \text{ bits}}$$

vol. 31, n° 1, 1997
corresponding to the infinite integer sequence \( i_1, i_2, \ldots \). Denote then

\[
g(z) = \lfloor -\log_2(1 - z) \rfloor,
\]

\[
k(z) = 2(1 - (1 - z)2^{g(z)})
\]

and

\[
g_j(z) = g(k^{j-1}(z)) \quad (j = 1, 2, \ldots)
\]

where \( k^j \) denotes \( j \)-fold composition power of \( k \). See Figures 1-4.
Lemma 1: \( i_j = g_j(z) \) (\( j = 1, 2, \ldots \))

Proof: The lemma is proved by induction on \( j \). The case \( j = 1 \) is immediate. To proceed by induction it is first shown that if the integer sequence corresponding to \( z \) is \( i_1, i_2, \ldots \) then the integer sequence corresponding to \( k(z) \) is \( i_2, i_3, \ldots \). Indeed if \( g(z) = i_1 \) and

\[
z = \sum_{j=1}^{\infty} 2^{-i_1-\cdots-i_{j-1}-j+1}(1 - 2^{-i_j}) = 0.1 \overbrace{01 \cdots 01}^{i_1 \text{ bits}} \overbrace{01 \cdots 01}^{i_2 \text{ bits}} \overbrace{01 \cdots 01}^{i_j \text{ bits}} \cdots \]

then

\[
k(z) = 2(1 - (1 - z)2^{i_1}) = 2^{i_1+1}(2^{-i_1} - 1 + z)
\]

\[
= \sum_{j=2}^{\infty} 2^{-i_2-\cdots-i_{j-1}-j+2}(1 - 2^{-i_j}) = 0.1 \overbrace{01 \cdots 01}^{i_2 \text{ bits}} \overbrace{01 \cdots 01}^{i_3 \text{ bits}} \overbrace{01 \cdots 01}^{i_j \text{ bits}} \cdots \cdots
\]

It is easily checked that if \( 0 \leq z < 1 \) then also \( 0 \leq k(z) < 1 \). Replacing \( z \) by \( k(z) \) repeatedly one sees inductively that the integer sequence corresponding to \( k^{j-1}(z) \) is \( i_j, i_{j+1}, \ldots \) whence \( g_j(z) = g(k^{j-1}(z)) = i_j \). □

It follows from Lemma 1 that \( g_j(z_2) = i_j \) (\( j = 1, \ldots, m \)) which is denoted simply by \( i = g(z_2) \). By \( h(i) \) a “typical” value of \( z_2 \) corresponding to the integers \( i_1, \ldots, i_m \) is denoted and it is chosen to be the midpoint of the particular interval, i.e.,

\[
h(i) = \sum_{j=1}^{m-1} 2^{-i_1-\cdots-i_{j-1}-j+1}(1 - 2^{-i_j}) + 2^{-i_1-\cdots-i_{m-1}-m}(2 - 3 \cdot 2^{-i_m-1}).
\]

Thus the binary representation of \( h(i) \) is

\[
h(i) = 0.1 \overbrace{101 \cdots 101}^{i_1 \text{ bits}} \overbrace{101 \cdots 101}^{i_2 \text{ bits}} \overbrace{101 \cdots 101}^{i_m \text{ bits}} \cdot
\]

and, by Lemma 1, \( g(h(i)) = i \). Now, to move the value of \( z_1 \) from one “typical” value to the next during first stage and using the value of \( z_2 \) one takes

\[
\frac{dz_2}{dt} = (h(g(z_2)) + P(g(z_2))) - h(g(z_2))s(t) \text{ and } z_1(0) = h(0, \ldots, 0, n).
\]
(The functions $s$ and $P$ appeared in property (D) in the previous section.) To update the value $z_2$ during second stage one then simply uses “dragging”:

$$\frac{dz_2}{dt} = \frac{3}{2} \sqrt{z_1 - z_2} s(t + 1) \text{ and } z_2(0) = h(0, \ldots, 0, n).$$

During the operation $z_2$ starts from some value within $z_1 \pm 1$ and within a unit interval of time is dragged to the value $z_1$ where it stays. There are of course several ways to get such a dragging operation.

