Bounds for the $3x+1$ Problem using Difference Inequalities

Ilia Krasikov
Brunel University
Department of Mathematical Sciences
Uxbridge UB8 3PH, United Kingdom

Jeffrey C. Lagarias
AT&T Labs - Research ,
Florham Park, NJ 07932-0971

(September 13, 2002 revision)

ABSTRACT

This paper studies difference inequality systems $\mathcal{I}_k$ for the $3x+1$ problem introduced by the first author in 1989, which depend on an integer parameter $k \geq 2$. These systems imply lower bounds for the number of integers below x which iterate to 1 under the $3x+1$ map. Previous results deriving such lower bounds gave away some information present in these inequalities. We give here an improvement which (apparently) extracts full information from the inequalities.

Taking the case $k = 11$, we deduce by computer-aided proof, that for any fixed positive integer a not divisible by 3, and for large enough x (depending on a), at least $x^{0.84}$ of integers below x have a in their forward orbit under the $3x+1$ function.

Mathematics Subject Classification (2000): 11B83 Primary; 26A18, 32H50 Secondary

Keywords: $3x+1$ problem, difference equations, dynamical system, iteration, nonlinear programming

1. Introduction

The $3x+1$ problem concerns the iteration of the $3x+1$ function $T(n) = n/2$ if n is an even integer, $(3n+1)/2$ if n is an odd integer. The well known $3x+1$ Conjecture asserts that all integers $n \geq 1$ eventually reach 1 under iteration of the $3x+1$ function. Results on this problem are surveyed in Lagarias [5] and Wirsching [9].

Let $\pi_1(x)$ count the number of integers below x that eventually reach 1 under this iteration. There are a number of methods known for establishing explicit lower bounds of the form $\pi_1(x) > x^\gamma$ for a positive constant γ. The first such bound was obtained in 1978 by Crandall [3], and other methods are described in [1], [2], [4], [6]. We remark that there is also a recent approach of Sinai [7] which gets information about preimages of the $3x+1$ map in an entropy sense, but at present this approach does not yield explicit lower bounds for $\pi_1(x)$. The strongest of the methods giving explicit bounds at present appears to be one introduced by the first author in 1989 ([4]), which uses systems of difference inequalities, and in this paper we consider it further.

This method formulates, for each $k \geq 2$, a system $\mathcal{I}_k$ of functional difference inequalities (mod 3^k), containing about 3^k variables, which certain functions, computed from $3x+1$ iterates, satisfy; they are specified in §2. One can establish an exponential lower bound for the growth rate of positive monotone solutions to these inequalities, and this translates into lower bounds for $\pi_1(x)$ of the form x^γ for some positive γ. The original paper [4] used the system $k=2$ to obtain a lower bound $x^{0.43}$ for the number of such integers. Later Wirsching [8] used the system $k=3$ to obtain the lower bound $x^{0.48}$, for all sufficiently large x.

In 1995 Applegate and Lagarias [2] introduced a nonlinear programming method to systematically deduce lower bounds from the difference inequalities $\mathcal{I}_k$. Their first step was to iterate the inequalities to obtain a derived system of difference inequalities $\mathcal{D}$ such that any positive, monotone solution to the original inequalities would remain a solution of the derived inequalities. This step can be done in many ways. To each such system of difference inequalities $\mathcal{D}$ they associated a parametrized family of auxiliary linear programs $L_k^{\mathcal{D}}(\lambda)$ depending on the parameter λ. This parameter lies in the interval $1 \leq \lambda \leq 2$, and the coefficients of the linear program depend nonlinearly on λ. If the derived system $\mathcal{D}$ contained only "retarded" variables (as defined below) then any positive feasible solution to the linear program for a fixed λ yields a rigorous exponential lower bound for the growth of any positive monotone solution of the system $\mathcal{D}$ with exponential growth constant λ; one then derives a lower bound $\pi_1(x) > cx^\gamma$ with $\gamma = \log_2 \lambda$. For a fixed system $\mathcal{D}$ the determination of the largest value of λ where a positive feasible solution exists, is a nonlinear programming problem. To obtain inequality systems with retarded variables, Applegate and Lagarias [2] found it necessary to apply at some point a "truncation" operation which weakens the inequalities and presumably weakens the lower bounds attained. Using the system $k=9$, and a particular sequence of derivations $\mathcal{D}$, a large computation yielded a lower bound $\pi_1(x) \geq x^{0.81}$ for all sufficiently large

x. Up to now this is the best asymptotic lower bound obtained for $\pi_1(x)$.

The object of this paper is to improve the method for extracting lower bounds from the difference inequalities $\mathcal{I}_k$ of Krasikov [4]. The nonlinear programming approach given in [2] did not apply directly to the original system of difference inequalities $\mathcal{I}_k$ (viewed as a derived system) because these inequalities contain terms with "advanced" variables (as defined below). Here we establish that the nonlinear program lower bounds derived directly from the original inequality system $\mathcal{I}_k$ do give legitimate lower bounds for the $3x+1$ function. The main theorem is stated in §2; it applies to a linear program family denoted $L_k^{NT}(\lambda)$ below. The proof is based on deriving from the original difference inequality system $\mathcal{I}_k$ an auxiliary inequality system from which advanced variables have been eliminated without using any truncation operations. The result is surprising because it is not obvious a priori that advanced variables can be eliminated.

The main theorem yields an immediate improvement of the current record exponent for lower bounds for the $3x+1$ problem, relying on computations already given in [2]. That paper reported computations for certain nonlinear programs L_λ^{NT} as a possible limit of one might hope from the nonlinear programming approach. In §6 we show that The system L_λ^{NT} for a given k has the same maximal admissible value of λ as the family of linear programs $L_k^{NT}(\lambda)$ studied here, and for $k = 9$ it yields a small improvement to the lower bound $x^{0.816}$. Using a further computation for $k = 11$, we report here in §6 the improved lower bound

$$\pi_1(x) > x^{0.84},$$

valid for all sufficiently large x.

The main theorem may prove useful for further work on the $3x + 1$ problem. The linear program families $L_k^{NT}(\lambda)$ have a relatively simple structure, although they are of exponential size in k. One hopes that a bound of the form $\pi_1(x) > x^{1-\epsilon}$ for any $\epsilon > 0$ can eventually be proved by considering $L_k^{NT}(\lambda)$ for arbitrarily large k, and understanding better the structure of the feasible solution sets to these linear program systems.

2. Main Result

We first recall the difference inequalities $\mathcal{I}_k$ of Krasikov [4]. We consider the $3x+1$ function $T(n)$, and for $a \not\equiv 0 \pmod 3$ and $x \geq 1$ we define the function

$$\pi_a(x) := \#\{n : 1 \leq n \leq x, \text{some } T^{(j)}(n) = a.\}$$

and the related function

$$\pi_a^*(x) := \#\{n : n \leq x, \text{ some } T^{(j)}(n) = a, \text{ all } T^{(i)}(n) \leq x \text{ for } 0 \leq i \leq j\}.$$

Note that $\pi_a^*(x) \leq \pi_a(x)$. For each residue class $m \pmod{3^k}$ with $m \neq 0 \pmod 3$, we define for $y \geq 0$ the function

$$\phi_k^m(y) := \inf\{\pi_a^*(2^y a) : \ a \equiv m \pmod{3^j} \text{ and } a \text{ not in a cycle}\}.$$

This function is well defined because there always exists some $a \equiv m \pmod{3^k}$ not in a cycle.

This definition immediately implies that for $k \geq 2$ and all $m \pmod{3^k}$, $m \not\equiv 0 \pmod 3$, these functions satisfy the three properties:

(P1) (*Positivity*) *For all* $y \geq 0$,

$$\phi_k^m(y) \geq 1.$$

(P2) (*Monotonicity*) *For* $y \geq 0$,

$$\phi_k^m(y) \text{ is a nondecreasing function of } y.$$

(P3) (*Minimization*) *For* $m \in [3^{k-1}]$ *and all* $y \geq 0$,

$$\phi_{k-1}^m(y) = \min[\phi_k^m(y), \phi_k^{m+3^{k-1}}(y), \phi_k^{m+2\cdot 3^{k-1}}(y)].$$

It is easy to see that

$$\phi_k^m(y) = \phi_k^{2m}(y-1) \ \text{ if } m \equiv 1 \pmod 3, \tag{2.1}$$

hence it suffices to study $\phi_k^m(y)$ for $y \equiv 2 \pmod 3$. For convenience in what follows we let $[3^k]$ denote the set of congruence classes

$$[3^k] := \{m \pmod{3^k} : \ m \equiv 2 \pmod 3\}. \tag{2.2}$$

The difference inequality system of Krasikov [4] can be put in the following form.

