Parikh's Theorem in Commutative Kleene Algebra

Mark W. Hopkins
Adaptive Micro Systems, Inc.
7840 North 86th St.
Milwaukee, Wisconsin 53224, USA
mwh@ams-i.com

Dexter C. Kozen
Department of Computer Science
Cornell University
Ithaca, New York 14853-7501, USA
kozen@cs.cornell.edu

Abstract

Parikh's Theorem says that the commutative image of every context free language is the commutative image of some regular set. Pilling has shown that this theorem is essentially a statement about least solutions of polynomial inequalities. We prove the following general theorem of commutative Kleene algebra, of which Parikh's and Pilling's theorems are special cases: Every finite system of polynomial inequalities $f_i(x_1, \ldots, x_n) \leq x_i$, $1 \leq i \leq n$, over a commutative Kleene algebra K has a unique least solution in K^n; moreover, the components of the solution are given by polynomials in the coefficients of the f_i. We also give a closed-form solution in terms of the Jacobian matrix of the system.

1 Introduction

Parikh's theorem [9] says that every context-free language is "letter-equivalent" to a regular set; formally, the commutative image of any context-free language is also the commutative image of some regular set, where the *commutative image* of a set A of strings over the finite alphabet $\{a_1, \ldots, a_k\}$ is the set of k-tuples

$$\{(\#a_1(x), \ldots, \#a_k(x)) \in \mathbb{N}^k \mid x \in A\} \quad \subseteq \quad \mathbb{N}^k,$$

where $\#a_i(x)$ is the number of occurrences of a_i in x. The k-tuples in $\mathbb{N}^k$ are often called *Parikh vectors*. For example, the context-free language $\{a^n b^n \mid n \geq 0\}$ is letter-equivalent to the regular set $(ab)^*$; these two sets have a common commutative image $\{(n, n) \mid n \geq 0\}$.

The usual combinatorial proofs of Parikh's theorem involve an induction on parse trees of context-free grammars. In this paper we prove the following general theorem of commutative Kleene algebra, of which Parikh's theorem is a special case:

Theorem 1.1 *Every system of inequalities*

$$f_i(x_1, \ldots, x_n) \quad \leq \quad x_i, \quad 1 \leq i \leq n, \tag{1}$$

where the f_i are polynomials in $K[x_1, \ldots, x_n]$ over a commutative Kleene algebra K, has a unique least solution in K^n; moreover, the components of the solution are given by polynomials in the coefficients of the f_i.

We might take the statement of Theorem 1.1 as a definition of *algebraic closure* in Kleene algebra, in which case the theorem says that any commutative Kleene algebra is algebraically closed.

Pilling [10] proves Theorem 1.1 in the special case of the commutative Kleene algebra $\mathsf{Reg}(\mathbb{N}^k)$, the algebra of regular sets of Parikh vectors, and argues that this is the essential content of Parikh's theorem. Indeed, context-free grammars are just systems of set inequalities, and the context-free languages they generate are the minimal solutions. For example, the context-free grammar $S \to aSb \mid \epsilon$ is essentially the system consisting of the single inequality $axb + 1 \leq x$ whose least solution in $2^{\{a,b\}^*}$ is the context-free language $\{a^n b^n \mid n \geq 0\}$. Under the assumption of commutativity, the inequality can be rewritten $abx + 1 \leq x$, whose least solution is the regular set $(ab)^*$. For a somewhat more difficult example, the context-free grammar $S \to \texttt{[}\, S \,\texttt{]} \mid SS \mid \epsilon$ generating the set of balanced strings of parentheses is essentially the system consisting of the single inequality $\texttt{[}\, x \,\texttt{]} + xx + 1 \leq x$. Under the assumption of commutativity, the inequality can be rewritten $\texttt{[ ]}x + x^2 + 1 \leq x$, whose least solution is the regular set $(\texttt{[ ]})^*$.

