Termination of Integer Linear Programs

Mark Braverman *

Department of Computer Science
University of Toronto

Abstract. We show that termination of a simple class of linear loops over the integers is decidable. Namely we show that termination of deterministic linear loops is decidable over the integers in the homogeneous case, and over the rationals in the general case. This is done by analyzing the powers of a matrix symbolically using its eigenvalues. Our results generalize the work of Tiwari [Tiw04], where similar results were derived for termination over the reals. We also gain some insights into termination of non-homogeneous integer programs, that are very common in practice.

1 Introduction

Termination analysis is one of the building blocks of automated verification. For a generic loop

while (*conditions*) { *commands* }

it is well known that the termination problem is undecidable in all but the most simple cases. Even when all the conditions and updates are given as piecewise linear functions, the problem of deciding termination of the loop remains undecidable since such programs can naturally simulate counter machines [Tiw04], and the problem of whether a counter machine terminates on all inputs is undecidable [BBK+01].

In view of the undecidability mentioned above, the efforts on practical termination analysis of loops have been concentrated on partial decision procedures. One approach is synthesizing a *ranking function*. Synthesis of ranking functions has been studied in [CSS03,BMP05a,BMP05b]. In some cases, one can even find a complete method for synthesis of *linear* ranking functions [PR04]. Even a complete synthesis method, however, can only establish existence of a certain way of proving termination, and not actually decide the termination problem itself. It is not hard to construct an example of a program that terminates but has no linear ranking function.

The termination problem appears to be much harder, and one can expect it to be decidable only in the simplest cases. In [Tiw04] termination has been shown to be decidable for loops of the form

while $(Bx > b)$ $\{ x \leftarrow Ax + c \}$

* Partially supported by an NSERC postgraduate scholarship

where $Bx > b$ represents a conjunction of linear inequalities over the state variables x, and $x \leftarrow Ax + c$ represents a (deterministic) linear update of each variable. The variables are interpreted over the reals $\mathbb{R}$, and there are no constraints on the initial conditions. Roughly speaking, [Tiw04] shows that only the subspace corresponding to eigenvectors of A with positive real eigenvalues is relevant to the termination problem. In the homogeneous case

while $(Bx > 0)$ $\quad \{\ x \leftarrow Ax\ \}$

it is immediate to see that if there is an eigenvector v of A such that $Av = \lambda v$, $\lambda > 0$ and $Bv > 0$, then the loop is non-terminating on v. The decision procedure depends on the fact that the inequality $Bx > 0$ is strict. More importantly, it depends on the fact that the variables are interpreted over the *reals*. As the following example illustrates, a program may be terminating over the integers, but not over the reals.

Example 1. Consider the homogeneous loop

$$\textbf{while}\ \ (4x + y > 0)\ \ \left\{\ \begin{pmatrix} x \\ y \end{pmatrix} \leftarrow \begin{pmatrix} -2 & 4 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\ \right\}$$

The matrix has two eigenvectors, $(-1-\sqrt{17}, 4)$ and $(-1+\sqrt{17}, 4)$ corresponding to eigenvalues $-1-\sqrt{17}$ and $-1+\sqrt{17}$, respectively.

The eigenvector $(-1+\sqrt{17}, 4)$ satisfies the loop condition, and corresponds to a positive eigenvalue. Hence the loop does not terminate over $\mathbb{R}$. However, the line $(-1+\sqrt{17}, 4)\alpha$ does not contain any rational points, and the loop outside this line is always dominated by the eigenvalue $-1-\sqrt{17} < 0$ that is bigger in absolute value than the other eigenvalue. At the limit, the orbit of (x, y) will alternate between the directions $(-1-\sqrt{17}, 4)$ and $(1+\sqrt{17}, -4)$. Hence the loop terminates on all integers. ■

The example highlights the difference between the integer and the real case. In general, it is not unusual to have differences between hardness of decidability of problems over the reals $\mathbb{R}$ and problems over the integers $\mathbb{Z}$. One notorious example is *quantifier elimination*. Given a quantified formula

$$Q_1 x_1 Q_2 x_2 \ldots Q_n x_n \quad f(x_1, \ldots, x_n),$$

there is an algorithm to decide its validity over $\mathbb{R}$ [Tar51], but not over $\mathbb{Z}$. In fact, by undecidability of Diophantine equations [Mat93], the formula above is undecidable even in the case when $Q_i = \exists$ for all i.

It has been conjectured in [Tiw04] that the termination of programs as above is still decidable when interpreted over the integers. In this paper we prove the following:

Theorem 1. *Let A, B_s, B_w be rational matrices and b_s, b_w, c be rational vectors. Then the termination problem of the loop*

while $(B_s x > b_s) \wedge (B_w x \geq b_w)$ $\quad \{\ x \leftarrow Ax + c\ \}$

is decidable when the variables range over the reals $\mathbb{R}$ or the rationals $\mathbb{Q}$. It is decidable over the integers $\mathbb{Z}$ in the homogeneous *case when $b_s, b_w, c = 0$.*

Theorem 1 settles the termination problem over the rationals for a linear loop with a deterministic update and no initial conditions in the most general form. Using Lemma 4 on linear combinations of sums of powers of complex units, we are able to deal with non-strict inequalities. Over the integers termination in the *non-homogeneous* case remains an intriguing open problem. We will return to it in Section 6.