Proposition 2.1. *Let $\alpha := \log_2 3 \simeq 1.585$. For each $k \geq 2$, the set of functions $\{\phi_k^m(y) : m \in [3^k]\}$ satisfy the following system $\mathcal{I}_k$ of difference inequalities, valid for all $y \geq 2$.*

(D1) If $m \equiv 2 \pmod 9$ then

$$\phi_k^m(y) \geq \phi_k^{4m}(y-2) + \phi_{k-1}^{(4m-2)/3}(y+\alpha-2). \tag{2.3}$$

(D2) If $m \equiv 5 \pmod 9$ then

$$\phi_k^m(y) \geq \phi_k^{4m}(y-2). \tag{2.4}$$

(D3) If $m \equiv 8 \pmod 9$ then

$$\phi_k^m(y) \geq \phi_k^{4m}(y-2) + \phi_{k-1}^{(2m-1)/3}(y+\alpha-1) \tag{2.5}$$

In these inequalities the functions $\phi_{k-1}^m(y)$ are defined by

$$\phi_{k-1}^m(y) := \min[\phi_k^m(y), \phi_k^{m+3^{k-1}}(y), \phi_k^{m+2\cdot 3^{k-1}}(y)]. \tag{2.6}$$

Proof. This follows from [4, Lemma 4], and appears in [2, Prop. 2.1]. ∎

We regard the system $\mathcal{I}_k$ of inequalities as expressed entirely in terms of the functions $\{\phi_k^m(y) : m \in [3^k]\}$, by using the minimum formulas (2.6). In that case all functions appearing are of the form $\phi_k^m(y+\beta_j)$ for various real numbers β_j. If $\beta_j \geq 0$ we call such a term *advanced*, while if $\beta < 0$ we call such a term *retarded*, since the terms have advanced arguments and retarded arguments respectively, in terms of the "time" variable y.

As mentioned earlier, Applegate and Lagarias [2] associated to $\mathcal{I}_k$ various auxiliary linear programs $L_k^{\mathcal{D}}(\lambda)$ depending on a parameter $\lambda > 1$; strictly positive feasible solutions for admissible linear programs for a given λ lead to exponential lower bounds for the functions $\phi_k^m(y) \geq c_0\lambda^y$. In this paper we study a linear program family, $L_k^{NT}(\lambda)$ which is similar in spirit [1] but not quite of the form used in [2]. However it is equivalent to the system L_λ^{NT} in [2] in the sense that matters for obtaining lower bounds, namely that $L_k^{NT}(\lambda)$ has a feasible solution for λ if and only if L_λ^{NT} has a strictly positive feasible solution for the same λ, as we show in §6.

[1] The linear program family $L_k^{NT}(\lambda)$ is a modification of L_λ^{NT} in [2]. It differs in having a different objective function variable, in minimizing rather than maximizing, and in having certain nonnegativity constraints modified to make them strictly positive.

The linear program family $L_k^{NT}(\lambda)$ is as follows.

$$L_k^{NT}(\lambda): \text{ Minimize } \; C_k^{max} \tag{2.7}$$

subject to:

(L0) *For all* $m \in [3^k]$,

$$1 \leq c_k^m \leq C_k^{max} \tag{2.8}$$

(L1) *For all* $m \in [3^k]$ *with* $m \equiv 2 \pmod 9$,

$$c_k^m \leq c_k^{4m} \lambda^{-2} + c_{k-1}^{\frac{4m-2}{3}} \lambda^{\alpha-2}\ . \tag{2.9}$$

(L2) *For all* $m \in [3^k]$ *with* $m \equiv 5 \pmod 9$,

$$c_k^m \leq c_k^{4m} \lambda^{-2}\ . \tag{2.10}$$

(L3) *For all* $m \in [3^k]$ *with* $m \equiv 8 \pmod 9$,

$$c_k^m \leq c_k^{4m} \lambda^{-2} + c_{k-1}^{\frac{2m-1}{3}} \lambda^{\alpha-1}\ . \tag{2.11}$$

(L4) *For all* $m \in [3^k]$,

$$c_{k-1}^m \leq c_k^m\ , \tag{2.12}$$

$$c_{k-1}^m \leq c_k^{m+3^k}\ , \tag{2.13}$$

$$c_{k-1}^m \leq c_k^{m+2\cdot 3^k}\ . \tag{2.14}$$

Note that the inequality signs in (L1)–(L3) go in the opposite direction from that in the difference inequalities $\mathcal{I}_k$, while (L4) goes in the same direction.

We call the variables $\{c_k^m : m \in [3^k]\}$ *principal variables* in $L_k^{NT}(\lambda)$, and the variables $\{c_{k-1}^m : m \in [3^{k-1}]\}$ *auxiliary variables*; the remaining variable C_k^{max} is the objective function variable. The objective function variable itself plays no role in determining feasibility of the linear program; the inequalities it appears in can always be satisfied by setting it equal to the maximum of the principal variables. If this linear program has any feasible solution, then this solution may be rescaled by a multiplicative constant so that $\min\{c_k^m\} = 1$, while decreasing

C_{max}, hence any optimum value of this linear program will have $\min\{c_k^m\} = 1$. Given a feasible solution, set

$$\bar{c}_{k-1}^m := \min\{c_k^m,\ c_k^{m+3^{k-1}},\ c_k^{m+2\cdot 3^{k-1}}\}. \tag{2.15}$$

The inequalities (D4) say that $c_{k-1}^m \le \bar{c}_{k-1}^m$. There are no lower bounds imposed on the auxiliary variables c_{k-1}^m, but given any feasible solution, there exists a positive feasible solution with the same principal variables and with auxiliary variables

$$c_{k-1}^m = \bar{c}_{k-1}^m \ge 1.$$

Indeed (D4) still holds for this choice of auxiliary variables and the remaining inequalities (D1)-(D3) stay the same or weaken.

The linear program $L_k^{NT}(\lambda)$ encodes advanced variables, and the theorems in [2] do not apply to it. Conjecture 4.1 of [2] asserts that the largest value of λ for which $L_k^{NT}(\lambda)$ has a positive feasible solution should give the largest possible exponential lower bound for positive, monotone functions Φ_k satisfying $\mathcal{I}_k$. Our main result is that $L_k^{NT}(\lambda)$ gives legitimate lower bounds for positive solutions for such functions $\phi_k^m(y)$.

Theorem 2.2. *Let $1 \le \lambda \le 2$ be such that the linear program $L_k^{NT}(\lambda)$ has a feasible solution with principal variables $\{c_k^m : m \in [3^k]\}$. Then for and all $m \in [3^k]$ and all $y \ge 0$,*

$$\phi_k^m(y) \ge \Delta_1 \cdot c_k^m \lambda^y, \tag{2.16}$$

in which

$$\Delta_1 := \frac{1}{4 \max\{c_k^m : m \in [3^k]\}}. \tag{2.17}$$

We believe that this result gives the largest exponential-type lower bound that can be extracted from the difference inequalities $\mathcal{I}_k$, for reasons given at the end of §6. However we have no rigorous proof of this assertion.

Theorem 2.2 is established as follows. In §3 we show that there exists a sequence of back substitutions of the difference inequalities into themselves that results in a difference inequality system from which all advanced variables have been eliminated. This results in a new system of difference inequalities $\mathcal{I}_k(EL)$. We show that all solutions ϕ_k^m of $\mathcal{I}_k$ which possess the positivity and monotonicity properties (P1) and (P2) will also be solutions of $\mathcal{I}_k(EL)$.

In §4 we consider linear programs. To each difference inequality system $\mathcal{D}$ (of a specified kind) we associate in a strictly deterministic way an auxiliary linear program family $L^{\mathcal{D}}(\lambda)$. Let $L_k^{EL}(\lambda)$ denote the linear program family attached to $\mathcal{I}_k(EL)$. The main result of §4 is the deduction that if the linear program $L_k^{NT}(\lambda)$ has a positive feasible solution with principal variables $\{c_k^m : m \in [3^k]\}$, then the linear program $L_k^{EL}(\lambda)$ with the same value of λ also has a positive feasible solution with the same principal variable values.

In §5 we show that any difference inequality system $\mathcal{D}$ in which only retarded variables appear has the property that positive feasible solutions to the auxiliary linear program $L^{\mathcal{D}}(\lambda)$ for fixed λ yields lower bounds of the form (2.16); the proof is similar to [2, Theorem 2.1]. It immediately follows that we get such lower bounds from the linear program family $L_k^{EL}(\lambda)$. We then prove Theorem 2.2, by combining this result with the main result of §4.

In §6 we present taxonomic data on the derived systems $L_k^{EL}(\lambda)$ for $2 \leq k \leq 5$ and information on positive feasible solutions the system $L_k^{NT}(\lambda)$ for $2 \leq k \leq 11$, computed by David Applegate, which yield the lower bound $\pi_a(x) \geq x^{0.84}$ for all sufficiently large x. The results of §4 imply that the linear program family $L_k^{EL}(\lambda)$ might conceivably give better exponential lower bounds than are obtainable from the linear program family $L_k^{NT}(\lambda)$. Numerical experiments show this is not the case for $2 \leq k \leq 5$; here $k = 5$ was the limit of computability for the system $L_k^{EL}(\lambda)$.

3. Eliminating Advanced Variables

We describe a recursive back-substitution procedure to eliminate "advanced" terms of the inequality system $\mathcal{I}_k$. We view the inequality system $\mathcal{I}_k$ as expressed entirely in terms of functions $\phi_k^m(y+\beta)$ by replacing each term involving any variable $\phi_{k-1}^{m'}(y+\beta')$ by the minimization expression on the right side of (2.6) in terms of ϕ_k^m functions.