Using results of [4, Lemma 7.1, p. 35] and [7, Section 2.3, p. 198] one can generalize Pilling's proof to any *-continuous Kleene algebra. However, the proof makes essential use of various infinitary properties such as the continuity of regular operators and the fact that a^* is the supremum of the a^n, $n \geq 0$.

Kuich [8] also gives a generalization of Parikh's theorem that holds for any commutative idempotent ω-continuous semiring. Kuich's result implies Pilling's, since $\mathsf{Reg}(\mathbb{N}^k)$

is embedded in the commutative idempotent ω-continuous semiring $2^{\mathbb{N}^k}$. Conversely, since every commutative idempotent ω-continuous semiring is a commutative Kleene algebra under the usual definition of the * operator

$$a^* \quad = \quad \sum_{n \geq 0} a^n,$$

Pilling's result, suitably generalized to *-continuous Kleene algebras, would imply Kuich's. But again, these proofs depend on the strong infinitary properties of *-continuous algebras.

Our result is a generalization of these results in that it holds in all commutative Kleene algebras. The main difference here is that Kleene algebra as defined in [5] has a *finitary* algebraic axiomatization consisting of finitely many equations and equational implications. Thus one might say that we are replacing the *analytic* arguments of Pilling and Kuich with *algebraic* arguments. The fact that we cannot argue combinatorially in the model $\mathsf{Reg}(\mathbb{N}^k)$ or use the infinitary properties of *-continuous algebras makes the proof more difficult, but also makes the result considerably stronger.

The situation is analogous to the fundamental theorem of algebra, which states that the complex numbers $\mathbb{C}$ are algebraically closed. The most common proof of this theorem, originally due to Gauss, depends on the analytic structure of $\mathbb{C}$ and uses second-order arguments (see e.g. [12]). However, one can give a first-order, purely algebraic proof of the more general result that if R is any real closed field (such as $\mathbb{R}$ or $\mathbb{A}$, the real algebraic numbers), then $R[i]$ is algebraically closed (see e.g. [11]). Like the fundamental theorem of algebra, our result also deals with solutions of polynomial systems, and our proof replaces arguments referring to the analytic or second-order structure of $\mathsf{Reg}(\mathbb{N}^k)$, embodied in the *-continuity axiom, with first-order equational arguments referring only to the finitary algebraic structure of commutative Kleene algebras.

Our development involves the definition of differential operators $\frac{\partial}{\partial x}$ on commutative Kleene algebras of polynomials and a version of Taylor's theorem:

$$f(x+d) \quad = \quad f(x) + f'(x+d) \cdot d.$$

Differential operators allow us to define the *Jacobian matrix* of a system of inequalities, which we use to give a closed form solution.

In Section 2 we review the definitions of Kleene algebra and commutative Kleene algebra. In Section 3 we discuss polynomials over a commutative Kleene algebra, define differential operators on a commutative Kleene algebra of polynomials, and develop some basic properties, culminating in a version of Taylor's theorem. In Section 4 we prove Theorem 1.1. In Section 5 we give a closed form solution in terms of the Jacobian matrix of a system of inequalities.

2 Commutative Kleene Algebra

Kleene algebra is the algebra of regular expressions [3, 1]. The axiomatization we adopt here is from [5]. A *Kleene algebra* is an algebraic structure $(K,\ +,\ \cdot,\ ^*,\ 0,\ 1)$ that is an idempotent semiring under $+,\ \cdot,\ 0,\ 1$ satisfying

$$1 + pp^* \quad = \quad p^* \tag{2}$$
$$1 + p^*p \quad = \quad p^* \tag{3}$$
$$q + pr \leq r \quad \rightarrow \quad p^*q \leq r \tag{4}$$
$$q + rp \leq r \quad \rightarrow \quad qp^* \leq r \tag{5}$$

where $\leq$ refers to the natural partial order on K:

$$p \leq q \quad \stackrel{\text{def}}{\iff} \quad p + q = q\ .$$

The operation $+$ gives the supremum with respect to the natural order $\leq$. Instead of (4) and (5), we might take the equivalent axioms

$$pr \leq r \quad \rightarrow \quad p^*r \leq r \tag{6}$$
$$rp \leq r \quad \rightarrow \quad rp^* \leq r\ . \tag{7}$$

These axioms say essentially that * behaves like the Kleene asterate operator of formal language theory or the reflexive transitive closure operator of relational algebra.