In practice, the programs are usually specified over integer variables, and it is encouraging to know that the termination of homogeneous loops as above is still decidable in this setting. Most of the paper is dedicated to proving Theorem 1.

Acknowledgments: I would like to thank Marsha Chechik and Arie Gurfinkel for encouraging me to work on the problem. I am also grateful to Arie Gurfinkel for his many useful comment on preliminary versions of the paper. I would like to thank Ilia Binder for our useful discussions on applying the ergodic theorem. Finally, I would like to thank the anonymous referees for the many useful suggestions for improvements of the paper.

2 Proof Outline of Theorem 1

The main part of the proof is in deciding termination over $\mathbb{Q}$ for homogeneous programs (i.e. programs for which $b_s, b_w, c = 0$). Unlike the termination analysis over $\mathbb{R}$ [Tiw04], we cannot ignore the vectors corresponding to negative and complex eigenvalues. As illustrated in the following example, it is possible that there are no rational points on the non-terminating subspace S^+ corresponding to the positive eigenvalues of A, but there is a rational vector outside S^+ very close to it, and on which the loop is still non-terminating.

Example 2. Consider the loop

$$\textbf{while}\ \ (4x - 5y > 0)\ \ \left\{\ \begin{pmatrix} x \\ y \end{pmatrix} \leftarrow \begin{pmatrix} 2 & 4 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}\ \right\}$$

The matrix has two eigenvectors, $(1+\sqrt{17}, 4)$ and $(1-\sqrt{17}, 4)$ corresponding to eigenvalues $1+\sqrt{17}$ and $1-\sqrt{17}$, respectively. The only eigenvector in S^+ is $v_1 = (1+\sqrt{17}, 4)$, which satisfies the loop condition, but contains no rational points. However, the orbit of a rational perturbation q_1 of v_1 converges to the direction of v_1 at the limit. Hence it is possible to choose q_1 that is a nonterminating rational initial condition, and the loop is non-terminating over $\mathbb{Q}$ despite the fact that there are no rational points in S^+. The point $q_1 = (9, 7)$ is an example of a specific such value. Note that $\left|\frac{9}{7} - \frac{1+\sqrt{17}}{4}\right| < 0.005$, which means that q_1 is a good rational approximation of v_1. ■

We will see that the set N of real points for which the program is non-terminating is a convex cone. Hence it has a dimension and a unique minimal linear space S_{min} containing it. The rough outline of the procedure for finding a rational point in N (i.e. in $\mathbb{Q}^n \cap N$) is as follows:

Termination (*loop* P)
compute S_{min}
$Q_{min} \leftarrow S_{min} \cap \mathbb{Q}^n$
if $Q_{min} = \emptyset$
    return *terminating*
if $dim(Q_{min}) = dim(S_{min})$
    return *non-terminating*
else
    reduce the loop to a loop P' on the subspace Q_{min}
    run **Termination**(P')

At each iteration, S_{min} is the current feasible real subspace, and Q_{min} its rational subspace. We continuously update both until their dimensions match or until Q_{min} becomes empty. If the dimensions match, we know that Q_{min} is dense in N, and we can return *non-terminating*. If Q_{min} becomes empty, we can return *terminating*. At each iteration we reduce the dimension of the loop by at least 1, hence the algorithm terminates. The crucial step in the computation is the ability to compute S_{min} at each step of the iteration.

Running the procedure on Example 1 above, we would obtain that S_{min} is the one-dimensional space $span\{(-1+\sqrt{17}, 4)\}$, and $Q_{min} = \{0\}$, thus outputting **terminating**. On the other hand, for Example 2 above we would obtain $S_{min} = \mathbb{R}^2$, and $Q_{min} = \mathbb{Q}^2$, thus $dim(Q_{min}) = dim(S_{min})$, and we output **non-terminating**.

3 Preliminaries

3.1 Linear Algebra

We will see that symbolically powering the matrix A is an essential step in deciding termination of the loop. If A is similar to some matrix D via $A = P^{-1}DP$ then

$$A^n = (P^{-1}DP)^n = (P^{-1}DP)(P^{-1}DP)\dots(P^{-1}DP) = P^{-1}D^nP.$$

Hence powering the matrix A is as hard as powering the matrix D. We would like to make D as simple as possible. It is well known from linear algebra [HK71] that any A can be transformed into *Jordan canonical form*:

Lemma 2 (Jordan canonical form). *For any matrix $A \in \mathbb{C}^{n\times n}$ there is a matrix P, and a matrix D of the form $D = Diag(J_1, J_2, \dots, J_N)$ with each block J_i having the form*

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix},$$

where λ_i is an eigenvalue of A and $A = P^{-1}DP$. Moreover, if A is an algebraic matrix, then D and P are also algebraic matrices and their entries can be computed from the entries of A.

Next, we explicitly write the n-th power of the matrix D. The formula can be proved by induction on n.