We start with a single inequality (D3) of the system $\mathcal{I}_k$ associated to a fixed $m \in [3^k]$, $m \equiv 8 \pmod 9$, and perform a recursive back-substitution process of the inequalities $\mathcal{I}_k$ into its right-hand side. At the lth-stage of this process we will have an inequality $I_k^m(l)$ whose left side is $\phi_k^m(y)$ and whose right side is a nested series of minimizations of various functions

$\phi_k^{m'}(y+\beta')$. The step from $I_k^m(l)$ to $I_k^m(l+1)$ has two substeps. First, one picks an advanced term $\phi_k^{m'}(y+\beta')$, $\beta' \geq 0$ appearing in the right side of $I_k^m(l)$ and replaces it with the right side of the inequality of the system $\mathcal{I}_k$ having left side $\phi_k^{m'}(y')$, where we take $y' = y + \beta'$. (This is called "splitting" a term in [2].) A new minimization term may appear in this process, which contains three terms

$$\phi_k^{m'}(y+\beta''), \quad \phi_k^{m'+3^{k-1}}(y+\beta''), \quad \phi_k^{m'+2\cdot 3^{k-1}}(y+\beta''') \,. \tag{3.1}$$

The second substep in obtaining $I_k^m(l+1)$ is to apply a deletion rule described below, which, if $\beta'' \geq 0$, may remove up to two of these terms. The resulting inequality after the deletion substep is $I_k^m(l+1)$.

At each stage in this process the inequality $I_k^m(l)$ has $\phi_k^m(y)$ on its left side and a sum of nested minimization terms on its right side, involving various functions $\phi_k^m(y+\beta_j)$; it will have each $\beta_j \geq -2$, because we will only substitute for terms $\phi_k^m(y+\beta_j)$ with $\beta_j \geq 0$, and the formulas (D1)–(D3) produce new terms $\phi_k^{m'}(y+\beta_j')$ which have $\beta_j' \geq \beta_j - 2$. The structure of the right side of an inequality $I_k^m(l)$ is described by a directed rooted labelled tree $\mathcal{T}_k^m(l)$, in which the root mode is labelled with the left side $\phi_k^m(y)$ of the original inequality, each node is either a p-node (for "principal") or an m-node (for "minimization"). The initial tree for the inequality I_k^m for an $m \in [3^k]$ with $m \equiv\ 8 \pmod 9$ is pictured in Figure 1.

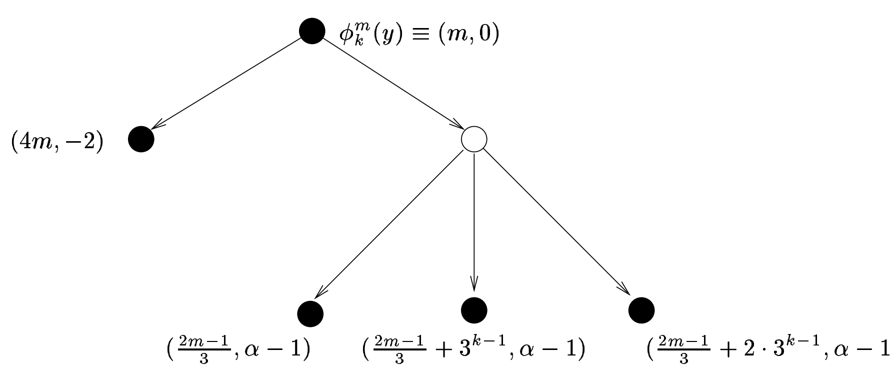

Figure 1: Rooted tree for inequality (D3).

Here p-nodes are indicated by solid points and m-nodes by circled points. Each p-node is labelled by data (m, β) specifying the function $\phi_k^m(y+\beta)$ with m viewed $\pmod{3^k}$, while each m-node is assigned the label (m, β) of the p-node of which it is a child. The root node is a p-node and has label $(m, 0)$. The inequality $I_k^m(l)$ is uniquely specified by the tree $\mathcal{T}_k^m(l)$ and vice-versa; the root node specifies the left side $\phi_k^m(y)$ of the inequality $I_k^m(l)$, leaf nodes specify functions appearing in the right side, and the internal tree structure specifies the nested sequence of additions and minimizations comprising this right side of the inequality.

A step from $\mathcal{T}_k^m(l)$ to $\mathcal{T}_k^m(l+1)$ consists of picking a leaf node with label (m', β') which has $\beta' \geq 0$ and changing the tree in the following two substeps. First we attach to the leaf node (as root node) the directed tree associated to the formula (D1)–(D3) of $\phi_k^{m'}(y')$ with variable y' and then changing variables $y' = y + \beta'$. We term this "splitting" the leaf node, following [2]. The tree $\tilde{\mathcal{T}}_k^m(l+1)$ that results has a new p-node labelled $(4m', \beta' - 2)$, and may or may not have a new m-node with three new leaf nodes (3.1) depending from it. If there is no m-node this tree will be $\mathcal{T}_k^m(l+1)$. Second, if there is a new m node, we apply the *deletion rule* given below to $\tilde{\mathcal{T}}_k^m(l+1)$ to remove some (possibly empty) subset of the three leaves in a m-term.

Deletion Rule. *Suppose that a leaf node $\phi_k^{m'}(y + \beta')$, has label (m', β') with $\beta' \geq 0$. Delete this leaf node if on the directed path from the root node $\phi_k^m(y)$ to it there is a p-node having label (m, β) with $m = m'$ and $\beta < \beta'$. If this rule removes all three leaves, then delete the m-node above it as well.*

After the deletion rule is applied to $\tilde{\mathcal{T}}_k^m(l+1)$, the tree that results is $\mathcal{T}_k^m(l+1)$, and the inequality correspoding to it is $I_k^m(l+1)$. All leaf nodes on the new tree $\mathcal{T}_k^m(l+1)$ are p-nodes, so the process can continue. We justify the deletion rule in Theorem 3.2 below. Actually it never occurs that all three leaves are deleted, but we do not need this result in what follows.

The back-substitution process is not completely specified, in that one has the freedom to choose to split any leaf node carrying an advanced term. However the order of splitting does not matter as the following result asserts.

Theorem 3.1. *Let $k \geq 2$, and take $m \in [3^k]$ with $m \equiv 8 \pmod 9$. The back-substitution process applied to $\phi_k^m(y)$ halts after a finite number of steps at an inequality $I_k^m(l)$ having no advanced terms on its right side. The number of steps l and the final inequality $I_k^m(l)$ are independent of the order in which advanced terms are split; let $I_k^m(EL)$ denote this final inequality.*

Proof. We first show that the back-substitution procedure always halts. We suppose not, and obtain a contradiction. Let $\mathcal{T}_l \equiv \mathcal{T}_k^m(l)$ denote the rooted labelled tree associated to the inequality $I_k^m(l)$ for $l = 1, 2, \ldots$. Then we have an infinite sequence of trees, each containing the last as a subtree having the same root, and the process defines an infinite limiting tree $\mathcal{T}_\infty$. Without loss of generality we can suppose that $\mathcal{T}_\infty$ has the property that in it all nodes that can be split are split, if necessary by doing additional splittings of any advanced nodes that were missed, using transfinite induction. By Konig's infinity lemma there is an infinite directed path in $\mathcal{T}_\infty$ starting from the root. Along that path there is some residue class $m' \in [3^k]$ that occurs as a label infinitely often. Let $\{(m', \beta_j) : \ j = 1, 2, \ldots\}$ be the successive labels of the p-nodes on this path having residue class $m' \pmod{3^k}$, starting from the root. We must have each $\beta_j \geq 0$ (or the process halts) and also

$$\beta_1 > \beta_2 > \beta_3 > \cdots \ , \tag{3.2}$$

because the deletion rule would have removed the p-node labelled (m', β_j) if $\beta_j \geq \beta_i$ for some $j > i$.

The tree $\mathcal{T}_\infty$ has a recursive self-similar structure, using the fact that all nodes that could be split were split. Consider the subtree $\mathcal{T}_\infty[j]$ grown starting from the root node $\phi_k^{m'}(y + \beta_j)$ along this chain, using the new variable $y_j = y + \beta_j$. These subtrees are all identical, and $\mathcal{T}_\infty[2]$ is obtained from $\mathcal{T}_\infty[1]$ by shifting the argument of y by $\delta = \beta_2 - \beta_1 > 0$. The isomorphism of $\mathcal{T}_\infty[2]$ and $\mathcal{T}_\infty[1]$ identifies $\mathcal{T}_\infty[j]$ with $\mathcal{T}_\infty[j-1]$, and therefore, by induction on $j \geq 2$, we obtain $\beta_j - \beta_{j-1} = \delta$. Thus $\beta_j = \beta_1 + (j-1)\delta$ for all $j \geq 2$, hence $\beta_j < 0$ for sufficiently large j, which contradicts all $\beta_j \geq 0$.

The back-substitution process halts at a unique tree, regardless of the order leaf nodes are split, because the back-substitution process on a given leaf node $\mathbf{v}$ does not depend on any other leaf nodes, but only on the path from the root node to $\mathbf{v}$. One grows out all leaf nodes until they halt, and the total number of steps l until halting is independent of the order of growth. ∎

Theorem 3.2. *Let $\mathcal{I}_k(EL)$ denote the difference inequality system consisting of the inequalities* (D1), (D2) *of $\mathcal{I}_k$ plus the complete set of inequalities $\{I_k^m(EL) : m \in [3^k],\ m \equiv 8 \pmod{9}\}$. If $\Phi_k = \{\phi_k^m(y) : m \in [3^k]\}$ is any set of functions in which each $\phi_k^m(y)$ is strictly* *positive and nondecreasing on $\mathbb{R}_{\geq 0}$ and satisfies the inequality system $\mathcal{I}_k$ for all $y \geq 2$, then Φ_k also satisfies the inequalities $\mathcal{I}_k(EL)$ for all $y \geq 2$.*

Proof. It suffices to show that if the set $\Phi_k := \{\phi_k^m(y) : m \in [3^k]\}$ of positive nondecreasing functions on $\mathbb{R}^+ = \{y \geq 0\}$ satisfies $\mathcal{I}_k$ for all $y \geq 2$, then they satisfy each inequality $I_k^m(l)$ for each $l \geq 1$, for all $y \geq 2$.