A Kleene algebra is **-continuous* if it satisfies the additional infinitary axiom

$$pq^*r \quad = \quad \sup_{n \geq 0} pq^n r,$$

where the supremum on the right-hand side is with respect to the natural order $\leq$. A Kleene algebra or *-continuous Kleene algebra is *commutative* if it satisfies the additional axiom $pq = qp$.

Kleene algebras play a prominent role in dynamic logic and other program logics. Standard models include the family of regular sets over a finite alphabet; the family of binary relations on a set; and the family of $n \times n$ matrices over another Kleene algebra. Other more unusual interpretations include the min,+ algebra used in shortest path algorithms and models consisting of convex polyhedra used in computational geometry [2]. All naturally occurring models are *-continuous. Commutativity assumptions also arise in practice [6].

The following are some typical identities of Kleene algebra:

$$(p^*q)^*p^* \quad = \quad (p+q)^* \tag{8}$$
$$p(qp)^* \quad = \quad (pq)^*p \tag{9}$$
$$p^* \quad = \quad (pp)^*(1+p)\ . \tag{10}$$

All the operators are monotone with respect to $\leq$. In other words, if $p \leq q$, then $pr \leq qr, rp \leq rq, p+r \leq q+r$, and $p^* \leq q^*$ for any r.

The following is a theorem of commutative Kleene algebra that does not hold in Kleene algebra in general:

$$(p+q)^* \quad = \quad p^*q^*. \tag{11}$$

Using this, one can prove a normal form theorem that says that every expression is equivalent to a sum $y_1 + \cdots + y_n$, where each y_i is a product of atomic symbols and expressions of the form $(a_1 \cdots a_k)^*$, where the a_i are atomic symbols. For example,

$$(((ab)^*c)^* + d)^* \quad = \quad d^* + (ab)^*c^*cd^*.$$

This normal form was observed by Pilling [10] in the context of $\mathsf{Reg}(\mathbb{N}^k)$, but using (11) it is easily shown to hold in all commutative Kleene algebras.

The equational theory of Kleene algebras and *-continuous Kleene algebras coincide [5], but their Horn theories do not; indeed, the Horn theory of *-continuous Kleene algebras is Π^1_1-complete [7].

See [5] for a more thorough introduction to Kleene algebra.

3 Polynomials and Differential Operators

3.1 Polynomials over a commutative Kleene algebra

If K is a commutative Kleene algebra, we denote by $K[\mathbf{x}]$ the commutative Kleene algebra of polynomials in indeterminates $\mathbf{x}$ over K. These are very much like polynomials over a ring or field. We can think of a polynomial as a regular expression over K and $\mathbf{x}$ reduced modulo the axioms of commutative Kleene algebra and the diagram of K (the set of ground identities that hold in K). Typical examples of polynomials are

$$(ax+by)^*$$
$$1 + (ax^*b^*)^* + bx + cy$$
$$a + xy(bxy)^*,$$

where x, y are indeterminates and $a, b, c \in K$.

Formally, $K[\mathbf{x}]$ is defined to be the direct sum (coproduct) of K with the free commutative Kleene algebra on generators $\mathbf{x}$ in the category of commutative Kleene algebras. The most significant property of polynomials is that any pair of maps h, h', where $h : K \to L$ is a Kleene algebra homomorphism and $h' : \mathbf{x} \to L$ is a set function, extend simultaneously and uniquely to a Kleene algebra homomorphism $\widehat{h} : K[\mathbf{x}] \to L$. When h is the identity on K, the map $\widehat{h}$ is just polynomial evaluation; intuitively, applying $\widehat{h}$ can be regarded as substituting the values $h'(x)$ for the indeterminates $x \in \mathbf{x}$ and then evaluating the resulting expression.