Lemma 3 *For a matrix $D = Diag(J_1, \ldots, J_N)$ in Jordan canonical form, its n-th power is given by $D^n = Diag(J_1^n, J_2^n, \ldots, J_N^n)$, where*

$$J_i^n = \begin{pmatrix} \lambda_i^n & n\lambda_i^{n-1} & \binom{n}{2}\lambda_i^{n-2} & \ldots & \binom{n}{N_i-1}\lambda_i^{n-(N_i-1)} \\ 0 & \lambda_i^n & n\lambda_i^{n-1} & \ldots & \binom{n}{N_i-2}\lambda_i^{n-(N_i-2)} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & \lambda_i^n & n\lambda_i^{n-1} \\ 0 & 0 & \ldots & 0 & \lambda_i^n \end{pmatrix},$$

where N_i is the dimension of the block J_i, and $\binom{n}{k} = 0$ if $n < k$.

3.2 A Lemma about Complex Units

Let $\zeta \neq 1$ be a complex number on the unit circle, that is, $|\zeta| = 1$. It is easy to see that the orbit $\zeta, \zeta^2, \zeta^3, \ldots$ will visit the negative half of the complex plane infinitely often. We need a generalization of this fact to a linear combination of such ζ's.

Lemma 4 *Let $\zeta_1, \zeta_2, \ldots, \zeta_m \in \mathbb{C}$ be a collection of distinct complex numbers such that $|\zeta_i| = 1$ and $\zeta_i \neq 1$ for all i. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be any complex numbers. Denote*

$$z_n = \alpha_1\zeta_1^n + \alpha_2\zeta_2^n + \ldots + \alpha_m\zeta_m^n.$$

Then one of the following is true:

1. *the real part $Re(z_n) = 0$ for all n; or*
2. *there is a $c < 0$ such that $Re(z_n) < c$ for infinitely many n's.*

We will be interested in the case when $z_n \in \mathbb{R}$ are all reals. In this case we have $Re(z_n) = z_n$ for all n, and the lemma applies directly to z_n.

Proof. Due to space constraints, we will only present a proof idea here. First of all, we can write

$$y_n = 2Re(z_n) = z_n + \bar{z}_n = \alpha_1\zeta_1^n + \ldots + \alpha_m\zeta_m^n + \bar{\alpha}_1\bar{\zeta}_1^n + \ldots + \bar{\alpha}_m\bar{\zeta}_m^n \in \mathbb{R}.$$

After collecting together terms where $\zeta_i = \bar{\zeta}_j$, we see that the claim for y_n is equivalent to the claim for the z_n, but now $y_n \in \mathbb{R}$ for all n. Hence it suffices to prove the lemma under the assumption $z_n \in \mathbb{R}$. We actually show that if z_n is not syntactically 0, then the second possibility above holds.

The two key claims of the proof are that

1. The cumulative sum of the z_n is bounded from above: $\left|\sum_{n=0}^{N} z_n\right| \leq C_1$, where $C_1 > 0$ is some explicit constant.
2. The sum of absolute values $|z_n|$ is bounded from below: $\sum_{n=N+1}^{N+m} |z_n| > C_2$ for each N for some explicit constant $C_2 > 0$.

Both claims are not too hard to prove, and together they yield the statement of the lemma: Choose an integer K such that $K \cdot C_2 > 4C_1$. Then for any N we have by the first claim

$$\sum_{n=N+1}^{N+Km} z_n = \sum_{n=0}^{N+Km} z_n - \sum_{n=0}^{N} z_n < 2C_1.$$

On the other hand, by the second claim we have

$$\sum_{n=N+1}^{N+Km} |z_n| = \sum_{i=0}^{K-1} \sum_{n=N+mi+1}^{N+mi+m} |z_n| > K \cdot C_2 > 4C_1.$$

These together imply that

$$\sum_{n=N+1,\ z_n<0}^{N+Km} z_n < -C_1.$$

Hence there is an $n \in \{N+1, \ldots, N+Km\}$ such that $z_n < -C_1/(Km)$.

Set $c = -C_1/(Km)$. We have just seen that there is a z_n satisfying $z_n < c$ among any Km consecutive elements. This completes the proof of Lemma 4.

Remark: It is also possible to give a less constructive proof of Lemma 4 using ergodic theory. ■

4 Termination over $\mathbb{Q}$ and $\mathbb{R}$ in the Homogeneous Case

In this section we assume that the loop is homogeneous, that is $c, b_s, b_w = 0$. Let N be the set of nonterminating points of the program over $\mathbb{R}^n$. We are interested in determining whether N and $N \cap \mathbb{Q}^n$ are empty.

For a point $z \in \mathbb{R}^n$ we consider the evolution of the loop with initial variables vector z. We denote the value of the variables after i iterations by $z(i) = A^i z$. In particular $z(0) = z$. $z \in N$ if and only if $z(i)$ satisfies the loop conditions $B_s z(i) > 0$ and $B_w z(i) \geq 0$ for all $i \geq 0$.

First, we note that N is a convex cone.

Lemma 5 *Assuming $N \neq \emptyset$, N must be a convex cone. That is, for every $x, y \in N$ and $\lambda > 0$, $\lambda x \in N$, and the line segment connecting x to y belongs to N.*

Proof. Since the loop is homogeneous, the execution on x will run for exactly as long as the execution on λx. In particular if the loop does not terminate on x, it will not terminate on λx. Suppose that initially z is on the line segment connecting x with y. Then $z(0) = z = \alpha x + (1-\alpha)y$ for some $\alpha \in [0,1]$. On the n-th iteration we have

$$z(n) = \alpha x(n) + (1-\alpha)y(n),$$

is still on the line segment connecting $x(n)$ with $y(n)$, and

$$\begin{aligned} B_s z(n) &= \alpha B_s x(n) + (1-\alpha) B_s y(n) > 0, \\ B_w z(n) &= \alpha B_w x(n) + (1-\alpha) B_w y(n) \geq 0, \end{aligned}$$

because the loop does not terminate on both x and y. ■

N is a convex body in $\mathbb{R}^n$ and as such, has a dimension d_N, which is the rank of the smallest subspace containing N. Determining the minimum linear space $S_{min} = span\{N\}$ containing N is central to the construction.