We prove, by induction on $l \geq 1$, that the set Φ_k satisfies $\mathcal{T}_k^m(l)$. The base case $l = 1$ holds because $\mathcal{T}_k^m(l)$ has only one internal p-node, its root node, and the corresponding inequality $I_k^m(1)$ is a member of $\mathcal{I}_k$. Now suppose the induction hypothesis holds for $\mathcal{T}_k^m(l)$, and consider $\mathcal{T}_k^m(l+1)$. To obtain $\mathcal{T}_k^m(l+1)$ we first split a leaf of $\mathcal{T}_k^m(l)$ to obtain a tree $\tilde{\mathcal{T}}_k^m(l+1)$ and then, if a new m-node was added, we apply the deletion rule to the three vertices of that m-node. The splitting procedure yielding $\tilde{\mathcal{T}}_k^m(l+1)$ substitutes an inequality of $\mathcal{I}_k$, hence Φ_k automatically satisfies $\tilde{\mathcal{T}}_k^m(l+1)$.

Consider the deletion step, applied to the three leaf-node labels of $\tilde{\mathcal{T}}_k^m(l+1)$ inside the min-term

$$f(y) := \min[\phi_k^{m'}(y+\beta'), \phi_k^{m'+3^{k-1}}(y+\beta'), \phi^{m'+2\cdot 3^{k-1}}(y+\beta')] \ . \tag{3.3}$$

The leaf node (m', β') is to be deleted if earlier in its directed path from the root appears a p-node with label $\phi_k^{m'}(y+\beta)$ with $\beta < \beta'$.

To justify the deletion rule, note that the inequality associated to each tree $\mathcal{T}_k^m(l)$ for fixed functions Φ_k and a fixed value $y \geq 2$, can be written as a sum of terms corresponding to a subset of leaves of the tree which are specified by choosing one of the terms in each min-term that attains the minimum. (This choice is usually unique once the functions Φ_k and the value y are specified, unless two terms in a min-term have equal values.) We call this set of leaves a *critical assignment*, the leaves in it *critical leaves*, and the set of paths to these leaves *critical paths*.

Claim. *To each internal p-vertex $\mathbf{v}$ of the tree with label $\phi_k^m(y+\beta)$, and for each fixed value of $y \geq 2$, exactly one of two possibilities occurs.*

(a) There are no critical assignments $\mathcal{A}$ having a critical path passing through $\mathbf{v}$.

(b) There is at least one critical assignment $\mathcal{A}$ with a path passing through $\mathbf{v}$. For any such

assignment

$$\phi_k^m(y+\beta) \geq \sum_{\mathbf{w}\in\mathcal{A}_{\mathbf{v}}} \phi_k^{m(\mathbf{w})}(y+\beta(\mathbf{w})) \ , \tag{3.4}$$

where $\mathcal{A}_{\mathbf{v}}$ *denotes the set of critical leaves in* $\mathcal{A}$ *whose paths pass through* $\mathbf{v}$.

Warning: For fixed k, m, β which of case (a) or (b) occurs depends on the value of y. The key content of the claim is the inequality (3.4) given in case (b).

We will prove the claim by induction on l, and justify the deletion rule at the same time. Now (3.4) holds for the base case $l = 1$ where the only internal p-node is the root node, and (3.4) is then an inequality in $\mathcal{I}_k$. We assume it holds for $\mathcal{T}_k^m(l)$ and wish to prove it for $\mathcal{T}_k^m(l+1)$. First of all, the relations (a), (b) hold for $\tilde{\mathcal{T}}_k^n(l+1)$. They hold for internal p-nodes inherited from $\mathcal{T}_k^m(l)$, because we have back-substituted $\mathcal{I}_k$ on the right side of (3.4). We have added one new internal p-node $\mathbf{v}^*$, the one that was split; let its label be (m^*, β^*). For the new internal p-node $\mathbf{v}^*$ condition (3.4) in (b) directly expresses the $\mathcal{I}_k$ inequality substituted; thus (a) and (b) hold for this node in $\tilde{\mathcal{T}}_k^n(l+1)$.

We now consider $\mathcal{T}_k^n(l+1)$. If the deletion rule does nothing then the induction step holds for $\mathcal{T}_k^m(l+1)$. The deletion rule can be applied only if $\mathbf{v}^*$ has an m-node depending on it, and this occurs only when $\beta^* \geq 0$. We call a vertex $\mathbf{v}$ of $\tilde{\mathcal{T}}_k^m(l+1)$ *totally non-critical* if no critical path passes through it, for any critical assignment $\mathcal{A}$ for any $y \geq 2$; that is, case (a) holds for $\mathbf{v}$ for all $y \geq 2$. We can safely delete all totally non-critical vertices in $\tilde{\mathcal{T}}_k^m(l+1)$, and property (b) will still hold for the resulting tree $\mathcal{T}'$. (The property that a vertex in a tree is totally non-critical is hereditary in the sense that all vertices below a totally non-critical vertex are also totally non-critical.)

We now show that, for those sets of functions Φ_k that are positive and monotone, all vertices removed by the deletion rule are totally non-critical. Suppose the deletion rule applies to the leaf vertex $\mathbf{w}$ with label (m', β') of $\tilde{\mathcal{T}}_k^m(l+1)$, and let $\mathbf{v}$ be a p-vertex on its directed path that has label (m', β) with $\beta \leq \beta'$. To show that $\mathbf{w}$ is totally non-critical, we argue by contradiction. Suppose not, so that there is some $y \geq 2$ and a critical assignment $\mathcal{A}$ containing $\mathbf{w}$ as a critical leaf. Formula (3.4) of (b) applies to $\mathbf{v}$ to give $\phi_k^{m'}(y+\beta) \geq \sum_{(\tilde{m},\tilde{\beta})\in\mathcal{A}_{\mathbf{v}}} \phi_k^{\tilde{m}}(y+\tilde{\beta})$.

We deduce

$$\phi_k^{m'}(y+\beta) \geq \phi_k^{m'}(y+\beta')\ , \tag{3.5}$$

because $\phi_k^{m'}(y+\beta')$ is the contribution of $\mathbf{w} \in \mathcal{A}_{\mathbf{v}}$. However there is at least one more critical path in the sum $\mathcal{A}_{\mathbf{v}}$; namely one which passes through the last p-vertex $\mathbf{v}^*$ in the path before $\mathbf{w}$, and goes to its direct p-node descendant with label $(4m^*, \beta^* - 2)$. Since $\beta^* \geq 0$ and $y \geq 2$ we have $\phi_k^{4m^*}(y + \beta^* - 2) > 0$ by positivity and monotonicity of Φ_k. We conclude that (3.5) can be sharpened to strict inequality

$$\phi_k^{m'}(y+\beta) > \phi_k^{m'}(y+\beta')\ . \tag{3.6}$$

Since $\beta < \beta'$, this violates monotonicity of Φ_k, the desired contradiction.

Thus, the vertices removed by the deletion rule are totally non-critical. It follows that for the resulting tree $\mathcal{T}_k^m(l+1)$, the criteria (a), (b) and (3.4) hold for all p-vertices, for the functions Φ_k, for all $y \geq 2$. This completes the claim's induction step, and proves the claim.

Now we may apply (3.4) to the root vertex $\mathbf{v}$ for all critical assignments $\mathcal{A}$ for all $y \geq 2$ is equivalent to saying that the Φ_k satisfy the inequality $I_k^m(l+1)$ associated to $\mathcal{T}_k^m(l+1)$ for all $y \geq 2$. This completes the main induction step. ∎

Remark. The inequality system $\mathcal{I}_k(EL)$ involves nested minimization to a depth $d(k)$ which grows exponentially with k. The exponential growth occurs because the deletion rule requires a node with label (m', β') to lie on a path containing another node (m, β) with $m \equiv m' (\bmod\ 3^k)$, and if the values m are randomly distributed on the path one expects the path to have length comparable to 3^k. We present statistics in Table 1 on the size of this inequality system $\mathcal{I}_k(EL)$ for $2 \le k \le 5$, computed by D. Applegate. We measure the size in two ways: the depth of nested minimizations, and the total of the number of terms that appear in such an inequality. The data is for the term $\phi_k^m(y)$ that had the largest expansion under the elimination procedure.

k	depth	# (literals)
2	3	8
3	10	84
4	41	12829
5	> 226	$> 10^9$

Table 1: Statistics on $\mathcal{I}_k(EL)$ Inequalities

4. Linear Programs

We associate to a general difference inequality system $\mathcal{D}_k$ (of a sort described below) a family of linear program $L_k^{\mathcal{D}}(\lambda)$, as follows. We suppose that $\mathcal{D}_k$ consists of inequalities $\{D_k^m : [m] \in 3^k\}$ in which each inequality D_k^m is described by a rooted labelled tree $\mathcal{T}_k^m$ of the type considered in §3, involving variables $\{c_k^m : [m] \in 3^k\}$. The linear program has the basic form:

$$L_k^{\mathcal{D}}(\lambda): \text{ Minimize } C_{max} \tag{4.1}$$

subject to, for all $m \in [3^k]$,

$$1 \le c_k^m \le C_k^{max},$$

together with all inequalities associated to each tree $\mathcal{T}_k^m$ as specified below.