If $\mathbf{x} = x_1, \ldots, x_n$ and $\mathbf{a} = a_1, \ldots, a_n$, we write $f(\mathbf{a})$ or $f(\mathbf{x})\mid_{\mathbf{x}=\mathbf{a}}$ for the value of f evaluated at $x_i \mapsto a_i, 1 \leq i \leq n$.

3.2 Differential Operators

A map $D : K \to K$ on a commutative Kleene algebra K is called a *differential operator* if for all $x, y \in K$,

$$\begin{array}{rcl} D(x+y) & = & Dx + Dy \\ D(xy) & = & xDy + yDx \\ D(x^*) & = & x^*Dx \\ D0 & = & D1 \;=\; 0. \end{array} \tag{12}$$

For example, in $\mathsf{Reg}(\mathbb{N}^k)$, for every $1 \leq i \leq k$, the map

$$A \quad \mapsto \quad \{(a_1, \ldots, a_{i-1}, a_i - 1, a_{i+1}, \ldots, a_k) \mid (a_1, \ldots, a_k) \in A,\ a_i > 0\}$$

is a differential operator.

Theorem 3.1 *Any differential operator $D : K \to K$ and set function $D : \mathbf{x} \to K$ have a unique joint extension to a differential operator $D : K[\mathbf{x}] \to K[\mathbf{x}]$.*

Proof. The given maps D can be extended by induction to $D : K[\mathbf{x}] \to K[\mathbf{x}]$ using (12); but we must take care that the extended D is well-defined on equivalence classes modulo the axioms of commutative Kleene algebra and the diagram of K. That the extended D respects the diagram of K follows from the fact that the given $D : K \to K$ is a differential operator on K. To prove that D respects the commutative Kleene algebra axioms requires a case for each axiom. We argue the cases $a^* = 1 + aa^*$ and $ab \leq b \to a^*b \leq b$ explicitly.

For the case $a^* = 1 + aa^*$,

$$\begin{array}{rcl} D(1 + aa^*) & = & D1 + D(aa^*) \\ & = & 0 + aD(a^*) + a^*Da \\ & = & aa^*Da + a^*Da \\ & = & a^*Da \\ & = & D(a^*). \end{array}$$

For the case $ab \leq b \to a^*b \leq b$, suppose $ab \leq b$. By the induction hypothesis, $D(ab) \leq Db$, and we wish to show that $D(a^*b) \leq Db$. From $D(ab) \leq Db$ we have that $aDb + bDa \leq Db$, thus by (6) we have that $a^*Db \leq Db$, and by (4) we have that $a^*bDa \leq Db$. Therefore

$$\begin{array}{rcl} D(a^*b) & = & a^*Db + bD(a^*) \\ & = & a^*Db + ba^*Da \\ & \leq & Db. \end{array}$$

□

In particular, for $x \in \mathbf{x}$, we define a certain differential operator $\frac{\partial}{\partial x} : K[\mathbf{x}] \to K[\mathbf{x}]$ as follows. The value of $\frac{\partial}{\partial x}$ applied to $f \in K[\mathbf{x}]$ is denoted $\frac{\partial f}{\partial x}$ or $\frac{\partial f}{\partial x}(\mathbf{x})$. We define $\frac{\partial}{\partial x}$ to be the unique differential operator such that $\frac{\partial x}{\partial x} = 1$, $\frac{\partial y}{\partial x} = 0$ for $y \in \mathbf{x} - \{x\}$, and $\frac{\partial a}{\partial x} = 0$ for $a \in K$.