4.1 Finding the Minimum Space $S_{min} \supset N$

Intuition: N is a convex cone. If we consider N as a subset on S_{min}, we see that it has an interior $int(N)$, and for any point x in the interior small perturbations $x + \varepsilon v$ of x remain in N if and only if $v \in S_{min}$. The v's for which $x + \varepsilon v$ is in N span S_{min}. We first find such an x, we call z_{max}, and then generate all the small perturbations that leave z_{max} in N in order to get a linear basis for S_{min}.

We are interested in the behavior of the loop with initial condition $z(0)$. In particular, we would like to know whether $z(i) = A^i z(0)$ always satisfies the loop conditions. Since we know the Jordan canonical form of A, we can explicitly write the **while** condition after i steps as

$$\begin{cases} B_s A^i z(0) > 0 \\ B_w A^i z(0) \geq 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} B_s P^{-1} D^i P z(0) > 0 \\ B_w P^{-1} D^i P z(0) \geq 0 \end{cases} \tag{1}$$

where $D = PAP^{-1} = Diag(J_1, \ldots, J_N)$ is the Jordan canonical form of A.

Our next goal is to use (1) to write the conditions on $z(i)$ in an explicit form. Let $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_r$ be the absolute values of the eigenvalues of A sorted in the increasing order. We only consider the nonzero eigenvalues here. Let $\{\zeta_{ij}\}$ be complex numbers on the unit circle, $|\zeta_{ij}| = 1$, and $\zeta_{ij} \neq 1$ such that the eigenvalues of A are a subset of

$$\{\lambda_1, \lambda_1\zeta_{11}, \lambda_1\zeta_{12}, \ldots, \lambda_1\zeta_{1m_1}, \lambda_2, \lambda_2\zeta_{21}, \lambda_2\zeta_{22}, \ldots, \lambda_2\zeta_{2m_2}, \ldots,$$

$$\lambda_r, \lambda_r\zeta_{r1}, \lambda_r\zeta_{r2}, \ldots, \lambda_r\zeta_{rm_r}\}.$$

The ζ_{ij} are the arguments of the corresponding eigenvalues. By Lemma 3, symbolically, D^i is a linear combination of

$$\{\lambda_1^i, \lambda_1^i\zeta_{11}^i, \lambda_1^i\zeta_{12}^i, \ldots, \lambda_1^i\zeta_{1m_1}^i, i\lambda_1^{i-1}, i\lambda_1^{i-1}\zeta_{11}^{i-1}, i\lambda_1^{i-1}\zeta_{12}^{i-1}, \ldots, i\lambda_1^i\zeta_{1m_1}^{i-1}, \ldots,$$

$$\binom{i}{n_1-1}\lambda_1^{i-(n_1-1)}, \binom{i}{n_1-1}\lambda_1^{i-(n_1-1)}\zeta_{11}^{i-(n_1-1)}, \ldots,$$

$$\binom{i}{n_1-1}\lambda_1^{i-(n_1-1)}\zeta_{1m_1}^{i-(n_1-1)}, \lambda_2^i, \lambda_2^i\zeta_{21}^i, \lambda_2^i\zeta_{22}^i, \ldots, \lambda_2^i\zeta_{2m_2}^i, \ldots$$

$$\lambda_r^i, \lambda_r^i\zeta_{r1}^i, \lambda_r^i\zeta_{r2}^i, \ldots, \lambda_r^i\zeta_{rm_r}^i, \ldots, \binom{i}{n_r-1}\lambda_r^{i-(n_r-1)},$$

$$\binom{i}{n_r-1}\lambda_r^{i-(n_r-1)}\zeta_{r1}^{i-(n_r-1)}, \ldots, \binom{i}{n_r-1}\lambda_r^{i-(n_r-1)}\zeta_{rm_r}^{i-(n_r-1)}\}$$

Thus we can rewrite (1) as a set of conditions on the initial $z(0)$ of the form

$$Cond_k(z(0),i) = \lambda_1^i(C_{k11} + \zeta_{11}^i N_{k111} + \zeta_{12}^i N_{k112} + \ldots + \zeta_{1m_1}^i N_{k11m_1})z(0) +$$

$$i\lambda_1^{i-1}(C_{k12} + \zeta_{11}^{i-1} N_{k121} + \zeta_{12}^{i-1} N_{k122} + \ldots + \zeta_{1m_1}^{i-1} N_{k12m_1})z(0) + \ldots +$$

$$\binom{i}{n_1-1}\lambda_1^{i-(n_1-1)}(C_{k1n_1} + \zeta_{11}^{i-(n_1-1)} N_{k1n_11} + \zeta_{12}^{i-(n_1-1)} N_{k1n_12} + \ldots +$$