The LP-inequality system associated to a given tree $\mathcal{T}$ involves the principal variables $\{c_k^m : m \in [3^k]\}$ and certain auxiliary variables $\{a_{\mathbf{v}} : \mathbf{v} \text{ an } m\text{-vertex of } \mathcal{T}\}$. These auxiliary variables are distinct for different trees $\mathcal{T}_k^m$. We associate to each node $\mathbf{w}$ the label $(m(\mathbf{w}), \beta(\mathbf{w}))$ which consists of a *residue class* $m(\mathbf{w})$ and a *weight* $\beta(\mathbf{w})$. For a p-node $\mathbf{w}$ these labels are determined by its associated function $\phi_k^{m(\mathbf{w})}(x + \beta(\mathbf{w}))$ with $m(\mathbf{w})$ determined $(\bmod\ 3^k)$. For an m-node it is taken from the node function of any of its children, where we view $m(\mathbf{w})$ (mod 3^{k-1}) in this case, noting that $m(\mathbf{w})$(mod 3^{k-1}) is the same for all the child nodes. To specify the inequalities, we subdivide the tree $\mathcal{T}$ into levels: we say that a vertex $\mathbf{w}$ is at *m-depth d* if there are exactly $d-1$ internal m-nodes on the path from the root node to $\mathbf{w}$ (not counting $\mathbf{w}$ itself). The LP inequalities associated to $\mathcal{T}$ are in one-one correspondence with the leaf nodes of $\mathcal{T}$. To each leaf node $\mathbf{w}$ we assign a rooted subtree $\mathcal{T}_{\mathbf{w}}$ which consists of:

(1) The terminal part of the path from the root node to the leaf node. If an m-node occurs on the path, then it consists of that part of the path from the final m-node to the leaf; if no m-node occurs then it is the entire path from the root. We denote this path $\mathcal{P}_{\mathbf{w}}$ and call its top node the $\mathbf{w}$-root node. Every vertex on $\mathcal{P}_{\mathbf{w}}$ is a p-node except possibly the $\mathbf{w}$-root node.

(2) All other children of any p-node on the path $\mathcal{P}_{\mathbf{w}}$. These other children are all m-nodes. A typical subtree $\mathcal{T}_{\mathbf{w}}$ is pictured in Figure 2.

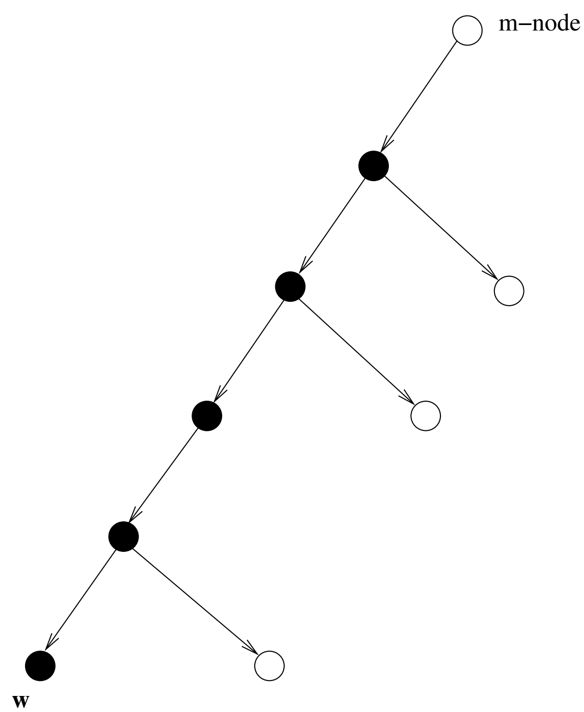

Figure 2: Subtree $\mathcal{T}_{\mathbf{w}}$ of the leaf node $\mathbf{w}$ (m-nodes are circled)

All the edges of $\mathcal{T}$ are partitioned among the $\mathcal{P}_{\mathbf{w}}$ and each $\mathcal{P}_{\mathbf{w}}$ contains exactly one leaf node. The trees in this partition are also in one-one correspondence with: either the root node

$\mathbf{v}$ or a pair $(\mathbf{v}, \mathbf{v}')$ consisting of an m-node $\mathbf{v}$ and one of its children $\mathbf{v}'$.

The LP-inequality associated to the unique leaf node $\mathbf{w}$ having no m-nodes on its path is of the form

$$c_k^m \le \lambda^{\beta(\mathbf{w})} c_k^{m(\mathbf{w})} + \sum_{\substack{\mathbf{v} \in \mathcal{T}_\mathbf{w} \\ m\text{-node}}} \lambda^{\beta(\mathbf{v})} a_\mathbf{v} \tag{4.2}$$

where $m = m(\mathbf{v}_0)$ for the root vertex $\mathbf{v}_0$. For all leaf nodes $\mathbf{w}$ such that $\mathcal{T}_\mathbf{w}$ has a node $\mathbf{v}_0$ as $\mathbf{w}$-root node, the associated LP-inequality is

$$\lambda^{\beta(\mathbf{v}_0)} a_{\mathbf{v}_0} \le \lambda^{\beta(\mathbf{w})} c_k^{m(\mathbf{w})} + \sum_{\substack{\mathbf{v} \in \mathcal{T}_\mathbf{w} \\ m\text{-node} \\ \mathbf{v} \neq \mathbf{v}_0}} \lambda^{\beta(\mathbf{v})} a_\mathbf{v} \ . \tag{4.3}$$

Note that the direction of this LP-inequality (4.2) for the root node is opposite to that of the $\phi_k^m(y)$-inequality.

For the original difference inequality system $\mathcal{I}_k$, the linear program $L_k^{\mathcal{I}}(\lambda)$ produced in this way is equivalent to $L_k^{NT}(\lambda)$ in the following sense: to every feasible solution of $L_k^{NT}(\lambda)$ with principal variables $\{c_k^m\}$ there corresponds a feasible solution to $L_k^{\mathcal{I}}(\lambda)$ with the same principal variables, and vice-versa. To see this, we note that $L_k^{NT}(\lambda)$ has auxiliary variables c_{k-1}^m, while $L_k^{\mathcal{I}}(\lambda)$ has auxiliary variables $a_\mathbf{v}$ in one-one correspondence with c_k^m for $m \equiv 2$ or $8 \pmod 9$; the LP-inequalities in $L_k^{\mathcal{I}}(\lambda)$ on these variables are equivalent to

$$a_\mathbf{v} \le \bar{c}_{k-1}^{\frac{4m-2}{3}} \quad \text{or} \quad a_\mathbf{v} \le \bar{c}_{k-1}^{\frac{2m-1}{3}} \ , \tag{4.4}$$

according as $m \equiv 2 \pmod 9$ or $m \equiv 8 \pmod 9$, respectively, where $\bar{c}_{k-1}^m := \min_{0 \le j \le 2}\{c_k^{m+j3^{k-1}}\}$. The correspondence between feasible solutions of $L_k^{NT}(\lambda)$ and $L_k^{\mathcal{I}}(\lambda)$ is obtained by setting

$$a_\mathbf{v} = \bar{c}_{k-1}^{\frac{4m-2}{3}} \quad \text{or} \quad a_\mathbf{v} = \bar{c}_{k-1}^{\frac{2m-1}{3}} \ , \tag{4.5}$$

according as $m \equiv 2 \pmod 9$ or $m \equiv 8 \pmod 9$, respectively.

We let $L_k^{EL}(\lambda)$ denote the family of linear programs associated to the derived inequality system $\mathcal{I}_k^m(EL)$ of Theorem 3.2.

Theorem 4.1. *Suppose for a given λ with $1 \le \lambda \le 2$ that the linear program $L_k^{NT}(\lambda)$ has a feasible solution with principal variables $\{c_k^m : m \in [3^k]\}$. Then the linear program $L_k^{EL}(\lambda)$ has a positive feasible solution with the same principal variables.*

Proof. We prove this by starting with the inequality system $\mathcal{D}_1 := \mathcal{I}_k$ and then successively producing inequality systems $\{\mathcal{D}_j : 1 \le j \le r\}$, in which $\mathcal{D}_{j+1}$ is obtained from $\mathcal{D}_j$ by a single back-substitution in one inequality, and ending at the final system $\mathcal{D}_r = \mathcal{I}_k(EL)$. For definiteness we choose to do the back-substitution procedure on each inequality I_k^m, for $m \in [3^k]\}$ in order until it halts, as guaranteed by Theorem 3.1, and go to the next m, in the order $m = 2, 5, 8, \ldots, 3^k - 1$.

We prove by induction on $j \ge 1$ that if $\{c_k^m : m \in [3^k]\}$ yields a feasible solution of $L_k^{NT}(\lambda)$, then these same principal variable values occur in some positive feasible solution of $L_k^{\mathcal{D}_j}(\lambda)$. The base case $j = 1$ holds because the linear program $L_k^{\mathcal{D}}(\lambda)$ agrees with $L_k^{NT}(\lambda)$; when we assign the auxiliary variables $a_{\mathbf{v}}$ the values (4.5) we obtain a positive feasible solution with the given $\{c_k^m\}$.