For univariate polynomials $f, e \in K[x]$, we sometimes write f' for $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial x}(e)$ or $f'(e)$ for the result of evaluating the polynomial $\frac{\partial f}{\partial x}$ at $x \mapsto e$.

For example, if $f(x) = axb + x^2 + 1$, then $f'(x) = ab + x$.

Note that $\frac{\partial}{\partial x}(f(e))$ and $\frac{\partial f}{\partial x}(e)$ are different in general. The former refers to the result of evaluating $f(x)$ at $x \mapsto e$ first, then applying the differential operator $\frac{\partial}{\partial x}$ to $f(e)$; whereas the latter refers to the result of applying the differential operator $\frac{\partial}{\partial x}$ to $f(x)$ first, then evaluating the resulting polynomial $\frac{\partial f}{\partial x}(x)$ at $x \mapsto e$. These two expressions are related by the *chain rule*:

Theorem 3.2 (chain rule) *For $f, e \in K[x]$,*

$$\frac{\partial}{\partial x}(f(e)) \quad = \quad \frac{\partial f}{\partial x}(e) \cdot \frac{\partial e}{\partial x},$$

or in more conventional notation,

$$f(e(x))' \quad = \quad f'(e(x)) \cdot e'(x).$$

Proof. This is a straightforward induction on the structure of f. We argue the cases $f = gh$ and $f = g^*$ explicitly.

$$\begin{aligned}
(g(e)h(e))' &= g(e)h(e)' + h(e)g(e)' \\
&= g(e)h'(e)e' + h(e)g'(e)e' \\
&= (g(e)h'(e) + h(e)g'(e))e' \\
&= (gh' + hg')(e)e' \\
&= (gh)'(e)e'.
\end{aligned}$$

$$\begin{aligned}
(g(e)^*)' &= g(e)^* g(e)' \\
&= g(e)^* g'(e)e' \\
&= (g^* g')(e)e' \\
&= (g^*)'(e)e'.
\end{aligned}$$

□

For example, if $f(x) = ax + x^* + 1$, then

$$\begin{aligned}
f'(x) &= a + x^* \\
f(f(x)) &= a(ax + x^* + 1) + (ax + x^* + 1)^* + 1 \\
&= a^2x + ax^* + (ax)^* x^* + 1 \\
f(f(x))' &= (a^2x + ax^* + (ax)^* x^* + 1)' \\
&= a^2 + (a+1)x^*(ax)^* \\
f'(f(x))f'(x) &= (a + (ax + x^* + 1)^*)(a + x^*) \\
&= a^2 + (a+1)x^*(ax)^*.
\end{aligned}$$

We also have the following version of Taylor's theorem in commutative Kleene algebra.

Theorem 3.3 (Taylor's theorem) *For $f, d \in K[x]$,*

$$f(x+d) \quad = \quad f(x) + f'(x+d) \cdot d.$$

In particular, evaluating at $x \mapsto 0$,

$$f(d) \quad = \quad f(0) + f'(d) \cdot d.$$

Proof. This is again a straightforward induction on the structure of f. As before, we argue the cases $f = gh$ and $f = g^*$ explicitly. For the case $f = gh$,

$$\begin{aligned}
&gh(x+d) \\
&\quad = \quad g(x+d)h(x+d) \\
&\quad = \quad g(x+d)h(x) + g(x+d)h'(x+d)d \\
&\quad = \quad g(x)h(x) + g'(x+d)h(x)d \\
&\qquad\quad + g(x+d)h'(x+d)d,
\end{aligned}$$

and by symmetry,

$$\begin{aligned}
&gh(x+d) \\
&\quad = \quad g(x)h(x) + g'(x+d)h(x+d)d \\
&\qquad\quad + g(x)h'(x+d)d,
\end{aligned}$$

therefore by monotonicity,

$$\begin{aligned}
&gh(x+d) \\
&\quad = \quad g(x)h(x) + g'(x+d)h(x)d + g(x+d)h'(x+d)d \\
&\qquad\quad + g'(x+d)h(x+d)d + g(x)h'(x+d)d \\
&\quad = \quad g(x)h(x) + g(x+d)h'(x+d)d \\
&\qquad\quad + g'(x+d)h(x+d)d \\
&\quad = \quad gh(x) + (gh)'(x+d)d.
\end{aligned}$$