$$\zeta_{1m_1}^{i-(n_1-1)} N_{k1n_1m_1})z(0) + \ldots +$$

$$\lambda_r^i(C_{kr1} + \zeta_{r1}^i N_{kr11} + \zeta_{r2}^i N_{kr12} + \ldots + \zeta_{rm_1}^i N_{kr1m_r})z(0) +$$

$$i\lambda_r^{i-1}(C_{kr2} + \zeta_{r1}^{i-1} N_{kr21} + \zeta_{r2}^{i-1} N_{kr22} + \ldots + \zeta_{rm_r}^{i-1} N_{kr2m_r})z(0) + \ldots +$$

$$\binom{i}{n_r-1}\lambda_r^{i-(n_r-1)}(C_{krn_r} + \zeta_{r1}^{i-(n_r-1)} N_{krn_r1} + \zeta_{r2}^{i-(n_r-1)} N_{krn_r2} + \ldots +$$

$$\zeta_{rm_r}^{i-(n_r-1)} N_{krn_rm_r})z(0) \triangleright 0,$$

where $\triangleright \in \{>, \geq\}$. The coefficients $C_{kj\ell}$ and $N_{kj\ell t}$ are all algebraic vectors and can be computed explicitly. Moreover, in our case all the conditions and A are over the reals, hence every coefficient $\sum_{t=1}^{m_j} N_{kj\ell t}\zeta_{jt}^i$ will add up to a real number. A point $z(0)$ is in N if and only if the conditions $Cond_k(z(0),i)$ are satisfied for all k and for all $i = 0, 1, 2 \ldots$.

Using the Jordan canonical form of A, we can split the space $\mathbb{R}^n$ into the subspace S^+ corresponding to the positive eigenvalues of A, and the subspace S^o corresponding to the other eigenvalues. Each $v \in \mathbb{R}^n$ decomposes uniquely into a sum $v = v^+ + v^o$ such that $v^+ \in S^+$ and $v^o \in S^o$. If we write $Cond_k(z(0)^+, i)$ we get all $N_{kj\ell t}$'s equal to zero, since there are no vectors in S^+ corresponding to the complex eigenvalues. Similarly, in $Cond_k(z(0)^o, i)$ we get all $C_{kj\ell}$'s equal to zero.

Observe that the magnitude of the terms

$$C_{kj\ell} + \zeta_{j1}^i N_{kj\ell 1} + \zeta_{j2}^i N_{kj\ell 2} + \ldots \zeta_{jm_j}^i N_{kj\ell m_j}$$

remains bounded by a constant independent of i throughout the iteration. Hence the magnitude of the components of $Cond_k(z(0))$ as i tends to ∞ is primarily

dictated by the $\binom{i}{\ell-1}\lambda_j^{i-(\ell-1)}$ terms of the products. These terms have a clear dominance order as $i \to \infty$. For higher j the terms grow geometrically faster, because $\lambda_{j_1}^i \ll \lambda_{j_2}^i$ for $j_1 < j_2$. For the same j, terms with higher ℓ grow polynomially faster, because $\binom{i}{\ell_1-1} \ll \binom{i}{\ell_2-1}$ for $\ell_1 < \ell_2$. This yields a natural lexicographic order $\prec$ on the pairs of indexes $j\ell$:

$$Ind = \{0 \prec 11 \prec 12 \prec \ldots \prec 1n_1 \prec 21 \prec \ldots \prec 2n_2 \prec \ldots \prec r1 \prec \ldots \prec rn_r\}.$$

The term 0 is the smallest term, it is introduced for completeness in the case of non-strict inequalities. It does not correspond to any actual index.

Our first step is very similar to [Tiw04]: we solve the problem over the positive eigenspace S^+.

Lemma 6 *For every vector $z \in S^+$ the program with initial conditions $A^q z$ is non-terminating for some integer $q \geq 0$ if and only if there is a function*

$$index_z : k \mapsto index_z(k) \in Ind$$

which maps the condition $Cond_k(z)$ to the highest ranking nonzero $C_{k,index_z(k)}$. All higher ranking coefficients must be zero. In other words, for each k,

$$\begin{cases} C_{k,ind} z = 0, & \text{if } ind \succ index_z(k) \\ C_{k,ind} z > 0, & \text{if } ind = index_z(k) \end{cases}$$

In the case that the k-th inequality is strict we must have $index_z(k) \succ 0$.

Proof. First of all note that since z is in S^+, only the $C_{k,ind}$ (and no $N_{k,ind,t}$'s) appear in the expressions for $Cond_k(z,i)$.

It is obvious that for $A^q z$ to be non-terminating for some q the conditions $Cond_k(z,i)$ must be satisfied as $i \to \infty$. In particular, the highest ranking coefficient, which dominates the behavior as i goes to infinity must be positive (or all of them may be 0 in the case of a non-strict inequality). Note that $index_z$ is a well-defined function for each such z.

Conversely, if the $index_z(k)$ function as in the statement of the lemma exists, then the dominating term in each $Cond_k(z,i)$ has a positive coefficient. Hence the conditions $Cond_k(z,i)$ are satisfied for sufficiently large i. In particular, there is a q such that they are satisfied for $i \geq q$, making the program non-terminating on $A^q z$. ■

We denote the set of z's for which $A^q z \in N$ for some q by N^e – "eventually non-terminating". Those are the points which might be terminating, but become non-terminating after finitely many applications of A. Lemma 6 gives a characterization of N^e, and associates a unique function $index_z$ with each $z \in N^e$. We claim that there is a maximum such function.