As in the case of the original ODE, if a nonreversible counter machine is simulated, then the solutions are necessarily nonunique in the backward direction and this nonuniqueness appears during the dragging. Indeed, whenever a counter machine configuration has several possible predecessors, information of which one of them actually appeared will be lost after the dragging. However, if a reversible counter machine is simulated then the present configuration uniquely determines the previous one and dragging can be replaced by a reversible operation similar to the one for $z_1$ and a unique solution is obtained. It should be mentioned that reversible counter machines can simulate reversible Turing machines, see [11], and the latter are known to be computationally universal (see [7] or [11]). In any case, as in [11], the solutions $z_1(t)$ and $z_2(t)$ are smooth and forward unique.

The behaviour of the original ODE

$$\frac{d\mathbf{q}}{dt} = \mathbf{Q}(\mathbf{q}(t))$$

is now faithfully emulated and occurrence of the event $z_1 = 0.5$ is undecidable. Of course the infinite time interval $[0, \infty)$ can be replaced by a finite half open one, say the interval $[0, \pi/2)$ obtained via the change of variable $t = \tan u$. Thus the following theorem is proved.

**Theorem 1:** There exists an explicitly given ODE pair

$$\begin{align*}
y_1' &= f_1(y_1, y_2, t) \\
y_2' &= f_2(y_1, y_2, t)
\end{align*}$$

for which EDP is undecidable in the time interval $[0, \infty)$ (or in a finite half open time interval $[0, T]$). The solutions $y_1(t)$, $y_2(t)$ are forward unique and smooth (and can be chosen to be unique at the expense of getting much more complicated $f_1$ and $f_2$). ☐
The right hand side of the ODE in Theorem 1 is bounded but discontinuous. To see this consider the fact that for \( i = i_1 = (0, n, 0, i_4, \ldots, i_m) \) and \( i = i_2 = (0, n, 1, i_4, \ldots, i_m) \) the typical values \( h(i_1) \) and \( h(i_2) \) can be arbitrarily close and yet the values of \( h(i_1 + P(i_1)) - h(i_1) \) and \( h(i_2 + P(i_2)) - h(i_2) \) may differ considerably depending on whether or not the value of \( i_1 \) is raised from 0 to 1. Let \( K(x) \) be defined by

\[
\begin{cases} 
  e^{1-1/x} & \text{if } x > 0 \\
  0 & \text{if } x \leq 0,
\end{cases}
\]

and take the smooth sigmoid

\[
L(x) = 1 - K(1 - K(x)).
\]

Then a somewhat “more continuous” replacement of \( g_i(z) \) would be e.g.

\[
d_i(z) = d(k^{i-1}(z))
\]

where

\[
d(z) = [- \log_2(1 - z)] + L(2^m k(z)) - 1
\]

(see Figures 5 and 6). It remains an open problem whether Theorem 1 is valid for continuous \( f_1 \) and \( f_2 \).

![Figure 5. The graph of \( d(z) \) for \( m = 2 \)](image)

![Figure 6. The graph of \( d(k(z)) \) for \( m = 2 \)](image)

4. ODEs OF DIMENSION ONE

It is possible to still reduce dimension in the construct of the previous section. First it is noted that if \( i_1 \) and \( i_2 \) contain the counter counts of two
computationally consecutive steps of the counter machine to be simulated then the sums of elements of \( i_1 \) and \( i_2 \) differ by at most \( m \). Let then

\[ i_1, i_2, \ldots, i_p \]

be a sequence of computationally consecutive counts where the initial count \( i_1 \) is \((0, \ldots, 0, n)\). Using typical values this can be coded as the number

\[
b(i_1, \ldots, i_p) = \sum_{j=1}^{p} 2^{-(m+n+1)-(2m+n+1)-\cdots-((j-1)m+n+1)} h(i_j)
\]

\[
= \sum_{j=1}^{p} 2^{-l(n,j-1)} h(i_j)
\]

where

\[ l(n, j) = j(n + 1 + \frac{1}{2} m(j - 1)). \]

Denote further

\[ c(n, p, z) = 2^{l(n,p)} z - \lfloor 2^{l(n,p)} z \rfloor. \]

**Lemma 2:** If \( z \) is in the interval \( b(i_1, \ldots, i_p) \leq z \leq b(i_1, \ldots, i_{p+1}) \), then

\[ g(c(n, p - 1, z)) = i_p. \]