For the case $f = g^*$,

$$\begin{aligned}
&g(x+d)^* \\
&\quad = \quad (g(x) + g'(x+d)d)^* \\
&\quad = \quad g(x)^*(g'(x+d)d)^* \qquad \text{by (11)} \\
&\quad = \quad g(x)^* + g(x)^* g'(x+d)d(g'(x+d)d)^* \\
&\quad = \quad g(x)^* + g'(x+d)dg(x)^*(g'(x+d)d)^* \\
&\quad = \quad g(x)^* + g'(x+d)dg(x+d)^* \\
&\quad = \quad g(x)^* + g(x+d)^* g'(x+d)d \\
&\quad = \quad g(x)^* + (g^*)'(x+d)d.
\end{aligned}$$

□

Continuing with the example $f(x) = ax + x^* + 1$ above,

$$\begin{aligned}
f(x+d) &= a(x+d) + (x+d)^* + 1 \\
&= ax + ad + x^* d^* + 1 \\
f'(x+d) &= a + (x+d)^* \\
&= a + x^* d^* \\
f(x) + f'(x+d)d &= ax + x^* + 1 + ad + x^* d^* d \\
&= ax + ad + x^* d^* + 1.
\end{aligned}$$

We often wish to differentiate simultaneously with respect to a sequence of indeterminates $\mathbf{y} = y_1, \ldots, y_k$. We define an operator $\frac{\partial}{\partial \mathbf{y}}$ that when applied to an element $f \in K[\mathbf{x}]$ produces a row vector of length k whose i^{th} component is $\frac{\partial f}{\partial y_i}$. More generally, $\frac{\partial}{\partial \mathbf{y}}$ applied to a column vector consisting of m elements $f_1, \ldots, f_m \in K[\mathbf{x}]$ produces an $m \times k$ matrix whose i, j^{th} element is $\frac{\partial f_i}{\partial y_j}$.

By iterating Theorem 3.3, one can show that for $f \in K[\mathbf{x}]$ and $\mathbf{e} = e_1, \ldots, e_n$,

$$\begin{aligned} f(\mathbf{e}) &= f(0, \ldots, 0) + \tfrac{\partial f}{\partial x_1}(\mathbf{e})e_1 + \cdots + \tfrac{\partial f}{\partial x_n}(\mathbf{e})e_n \\ &= f(\mathbf{0}) + \tfrac{\partial f}{\partial \mathbf{x}}(\mathbf{e}) \cdot \mathbf{e}, \end{aligned}$$

where $\cdot$ denotes dot product of vectors. The same holds for a column vector $\mathbf{f} = f_1, \ldots, f_m$ of elements of $K[\mathbf{x}]$; here $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x})$ is an $m \times n$ matrix whose i, j^{th} element is $\frac{\partial f_i}{\partial x_j}$, and

$$\mathbf{f}(\mathbf{e}) = \mathbf{f}(\mathbf{0}) + \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{e}) \cdot \mathbf{e}, \tag{13}$$

where in this case $\cdot$ denotes matrix-vector multiplication. The $m \times n$ matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x})$ is called the *Jacobian* of $\mathbf{f}$.