Lemma 7 *There is a $z_{max} \in N \cap S^+$ with an index function $index_{z_{max}} = index_{max}$ such that for any $z \in N \cap S^+$, for all k,*

$$index_z(k) \preceq index_{max}(k).$$

Proof. First we note that N^e is convex. If $z_1, z_2 \in N^e$, then there is a q such that $A^q z_1, A^q z_2 \in N$. N is convex, hence the line segment I connecting $A^q z_1$ to $A^q z_2$ is in N. The line segment connecting z_1 to z_2 is mapped to I by A^q, hence it is in N^e.

Denote $z = (z_1 + z_2)/2$. Then it is easy to see that

$$index_z = \max(index_{z_1}, index_{z_2}).$$

Thus, there can be only one maximal index function, which is the maximum index function for some $z'_{max} \in N^e \cap S^+$. We can take sufficiently many iterations of z'_{max} to obtain $z_{max} \in N \cap S^+$.

Note that it is easy to compute z_{max} and $index_{max}$ by considering the constraint satisfaction problem corresponding to each index function and choosing the maximum feasible function and a corresponding z_{max}. In fact, a generic element y of $N \cap S^+$ satisfies $index_y = index_{max}$. ■

As mentioned in the beginning of the section, the main idea in finding the minimal space S_{min} containing N is that it is spanned by small perturbations of z_{max}. The claim is that a small perturbation of z_{max} is in N as long as we do not introduce any terms that are more dominant than the currently dominant ones.

Lemma 8 *For a vector $v \in \mathbb{R}^n$ there is an $\varepsilon \neq 0$ such that $z_{max} + \varepsilon v \in N$ if and only if for all k*

$$\begin{cases} C_{k,ind} v = 0, & \text{for } ind \succ index_{max}(k) \\ N_{k,ind,t} v = 0, & \text{for } ind \succ index_{max}(k), \text{ for all } t \end{cases}$$

Proof. **The "if" direction.** $Cond_k(z_{max}, i)$ is dominated by the $C_{k,index_{max}(k)} z$ term for all i. It will remain positive if we add εv to it for some small ε. By the condition it will remain dominating, since no non-zero higher order terms are introduced by adding εv.

The "only if" direction. We first show by contradiction that the first condition must hold. Suppose that there is a v and ε such that

$$y = z_{max} + \varepsilon v \in N,$$

but $C_{k,ind} v \neq 0$ for some k and $ind \succ index_{max}(k)$. Decompose $y = y^+ + y^o$, so that $y^+ \in S^+$ and $y^o \in S^o$. Then $C_{k,ind} y = C_{k,ind} y^+$. There are two cases:

Case 1: For each k, the highest-ranking non-zero $C_{k,ind} y^+$ is positive. In this case $y^+ \in N^e$ by Lemma 6. By the definition of $index_{max}$ we get $C_{k,ind} y = C_{k,ind} y^+ = 0$ for all $ind \succ index_{max}$. Hence $C_{k,ind} v = (C_{k,ind} y)/\varepsilon = 0$, contradiction.

Case 2: There is a k such that the highest-ranking non-zero $C_{k,ind} y = C_{k,ind} y^+$ is negative. In this case the dominating term of $Cond_k(y, i)$ has the coefficient

$$C_{k,ind} y + \zeta^i_{k1} N_{k,ind,1} y + \zeta^i_{k2} N_{k,ind,2} y + \ldots + \zeta^i_{km_k} N_{k,ind,m_k} y.$$

By Lemma 4 the expression will be negative below $C_{k,ind}y$ infinitely often, hence $Cond_k(y,i)$ will be violated infinitely often, contradiction.

Now suppose that for some k and $ind \succ index_{max}(k)$ the second condition is violated. We already know that $C_{k,ind}y = 0$, and the dominating term of $Cond_k(y,i)$ has the coefficient

$$C(i) = \zeta^i_{k1}N_{k,ind,1}y + \zeta^i_{k2}N_{k,ind,2}y + \ldots + \zeta^i_{km_k}N_{k,ind,m_k}y,$$

which is not identically 0. By Lemma 4 we know that there is a $c < 0$ such that $C(i) < c$ infinitely often. Since this is a dominating term, it will cause $Cond_k(y,i)$ to be violated infinitely often, contradiction. ■

Solving the constraint system from Lemma 8 gives us a linear basis for S_{min}. The computation is done entirely symbolically over algebraic numbers. Note that we do not need to know ε from Lemma 8, but merely that such an ε exists. This solves the termination problem over $\mathbb{R}$. Our goal now is to tackle the problem over $\mathbb{Q}$. If $S_{min} = \emptyset$, we can return **terminates**, otherwise we need to find the rational subspace of S_{min}.

4.2 Looking for Rational Points in S_{min}

If the parameters of the loop are given by rationals, then the spanning vectors of S_{min} can be produced as explicit algebraic numbers. Denote by L_S the base vectors for S_{min} presented as algebraic numbers in some finite degree extension $\mathbb{Q}(\alpha)$ of $\mathbb{Q}$. By viewing $\mathbb{Q}(\alpha)$ as a finite-dimensional vector space over $\mathbb{Q}$ we can find the maximum space Q_{min} of *rational vectors* spanned by L_S. For further details about computations with algebraic numbers see [Bhu93,Loos83,Yap00]. We illustrate finding the rational subspace with the following simple example.