For the induction step, first note that in going from $\mathcal{D}_j$ to $\mathcal{D}_{j+1}$, we "split" one leaf vertex $\mathbf{w}$ of a particular tree $\mathcal{T}_k^m(l)$, leaving all other trees alone, and then perform a deletion operation. The vertex $\mathbf{w}$ being a p-node, has associated value $m(\mathbf{v}) \pmod{3^k}$. We let $\tilde{\mathcal{D}}_{j+1}$ denote the inequalities resulting from the splitting operation before the deletion step. It suffices to show that $L_k^{\tilde{\mathcal{D}}_{j+1}}$ has a feasible solution with the same principal variables, for the deletion step merely deletes linear programming inequalities, which preserves feasible solutions. The splitting step changes exactly one of the inequalities in $L_k^{\mathcal{D}_j}$; if it adds a new m-vertex $\mathbf{v}$, then it adds on up to three new inequalities, each involving the new auxiliary variable $a_{\mathbf{v}}$ for the added m-vertex. The corresponding tree is updated to $\mathcal{T}_k^m(l+1)$.

Let $m' = m(\mathbf{w})$. If $m' \equiv 5 \pmod 9$, the unique LP-inequality containing the term $c_k^{m'}\lambda^{\beta(\mathbf{w})}$ corresponding to $\mathbf{w}$ on its right side, has this term replaced by that of a new leaf vertex $\mathbf{w}'$ with $m(\mathbf{w}') = 4m'$, $\beta(\mathbf{w}') = \beta(\mathbf{w}) - 2$ and $\mathbf{w}'$ has the same depth no $\mathbf{w}$; its new term is $c_k^{4m'}\lambda^{\beta(\mathbf{w})-2}$. However by hypothesis $\{c_k^m\}$ satisfies $L_k^{NT}(\lambda)$, hence it satisfies the inequality

$$c_k^{m'} \le c_k^{4m'}\lambda^{-2} .$$

Thus we obtain $c_k^{m'}\lambda^{\beta(\mathbf{w})} \le c_k^{4m'}\lambda^{\beta(\mathbf{w}-2}$, so the right side of the new inequality (4.2) or (4.3) is less binding than before, and the solution remains feasible. If $m' \equiv 2 \pmod 9$, the term $c_k^{m'}\lambda^{\beta(\mathbf{w})}$ is replaced with

$$c_k^{4m'}\lambda^{\beta(\mathbf{w})-2} + a_{\mathbf{v}}\lambda^{\beta(\mathbf{w})} ,$$

where $\beta(\mathbf{v}) = \beta(\mathbf{w}) + \alpha - 2$, and $L_k^{\tilde{\mathcal{D}}_{j+1}}$ has three new inequalities

$$a_{\mathbf{v}} \lambda^{\beta(\mathbf{v})} \leq c_k^{m(\mathbf{v})+j3^{k-1}} \lambda^{\beta(\mathbf{v})} \tag{4.6}$$

for $0 \leq j \leq 2$, with $m(\mathbf{v}) = \frac{4m(\mathbf{w})-2}{3}$. We may choose

$$a_{\mathbf{v}} = \bar{c}_{k-1}^{m(\mathbf{v})} := \min_{0 \leq j \leq 2} \{c_k^{m(\mathbf{v})+j3^{k-1}}\} \tag{4.7}$$

and satisfy (4.6); the fact that $\{c_k^m\}$ satisfies $L_k^{NT}(\lambda)$ gives

$$c_k^{m'} \lambda^{\beta(\mathbf{w})} \leq c_k^{4m'} \lambda^{\beta(\mathbf{w})-2} + c_{k-1}^{\frac{4m'-1}{3}} \lambda^{\beta(\mathbf{w})+\alpha_2} = c_k^{4m'} \lambda^{\beta(\mathbf{w})-2} - a_{\mathbf{v}} \lambda^{\beta(\mathbf{v})} .$$

Thus the right side of the equation is less binding than before, so remains feasible. The case $m' \equiv 8 \pmod 9$ is handled by similar reasoning to the case $m' \equiv 2 \pmod 9$, so feasibility is maintained in this case. The induction step follows.

The final case of the induction step gives the inequality system $\mathcal{I}_k(EL)$, and the theorem follows. ∎

5. Lower Bounds For Difference Inequalities

We obtain exponential lower bounds for systems of positive nondecreasing functions Φ_k satisfying difference inequalities $\mathcal{D}$ without advanced variables, using an associated linear program $L_k^{\mathcal{D}}$. The following result is similar in spirit to [2, Theorem 2.1].

Theorem 5.1. *Let $\Phi_k := \{\phi_k^m(y) : m \in [3^k]\}$ be a set of positive nondecreasing functions on $\mathbb{R}_+ = \{y : y \geq 0\}$. Suppose that Φ_k satisfies a system $\mathcal{D}$ of difference inequalities specified by a set of rooted labelled trees $\{\mathcal{T}_k^m : m \in [3^k]\}$, such that all inequalities contain no advanced variables on their right side. If the associated linear program $L_k^{\mathcal{D}}(\lambda)$ for $\lambda > 1$ has a positive feasible solution with principal variables $\{c_k^m\}$ then, for all $m \in [3^k]$,*

$$\phi_k^m(y) \geq \Delta c_k^m \lambda^y \ , \quad \textit{for all} \ \ y \geq 0, \tag{5.1}$$

with

$$\Delta := \lambda^{-\nu} \frac{\min\{\phi_k^m(0)\}}{\max\{c_k^m\}}, \tag{5.2}$$

and ν is the largest backward time-shift of a variable in $\mathcal{D}$.

Proof. Suppose that the set of functions Φ_k satisfies the system $\mathcal{D} := \{D_k^m : m \in [3^k]\}$ of difference inequalities. Set

$$\mu := \min\{\beta : \phi_k^{m'}(y-\beta) \text{ appears on right side of some } D_k^m\}$$

and

$$\nu := \max\{\beta : \phi_k^{m'}(y-\beta) \text{ appears on right side of some } D_k^m\} .$$

The hypothesis of no advanced variables in $\mathcal{D}$ means that $\mu > 0$. Now the inequalities (5.1) hold for all $m \in [3^k]$, on the initial interval $[0, \nu]$, since the definition of Δ gives

$$\phi_k^m(y) \geq \phi_k^m(0) \geq \Delta \max\{c_k^m\}\lambda^\nu \geq \Delta c_k^m \lambda^y \text{ for } \ y \in [0, \nu], \tag{5.3}$$

using the monotonicity and inequality properties of $\phi_k^m(y)$.

We now prove that (5.1) holds for all $m \in [3^k]$ on the interval $y \in [0, \nu + j\mu]$ by induction on $j \geq 0$. It holds for the base case $j = 0$ by (5.3).

For the induction step, suppose that (5.1) holds for j and we are to prove it for $j + 1$. It suffices to consider a given $y \in [\nu + j\mu, \nu + (j+1)\mu]$. The induction step consists, schematically, of showing

$$\begin{aligned}
\phi_k^m(y) &\geq \sum_{D_k^m(EL)} \text{nested-min}[\phi_k^{m'}(y+\beta')] && (5.4)\\
&\geq \sum_{\mathcal{T}_k^m(EL)} \text{nested-min}[c_k^{m'}\lambda^{y+\beta'}] && (5.5)\\
&\geq \Delta c_k^m \lambda^y . && (5.6)
\end{aligned}$$

Here (5.4) represents schematically the inequality $D_k^m(EL)$, with the right side actually being a nested series of minimizations. Each function $\phi_k^{m'}(y + \beta')$ that appears on the right side of (5.4) has $-\nu \leq \beta' \leq -\mu$, hence

$$0 \leq j\mu \leq y + \beta' \leq \nu + j\mu,$$

so the induction hypothesis applies to each such term.

The induction hypothesis gives

$$\phi_k^{m'}(y + \beta') \geq \Delta c_k^{m'} \lambda^{y+\beta'} = \Delta \lambda^y (c_k^{m'} \lambda^{\beta'}).$$

Substituting these inequalities in (5.4) term by term yields the right side of (5.5), because the nested minimization on the right side of (5.4) involves only the operations of addition and minimization and these operations are both monotone in each variable appearing in them; also the structure $\mathcal{T}_k^m(EL)$ in (5.5) is the tree structure of the inequality $D_k^m(EL)$. Now let $f(y)$ represent the value of the right side of (5.5) as a function of y. Each minimization on the right side of (5.5) corresponds to a m-vertex $\mathbf{v}$ of $\mathcal{T}_k^m(EL)$; we let $f_{\mathbf{v}}(y)$ equal the value of this minimization expression as a function of y. Next we can apply the inequalities in $L_k^{EL}(\lambda)$ in a suitable order to prove that

$$f_{\mathbf{v}}(y) \geq \Delta\lambda^y (a_{\mathbf{v}}\lambda^{\beta(\mathbf{v})}).$$

for all m-vertices; the order starts with the innermost minimization and works outward. At the last step we reach the root vertex and obtain

$$f(y) \geq \Delta\lambda^y c_k^m \lambda^{\beta(\mathbf{w}_0)} = \Delta c_k^m \lambda^y,$$

since $\beta(\mathbf{w}_0) = 0$. This gives the right side of (5.6). Since this holds for all $k \in [3^m]$, this completes the induction step. ∎

We now prove the main Theorem 2.2 by combining the results of §3–§5.