We also have the following vector-vector and matrix-vector versions of the chain rule, Theorem 3.2:

$$\begin{aligned} \tfrac{\partial}{\partial z}(f(\mathbf{e})) &= \tfrac{\partial f}{\partial x_1}(\mathbf{e})\tfrac{\partial e_1}{\partial z} + \cdots + \tfrac{\partial f}{\partial x_n}(\mathbf{e})\tfrac{\partial e_n}{\partial z} \\ &= \tfrac{\partial f}{\partial \mathbf{x}}(\mathbf{e}) \cdot \tfrac{\partial \mathbf{e}}{\partial z} \\ \tfrac{\partial}{\partial z}(\mathbf{f}(\mathbf{e})) &= \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{e}) \cdot \tfrac{\partial \mathbf{e}}{\partial z}. \end{aligned}$$

The proof of these propositions is a straightforward generalization of the proof of Theorem 3.2.

4 A Generalization of Parikh's Theorem

In this section we prove Theorem 1.1. We first prove the result for $n = 1$, then extend it to arbitrary n. Many arguments in this section are inspired by those of Pilling [10] (see also [1]) but generalized to apply to arbitrary commutative Kleene algebras.

Theorem 4.1 *Let K be a commutative Kleene algebra and let $f(x) \in K[x]$. The unique least solution of the inequality $f(x) \leq x$ is*

$$f'(f(0))^* \cdot f(0). \tag{14}$$

Moreover, this holds uniformly over all homomorphic images of K.

For example, the context-free language $A = \{a^n b^n \mid n \geq 0\}$ is generated by the grammar $S \to aSb \mid \epsilon$, which translates to the one-dimensional system $axb + 1 \leq x$. Letting $f(x) = axb + 1$, we get $f'(x) = ab$ and $f(0) = 1$, thus (14) gives $(ab)^*$. This is a regular expression describing a regular set letter-equivalent to A.

Proof. First we argue that (14) is a solution to $f(x) \leq x$. It follows by a straightforward inductive argument that for any polynomial $h(x)$,

$$ac \leq bc \quad \to \quad h(a)c \leq h(b)c. \tag{15}$$

Applying this with $b = f(0)$, $c = f'(b)^*$, and $a = bc$,

$$\begin{aligned} f(a) &= f(bc) \\ &= f(0) + f'(bc)bc && \text{by Theorem 3.3} \\ &= b + f'(bc)bc \\ &\leq b + f'(b)bc && \text{by (15)} \\ &= b + f'(b)f'(b)^*b \\ &= f'(b)^*b && \text{by Kleene algebra} \\ &= a. \end{aligned}$$

Now we show that (14) is the least solution. Suppose y is any solution; thus $f(y) \leq y$. We wish to show that

$$f'(f(0))^* f(0) \leq y.$$

By (4), it suffices to show

$$f(0) + f'(f(0)) \cdot y \leq y.$$

But by monotonicity, $f(0) \leq f(y) \leq y$, and

$$\begin{aligned} & f(0) + f'(f(0)) \cdot y \\ &\leq f(0) + f'(y) \cdot y && \text{by monotonicity} \\ &= f(y) && \text{by Theorem 3.3} \\ &\leq y. \end{aligned}$$

The expression (14) gives the least solution of $f(x) \leq x$ uniformly over all homomorphic images of K because the axioms of Kleene algebra used in the proof hold universally under any interpretation. □

The uniformity condition of Theorem 4.1 may seem obvious, but it is actually a rather subtle point. The issue is that equations are preserved under homomorphisms, but in general Horn formulas (equational implications) are not. The homomorphic image $h(e)$ of a solution e of an inequality $f(x) \leq x$ is a solution of the homomorphic image of the inequality, because the inequality is equivalent to an equation $f(x) + x = x$; but that $h(e)$ is the *least* solution does not follow from the fact that e is least, since this property requires a Horn formula.

Proof of Theorem 1.1. We iterate the one-dimensional solution as follows. Consider the two-dimensional system

$$\begin{aligned} f(x, y) &\leq x \\ g(x, y) &\leq y. \end{aligned} \tag{16}$$

Viewing $K[x,y]$ as $K[x][y]$, first compute the least solution to the one-dimensional system $g(x,y) \leq y$ in $K[x]$; call it $h(x)$. Then compute the least solution a of $f(x,h(x)) \leq x$ in K.