Example 3. Consider the simple example when $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{2})$ and $L_S = \{v_1 = (1,0,\sqrt{2}), v_2 = (-\sqrt{2},1,0)\}$. We are looking for coefficients $\beta, \gamma \in \mathbb{Q}(\sqrt{2})$ for which $\beta v_1 + \gamma v_2 \in \mathbb{Q}^3$. By writing $\beta = \beta_1 + \beta_2\sqrt{2}$, $\gamma = \gamma_1 + \gamma_2\sqrt{2}$ with $\beta_1, \beta_2, \gamma_1, \gamma_2 \in \mathbb{Q}$ we obtain the conditions

$$\begin{cases} \beta + (-\sqrt{2})\gamma \in \mathbb{Q} \\ \gamma \in \mathbb{Q} \\ (\sqrt{2})\beta \in \mathbb{Q} \end{cases} \iff \begin{cases} \beta_2 - \gamma_1 = 0 \\ \gamma_2 = 0 \\ \beta_1 = 0 \end{cases}$$

Hence we must have $\gamma = \beta/\sqrt{2} \in \mathbb{Q}$, and the rational subspace of $span(L_S)$ is one dimensional, spanned by $\sqrt{2}v_1 + v_2 = (0,1,2)$. ■

There are three possible cases. The first one is that $dim(Q_{min}) = dim(S_{min})$. This means that the rational points are dense in the nonterminating set N, and hence there are nonterminating rational points, and we can return **non-terminating**.

If $dim(Q_{min}) = 0$, then the only potential nonterminating rational point is 0. It is trivial to check whether 0 is non-terminating in the homogeneous case:

we just need to check whether it satisfies the loop conditions. If it does we return **non-terminating**, otherwise return **terminating**.

The more difficult and interesting case is when $0 < d = dim(Q_{min}) < dim(S_{min})$. In this case there are some rational points in S_{min}, but we can no longer guarantee that any of them are in N, since they all lie in a proper subspace of S_{min}. The only thing we know is that *all* potential rational non-terminating points lie in Q_{min}. Denote by R_{min} the space of *real* vectors spanned by Q_{min}. Obviously $dim(R_{min}) = dim(Q_{min})$. We prove the following.

Lemma 9 *R_{min} is invariant under A, that is $Av \in R_{min}$ for any $v \in R_{min}$.*

Proof. First of all, the non-terminating set N is invariant under A, since if the loop is nonterminating on x, it is also nonterminating on Ax. N contains a linear basis for S_{min}, hence S_{min} is invariant under A.

Let q be any rational vector in Q_{min}. Aq is rational, and $Aq \in S_{min}$ by the invariance of S_{min}. Hence by the definition of Q_{min} (as containing all the rational vectors in S_{min}), $Aq \in Q_{min}$. The rational vectors of R_{min} span it, hence R_{min} is invariant under A. ■

R_{min} is a subspace invariant under A, and it has a rational basis $L_R = \{r_1, \ldots, r_d\}$. We can translate the action of A on R_{min} with respect to L_R, to obtain a $d \times d$ *rational* matrix A' such that

$$A : \alpha_1 r_1 + \ldots + \alpha_d r_d \mapsto \beta_1 r_1 + \ldots + \beta_d r_d,$$

where $(\beta_1, \ldots, \beta_d)^T = A'(\alpha_1, \ldots, \alpha_d)^T$. The conditions $B_s x > 0$ and $B_w x \geq 0$ can also be readily translated into *rational* conditions over the d-dimensional coefficient vector $(\alpha_1, \ldots, \alpha_d)$, where $x = \alpha_1 r_1 + \ldots + \alpha_d r_d \in R_{min}$. Thus we obtain a new loop, over d-dimensional vectors

while $(B'_s x > 0) \wedge (B'_w x \geq 0)$ $\quad \{\ x \leftarrow A'x\ \}$

and we need decide termination of the new loop over $\mathbb{Q}$. Note that we have reduced the dimension of the problem from n to $d < n$, and thus we will be able to decide termination over $\mathbb{Q}$ in the homogeneous case in at most n iterations.

5 The Integer and the Non-Homogeneous Cases

In the case the program is interpreted over the reals or the rationals, the transition from general termination to the homogeneous case is done exactly as in [Tiw04] by adding an extra auxiliary variable z. The program

while $(B_s x > b_s) \wedge (B_w x \geq b_w)$ $\quad \{\ x \leftarrow Ax + c\ \}$

always terminates if and only if the program

while $(B_s x > b_s z) \wedge (B_w x \geq b_w z) \wedge (z > 0)$ $\quad \{\ x \leftarrow Ax + cz,\ \ z \leftarrow z\ \}$

terminates. This is true both over $\mathbb{Q}$ and $\mathbb{R}$. If the first program does not terminate, then the second does not terminate with the same initial condition and $z = 1$. In the opposite direction, we can scale a nonterminating starting point of

the second program so that $z = 1$, and thus make it a nonterminating starting point for the first one.