**Proof:** If \( b(i_1, \ldots, i_p) \leq z \leq b(i_1, \ldots, i_{p+1}) \), i.e.,

\[
\sum_{j=1}^{p} 2^{-l(n,j-1)} h(i_j) \leq z \leq \sum_{j=1}^{p+1} 2^{-l(n,j-1)} h(i_j),
\]

then

\[
\sum_{j=1}^{p-1} 2^{-l(n,p-1)-l(n,j-1)} h(i_j) + h(i_p) \leq 2^{l(n,p-1)} z
\]

\[
\leq \sum_{j=1}^{p-1} 2^{-l(n,p-1)-l(n,j-1)} h(i_j) + h(i_p) + 2^{l(n,p-1)-l(n,p)} h(i_{p+1})
\]

and

\[
\lfloor 2^{l(n,p-1)} z \rfloor = \sum_{j=1}^{p-1} 2^{l(n,p-1)-l(n,j-1)} h(i_j).
\]
So
\[ h(i_p) \leq 2^{l(n,p-1)}z - [2^{l(n,p-1)}z] \leq h(i_p) + 2^{-pm-n-1}h(i_{p+1}) \]
and, by Lemma 1,
\[ g(c(n, p - 1, z)) = g(2^{l(n,p-1)}z - [2^{l(n,p-1)}z]) = i_p. \]

Thus one gets the initial value problem faithfully simulating the one used in the previous section:
\[ \frac{dz}{dt} = 2^{-l(n,[t]+1)}h(g(c(n, [t], z)) + P(g(c(n, [t], z)))) \]
\[ z(0) = h(i_1) = h(0, \ldots, 0, n). \]

The values of \( z \) start from \( h(i_1) \) and then move piecewise linearly through the values
\[ b(i_1) = h(i_1), b(i_1, i_2), b(i_1, i_2, i_3), \ldots \]
attained at \( t = 0, 1, 2, \ldots \). The event to be detected is
\[ g(c(n, [t], z)) = 1. \]

The computational history is preserved in the value of \( z \), so the value \( n \), given in the initial value \( z(0) \), can be recalled at any time during the simulation. Indeed, by Lemma 1,
\[ n = g_m(z). \]

Thus the initial value problem aimed at here is
\[ \frac{dz}{dt} = 2^{-l(g_m(z),[t]+1)}h(g(c(g_m(z), [t], z)) + P(g(c(g_m(z), [t], z)))) \]
\[ z(0) = h(i_1) = h(0, \ldots, 0, n) \]
and the event is
\[ g(c(g_m(z), [t], z)) = 1. \]

There is no need for the dragging used in the previous section, as the history of the computation is preserved. Thus the solution \( z(t) \) is unique. It
might be noted that the idea of preserving computational history to obtain reversible computing is an old one. It appears first in the work of Y. Lecerf [7] and then independently in [1]. The construct used in [11] is from [14].

A smooth solution $z(t)$ is obtained when the change of variable

$$t = a \int_0^u K(\sin^2 \pi w)dw$$

is made where $K$ is given in the previous section and the constant $a$ is chosen to satisfy

$$a \int_0^1 K(\sin^2 \pi w)dw = 1.$$ 

Note that then $[t] = [u]$. Again the infinite time interval $[0, \infty)$ can be replaced by a finite half open one. Thus the following theorem is proved.

**Theorem 2:** There exists an explicitly given ODE $y' = f(y, t)$ for which the EDP is undecidable in the time interval $[0, \infty)$ (or in a finite half open time interval $[0, T]$). The solution $y(t)$ is unique and smooth. □

Note that the price to be paid for getting Theorem 2 is the complicated structure of the event that must be used. As is easily seen, the event can be made simpler by adding one more dependent variable, and Theorem 1 follows.

The function $f$ in the theorem is discontinuous, for the same reason as that in Theorem 1.

**REFERENCES**

7. M. Y. LECERF, Machines de Turing réversibles. Récursive insolubilité en $n \in N$ de l’équation $u = \theta^n u$, où $\theta$ est un "isomorphisme de codes". *Comptes Rendus* 257, 1963, pp. 2597-2600.