Proof of Theorem 2.2. Theorem 3.2 shows that any set of positive nondecreasing functions Φ_k that satisfies the inequality system $\mathcal{I}_k$ also satisfies the derived inequality system $\mathcal{I}_k(EL)$ which has inequalities with no advanced variables on their right sides. The family of linear programs associated to this inequality system in §4 is denoted $L_k^{EL}(\lambda)$.

Suppose now that for a given $\lambda > 1$ the inequality system $L_k^{NT}(\lambda)$ has a feasible solution with principal variables $\{c_k^m : m \in [3^k]\}$. Theorem 4.1 established that the linear program $L_k^{EL}(\lambda)$ has a positive feasible solution with the same principal variables and the same value of λ.

Theorem 5.1 then applies to the system $\mathcal{I}_k(EL)$ to show that any positive feasible solution of $L_k^{EL}(\lambda)$ yields the bounds, for all $m \in [3^k]$,

$$\phi_k^m(y) \geq \Delta c_m^k \lambda^y \text{ for all } \; y \geq 0, \tag{5.7}$$

with

$$\Delta := \lambda^{-\nu} \frac{\min\{\phi_k^m(0)\}}{\max\{c_k^m\}} \tag{5.8}$$

where ν is the largest backwards timeshift.

For the system $\Phi_k = \{\phi_k^m(y) : m \in [3^k]\}$ coming from the $3x+1$ problem, we have by (P1) that $\phi_k^m(0) \geq 1$. We also have $\lambda \leq 2$ and the maximum retarded term $\nu \leq 2$. Thus we have

$$\Delta \geq \Delta_1 = \frac{1}{4 \max\{c_k^m\}}$$

which, with (5.7), implies the desired bound (2.16). ∎

Remark. Theorem 2.2 has the counterintuitive feature that iterating the inequalities seems potentially to strengthen, rather than weaken, the resulting exponential lower bound. It allows the possibility that the linear program $L_k^{EL}(\lambda)$ has a positive feasible solution for a larger value of λ than is obtainable using the original linear program family $L_k^{NT}(\lambda)$. However we believe this cannot occur, and that the exponent obtained from $L_k^{NT}(\lambda)$ is the largest possible for positive monotone solutions to the original difference inequalities $\mathcal{I}_k$. We discuss this further at the end of §6.

6. $3X+1$ Lower Bounds

We obtain lower bounds for the number $\pi_1(x)$ of integers below x that eventually iterate to 1 under the $3x+1$ function.

Theorem 6.1. *For each positive $a \not\equiv 0 (\mathrm{mod}\ 3)$ the function*

$$\pi_a(x) := |\{1 \leq n \leq x : \ \textit{Some} \ \ T^{(j}(n) = a.\}|$$

satisfies, for all sufficiently large $x \geq x_0(a)$,

$$\pi_a(x) \geq x^{0.84}.$$

Proof. This follows from Theorem 2.2, by finding a positive feasible solution by computer to the linear program family $L_k^{NT}(\lambda)$ for $k = 11$, for $\lambda = 1.7922310$, see Table 2 below. This yields the exponent $\gamma = \log_2 \lambda \approx 0.84175$. ∎.

Table 2 gives data on the bounds for the optimal λ for L_k^{NT} for $2 \leq k \leq 11$. For $1 \leq k \leq 9$ these are taken from [2]; the new values for $k = 10, 11$ were computed by D. Applegate. All values are rounded down in the last decimal place given, computations were to higher precision than shown.

k	γ_k	λ_k	C_k^{max}	$\bar{c}_{k,k}$	$\bar{c}_{k-1,k}$	$\bar{c}_{k,k} - \bar{c}_{k-1,k}$
2	0.4365	1.3534	1.8316	1.5237	1.0000	0.5237
3	0.6112	1.5275	3.4881	2.1014	1.6994	0.4020
4	0.6891	1.6122	5.4954	2.7869	2.4010	0.3858
5	0.7335	1.6627	9.0756	3.4648	3.0771	0.3876
6	0.7608	1.6944	12.8769	3.9667	3.5825	0.3841
7	0.7825	1.7201	20.1963	4.8122	4.4061	0.4061
8	0.8031	1.7449	29.1315	5.2028	4.8181	0.3846
9	0.8168	1.7615	43.3394	5.8102	5.4164	0.3937
10	0.8295	1.7771	64.9801	6.4567	6.0648	0.3919
11	0.8417	1.7922	98.4009	7.1552	6.7695	0.3856

Table 2: NLP Lower Bounds: No truncation of advanced terms

The last three columns in Table 2 present data on some average quantities formulated in [2]. Define

$$\bar{c}_{k,k} := \frac{1}{3^{k-1}} \sum_{m \in [3^k]} c_k^m$$

and

$$\bar{c}_{k-1,k} := \frac{1}{3^{k-2}} \sum_{m \in [3^{k-1}]} \bar{c}_{k-1}^m.$$

Adding up all the inequalities in $L_k^{NT}(\lambda)$ leads to

$$\bar{c}_{k,k} \leq \lambda^{-2}\bar{c}_{k,k} + \frac{1}{3}(\lambda^{\alpha-1} + \lambda^{\alpha-2})\bar{c}_{k-1,k}.$$

In [2] it was noted that a necessary and sufficient condition for a bound like $\pi_1(x) > x^{1-\epsilon}$ to hold for each $\epsilon > 0$ and all sufficiently large x is that $\lambda_k \to 2$ as $k \to \infty$, and this in turn would follow from the existence of feasible solutions with

$$\frac{c_{k-1,k}}{c_{k,k}} \to 1 \quad \text{as} \quad k \to \infty.$$

Table 2 gives more empirical data on these quantities.

The supremum of the exponential lower bounds that can be extracted from the linear program family $L_k^{NT}(\lambda)$ is given by λ_k, the supremum of values of λ for which $L_k^{NT}(\lambda)$ has a feasible solution. These values satisfy $\lambda_k \leq \lambda_{k+1}$, because given a feasible solution to $L_k(\lambda)$ with principal variables c_k^m one can define

$$c_{k+1}^{m+j\cdot 3^k} := c_k^m \quad \text{for} \quad 0 \leq j \leq 2,$$

and obtain a feasible solution to $L_{k+1}^{NT}(\lambda)$. It remains an open problem to show that the values λ_k are strictly increasing in k. As already noted in [2], showing that $\lambda_k \to 2$ as $k \to \infty$ would imply a lower bound $\pi_a(x) \geq x^{1-\epsilon}$ holds for each positive ϵ, for each $a \not\equiv 0 \pmod 3$ and all sufficiently large $x \geq x_0(a)$.

We now establish that the linear program system $L_k^{NT}(\lambda)$ used here is equivalent to the linear program system denoted L_λ^{NT} (for the same k) in [2] in the sense of Theorem 2.2; namely, the set of λ for which they have a strictly positive feasible solution coincide. To see this, observe first that if $L_k^{NT}(\lambda)$ has a feasible solution, then it has a strictly positive feasible solution. One may have to modify the auxiliary variables, which might be negative, while holding the principal variables fixed. However the auxiliary variables can be forced to their maximal values in terms of the principal variables without affecting feasiblity. Such a feasible solution has all values at least 1, so strict positivity is attained, and this solution also satisfies L_λ^{NT}. Conversely, given a positive feasible solution to L_λ^{NT}, it can be multiplicatively rescaled to have objective function value $c_1^2 = 1$, and this gives a feasible solution to $L_k^{NT}(\lambda)$, on taking $C_k^{max} := \max\{c_k^m\}$.

We conclude the paper by giving reasons to believe that the lower bound obtained in Theorem 2.2 is the largest one implied by the difference inequalities $\mathcal{I}_k$. This would follow if one could exhibit a positive monotone solution to $\mathcal{I}_k$ that has a growth rate matching the lower bound. Such a pure exponential lower bound could potentially be constructed from a solution to $L_k^{NT}(\lambda_k)$. Two conditions must hold:

(1) The supremum λ_k is attained. That is, $L_k^{NT}(\lambda_k)$ has a feasible solution.

(2) At the supremum value λ_k , there exists a feasible solution in which all of the principal inequalities (L1)-(L3) hold with equality.

If conditions (1), (2) hold, then the functions $\phi_k^m(y) = c_k^m \lambda_k^y$ would satisfy $\mathcal{I}_k$ with equality for all times $y \geq 2$, and would constitute a positive monotone solution to $\mathcal{I}_k$ attaining the best lower bound given by Theorem 2.2. Experimentally this is the case for $k \leq 11$.

Regarding condition (1), $L_k^{NT}(\lambda_k)$ could fail to have a feasible solution at the supremum value λ_k only if the objective function value as $\lambda \to \lambda_k$ from below diverges to ∞, so some variables c_k^m become unbounded. The numerical evidence up to $k = 11$ indicates that this does not happen. Regarding condition (2), the complementary slackness conditions for an optimal solution of a "generic" linear program of this type force all the principal inequalities (L1)-(L3) to hold with equality. In particular (2) must hold for any $L_k^{NT}(\lambda_k)$ having an optimal solution at which all variables c_k^m take distinct values. We think it likely that properties (1), (2) hold for all $k \geq 2$, but that this may be difficult to prove.