Note that in the homogeneous case we can scale any nonterminating solution, and hence termination over $\mathbb{Q}$ is always equivalent to termination over $\mathbb{Z}$. This is not true in the non-homogeneous case: termination over $\mathbb{Q}$ implies termination over $\mathbb{Z}$, but not vice versa. Thus it can only be used as a partial termination test. The termination problem over $\mathbb{Z}$ as well as termination of loops with initial conditions appears to be much harder and will be discussed in next section.

6 Further Directions and Open Problems

We have seen that termination of deterministic loops with no initial conditions is decidable over $\mathbb{Q}$ and over $\mathbb{Z}$ in the homogeneous case. On the other hand, by allowing the linear loop to be general enough one can easily make the termination problem undecidable. For example, having k different update functions depending on different conditions

while one of the k conditions is met for $1 \leq i \leq k$
  if $B_i x > d_i$ { $x \leftarrow A_i x + c_i$ }

is enough to make the termination problem undecidable, since this class of loops is sufficiently rich to allow encoding of counter machines [Tiw04].

This gives rise to natural open questions about termination of programs more general than the ones considered in this paper, but for which termination is still decidable. One such class are the programs discussed in [PR04]. They are similar to the ones described here, but have a nondeterministic inequality as an update:

while $(B_s x > b_s) \wedge (B_w x \geq b_w)$ { $x \leq Ax + c$ }

In [PR04] a complete linear ranking function generating algorithm is presented, but it still leaves the more general termination problem open over either $\mathbb{R}$, $\mathbb{Q}$ or $\mathbb{Z}$.

Another natural generalization is introducing initial conditions and the related problem of termination over $\mathbb{Z}$. It appears that to decide termination over $\mathbb{Z}$ it is necessary to be able to tell, given a point x_0, whether the program terminates on x_0 or not. Solving the termination problem on a given input would require a much sharper version of Lemma 4. In Lemma 4, we have shown that the expression $z_n = \sum_{i=1}^{m} \alpha_i \zeta_i^n$ always eventually falls below zero by at least some fixed amount c. It is even possible to compute the infimum of the expression using ergodic theory. However, this still falls short of solving the termination problem. Consider the following algebraic expression. Here $|\zeta| = 1$, $\zeta \neq 1$:

$$z(i) = Re(\zeta^i + 1 - 2^{-i}).$$

We would like to know whether $z(i)$ ever falls below 0. This depends on how close the orbit of ζ^i gets to -1. To answer this question some analysis of the continued fraction expansion of $\log \zeta$ seems to be needed.

We summarize the problems:

1. Given a deterministic linear loop P and an input x_0, does P terminate on x_0?
2. Given a deterministic linear loop P does it terminate on all integer inputs?
3. How much nondeterminism can be introduced in a linear loop with no initial conditions before termination becomes undecidable?

7 Conclusion

We have demonstrated a first termination decision procedure that works over the integers for simple homogeneous loop programs. Most programs in practice are specified over the integers, yet algorithms usually only work with the larger domain of real numbers because decision procedures are generally easier there.

We have gained new insights into termination of more general deterministic linear loops. We believe that techniques presented in the paper can be generalized using more refined analysis to obtain at least a good partial termination test over the integers for loops with initial constraints.

References

[BPR03] S. Basu, R. Pollack, M.F. Roy, Algorithms in Real Algebraic Geometry, Springer, 2003.

[Bhu93] M. Bhubaneswar, Algorithmic Algebra, Springer-Verlag, 1993.

[BBK+01] V.D. Blondel, O. Bournez, P. Koiran, C.H. Papadimitriou, J.N. Tsitsiklis, Deciding stability and mortality of piecewise affine dynamical system. Theoretical Computer Science, **255** (1-2), pp. 687696, 2001.

[BMP05a] A.R. Bradley, Z. Manna, H.B. Simpa, Linear ranking with reachability, in CAV 2005, pp. 491-504, 2005.

[BMP05b] A.R. Bradley, Z. Manna, H.B. Simpa, Termination analysis of integer linear loops, in CONCUR 2005, pp. 488-502, 2005.

[CSS03] M.A. Colón, S. Sankaranarayanan, H.B. Simpa, Linear invariant generation using non-linear constraint solving, in CAV 2003, LNCS 2725, pp. 420-432, 2003.

[HK71] K. Hoffman and R. Kunze, Linear Algebra, Prentice-Hall, 2nd ed., 1971.

[Loos83] R. Loos, Computing in Algebraic Extensions, in B. Buchberger, G.E. Collins et al. eds., Computer Algebra: Symbolic and Algebraic Computation, 2nd ed., Springer-Verlag, pp. 173-188, 1983.

[Mat93] Y. Matiyasevich, *Hilbert's Tenth Problem*, The MIT Press, Cambridge, London, 1993.

[PR04] A. Podelski, A. Rybalchenko, A complete method for synthesis of linear ranking functions, in B. Steffen and G. Levi (Eds.): VMCAI 2004, LNCS 2937, pp. 239-251, 2004.

[Tar51] A. Tarski, A Decision Method for Elementary Algebra and Geometry, 2nd ed. Berkeley, CA: University of California Press, 1951.

[Tiw04] A. Tiwari, Termination of linear programs, in R. Alur and D.A. Peled (Eds.): CAV 2004, LNCS 3114, pp. 70-82, 2004.

[Yap00] C.K. Yap, Fundamental Problems of Algorithmic Algebra, Oxford University Press, 2000.