The linear program $L_k^{NT}(\lambda_k)$ at the supremum value λ_k has a finite optimal objective function value provided that condition (1) holds, as we now assume. This value $\tilde{C}_k^{max}$ has an interesting meaning: it measures the minimal spread attainable in the values of c_k^m, while normalizing these variables by $\min\{c_k^m\} = 1$. This quantity shows up in the constant Δ_1 in Theorem 2.2. One may view the value $\tilde{C}_k^{max}$ as a quantitative measure of a rate of "mixing" between congruence classes $(\bmod\ 3^k)$ that the $3x+1$ function produces. The fourth column of Table 2 indicates that the quantity $\tilde{C}_k^{max}$ exists for $k \leq 11$, and it appears to grow exponentially with k.

Acknowledgments. The authors thank David Applegate for computations reported in Tables 1 and 2, and the reviewer for helpful comments.

Appendix: Inequalities for $k = 2$

The case $k = 2$ is the only case where the derived inequalities $\mathcal{I}_k(EM)$ and the linear program family $L_k^{EM}(\lambda)$ have sufficiently few terms to be easily written down. There are three functions $\Phi_2 := \{\phi_2^2(y), \phi_2^5(y), \phi_2^8(y) : y \geq 0\}$. Recall that $\alpha = \log_2 3 \approx 1.585$. The inequalities $\mathcal{I}_2$ are

$$\begin{aligned}
\phi_2^2(y) &\geq \phi_2^8(y-2) + \min[\phi_2^2(y+\alpha-2), \phi_2^5(y+\alpha-2), \phi_2^8(y+\alpha-2)]\,, \\
\phi_2^5(y) &\geq \phi_2^2(y-2)\ , \\
\phi_2^8(y) &\geq \phi_2^5(y-2) + \min[\phi_2^2(y+\alpha-1), \phi_2^5(y+\alpha-1), \phi_2^8(y+\alpha-1)]\,.
\end{aligned}$$

Of these, only the inequality for $\phi_2^8(y)$ contains advanced terms on its right side.

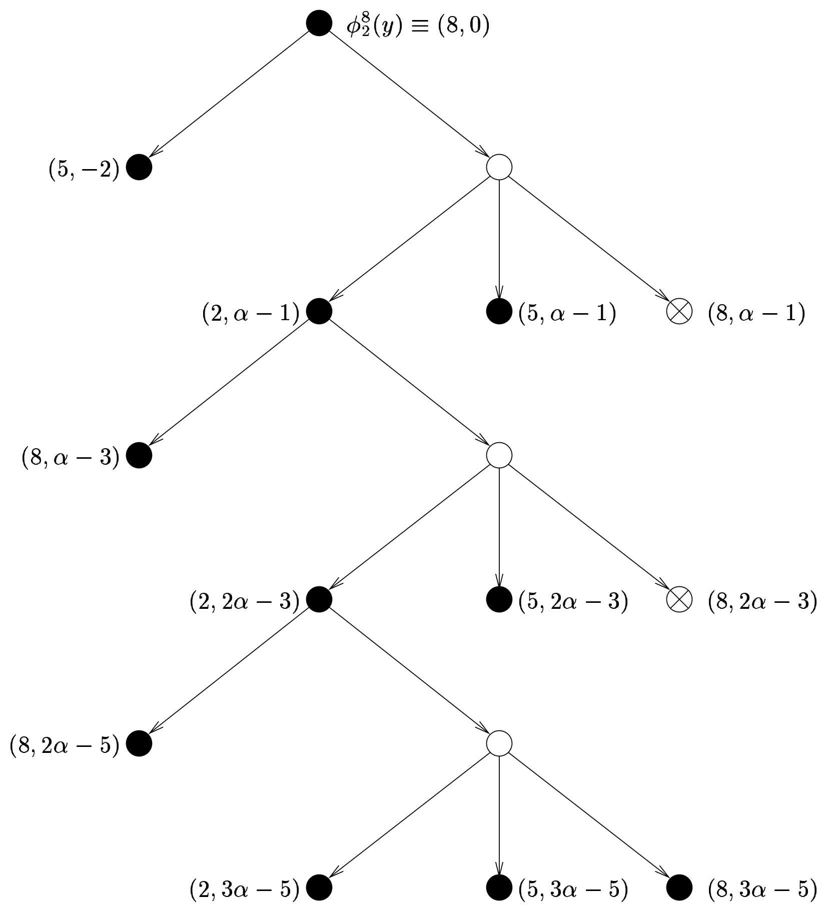

Figure 3: Tree $\mathcal{T}_2^8(EL)$ (Nodes marked $\otimes$ are deleted nodes)

The corresponding inequality $\mathcal{I}_2^8(EL)$ has three leaves of nested minimization. The corresponding tree $\mathcal{T}_2^8(EL)$ is pictured in Figure 3, with the deleted nodes indicated. The tree $\mathcal{T}_2^8(EL)$ has three m-nodes and eight leaf nodes. We let a_1, a_2, a_3 be the auxiliary variables for the leaf nodes, numbered as in Figure 3, and its associated inequalities are:

$$\phi_2^8(y) \geq \phi_2^5(y-2) + \min[\phi_2^8(y+\alpha-3) + M_1(y), \phi_2^2(y+\alpha-3)]\,,$$

in which

$$M_1(y) := \min[\phi_2^8(y+2\alpha-5) + M_2(y), \phi_2^5(y+2\alpha-5)]\,,$$

and

$$M_2(y) := \min[\phi_2^2(y+3\alpha-5), \phi_2^5(y+3\alpha-5), \phi_2^8(y+3\alpha-5)]\,.$$

The inequalities in the linear program $L_2^{EL}(\lambda)$ for the three trees $\mathcal{T}_2^m(EL)$ with $m = 2, 5$ and 8 are given in Table 3; they are associated to the leaves of these trees, identified by their labels in Table 3.

Tree	Leaf node label	Inequality
$\mathcal{T}_2^2(EL)$	$(8, -2)$	$c_2^2 \leq c_2^8\lambda^{-2} + a_1^{'}\lambda^{\alpha-2}$
	$(2,\ \alpha - 2)$	$a_1^{'}\lambda^{\alpha-2} \leq c_2^2\lambda^{\alpha-2}$
	$(5,\ \alpha - 2)$	$a_1^{'}\lambda^{\alpha-2} \leq c_2^5\lambda^{\alpha-2}$
	$(8,\ \alpha - 2)$	$a_1^{'}\lambda^{\alpha-2} \leq c_2^8\lambda^{\alpha-2}$
$\mathcal{T}_2^5(EL)$	$(2,\ -2)$	$c_2^5 \leq c_2^2\lambda^{-2}$
$\mathcal{T}_2^8(EL)$	$(5,\ -2)$	$c_2^8 \leq c_2^5\lambda^{-2} + a_1\lambda^{\alpha-1}$
	$(8,\ \alpha - 3)$	$a_1\lambda^{\alpha-1} \leq c_2^8\lambda^{\alpha-3} + a_2\lambda^{2\alpha-3}$
	$(2,\ \alpha - 3)$	$a_1\lambda^{\alpha-1} \leq c_2^2\lambda^{\alpha-3}$
	$(8,\ 2\alpha - 5)$	$a_2\lambda^{2\lambda-3} \leq c_2^8\lambda^{2\alpha-5} + a_3\lambda^{3\alpha-5}$
	$(2,\ 2\alpha - 5)$	$a_2\lambda^{2\alpha-3} \leq c_2^2\lambda^{2\alpha-5}$
	$(2,\ 3\alpha - 5)$	$a_3\lambda^{3\alpha-5} \leq c_2^2\lambda^{3\alpha-5}$
	$(5,\ 3\alpha - 5)$	$a_3\lambda^{3\alpha-5} \leq c_2^5\lambda^{3\alpha-5}$
	$(8,\ 3\alpha - 5)$	$a_3\lambda^{3\alpha-5} \leq c_2^8\lambda^{3\alpha-5}$

Table 3: $L_2^{EL}(\lambda)$ inequalities for trees $\mathcal{T}_2^m(EL)$.

References

[1] D. Applegate and J. C. Lagarias, *Density bounds for the* $3x+1$ *problem. I. Tree-search method*, Math. Comp. **64** (1995), 411-426.

[2] D. Applegate and J. C. Lagarias, *Density bounds for the* $3x+1$ *problem. II. Krasikov inequalities*, Math. Comp. **64** (1995), 427-438.

[3] R. E. Crandall, *On the* $3x+1$ *problem*, Math. Comp. **32** (1978), 1281-1292.

[4] I. Krasikov, *How many numbers satisfy* $3X+1$ *conjecture?*, Internat. J. Math & Math.Sci. **12** (1989), 791-796

[5] J. C. Lagarias, *The* $3x+1$ *problem and its generalizations*, Amer. Moth. Monthly **92** (1985), 3-21.

[6] J. W. Sander, *On the* $(3X+1)$-*Conjecture*, Acta Arithmetica **55** (1990), 241–248.

[7] Ya. G. Sinai, *Statistical* $(3X+1)$- *Problem*, preprint.

[8] G. J. Wirsching, *An improved estimate concerning* $3N+1$ *predecessor sets*, Acta Arith. **63** (1993), 205-210.

[9] G. J. Wirsching, *The Dynamical System Generated by the* $3n+1$ *Function,* Lecture Notes in Math. No. 1681, Springer-Verlag: Berlin 1998.

`email:  Ilia.Krasikov@brunel.ac.uk`
`        jcl@research.att.com`