We claim that $(a,h(a))$ is the desired least solution to (16) in K^2. Surely $f(a,h(a)) \leq a$ by the one-dimensional argument. Moreover, by the uniformity observation, we also have $g(a,h(a)) \leq h(a)$, since it is the image of $g(x,h(x)) \leq h(x)$ under the evaluation homomorphism $x \mapsto a$.

To show $(a,h(a))$ is the least solution, suppose (b,c) is any other solution. Then $f(b,c) \leq b$ and $g(b,c) \leq c$. Using the uniformity observation with the evaluation morphism $x \mapsto b$, we have that $h(b)$ is the least solution of $g(b,y) \leq y$. Then $h(b) \leq c$. But by monotonicity, $f(b,h(b)) \leq f(b,c) \leq b$. Since a is the least solution to $f(x,h(x)) \leq x$, we have that $a \leq b$. Again by monotonicity, $h(a) \leq h(b) \leq c$. Thus $(a,h(a)) \leq (b,c)$.

By iterating this process inductively, we can obtain the existence of a solution to any $n \times n$ system. □

5 A Closed Form Solution

The iterated construction of the previous section does not give a symmetric closed-form expression for any dimension greater than one. In this section we provide a symmetric closed-form solution.

Let K be a commutative Kleene algebra and consider an $n \times n$ system

$$\mathbf{f}(\mathbf{x}) \quad \leq \quad \mathbf{x} \tag{17}$$

where $\mathbf{x} = x_1, \ldots, x_n$ and $\mathbf{f} = \mathbf{f}(\mathbf{x}) = f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}) \in K[\mathbf{x}]$. Let $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ be the Jacobian of the system (17) as defined in Section 3.2. Define

$$\begin{array}{rcl} \mathbf{a}_0 & \stackrel{\text{def}}{=} & \mathbf{f}(\mathbf{0}) \\ \mathbf{a}_{k+1} & \stackrel{\text{def}}{=} & \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_k)^* \mathbf{a}_k. \end{array}$$

Theorem 5.1 *For sufficiently large finite N, the n-vector $\mathbf{a}_N$ is the least solution to* (17). *Moreover, this solution is uniform over all homomorphic images of K.*

Below we will derive an explicit single-exponential bound on N as a function of n.

Proof. We will prove the first statement of the theorem; the uniformity property will follow by the same considerations as in the proof of Theorem 4.1. The proof proceeds by induction on n. The basis $n = 1$ was given in Theorem 4.1.

Now suppose $n \geq 2$. Partition n as $m + (n-m)$ where $1 \leq m < n$. For an n-vector $\mathbf{b} = b_1, \ldots, b_n$, write

$$\begin{array}{rcl} \mathbf{b}^{\blacksquare\square} & \stackrel{\text{def}}{=} & b_1, \ldots, b_m \\ \mathbf{b}^{\square\blacksquare} & \stackrel{\text{def}}{=} & b_{m+1}, \ldots, b_n, \end{array}$$

and for an $n \times n$ matrix M, write

$$M^{\blacksquare\blacksquare}, \quad M^{\blacksquare\square}, \quad M^{\square\blacksquare}, \quad M^{\square\square}$$

for the upper left $m \times m$, upper right $m \times (n-m)$, lower left $(n-m) \times m$, and lower right $(n-m) \times (n-m)$ submatrices of M, respectively. To simplify notation, define

$$\begin{array}{rcl} \mathbf{y} & \stackrel{\text{def}}{=} & \mathbf{x}^{\blacksquare\square}, \\ \mathbf{z} & \stackrel{\text{def}}{=} & \mathbf{x}^{\square\blacksquare}, \end{array} \qquad \begin{array}{rcl} \mathbf{g} & \stackrel{\text{def}}{=} & \mathbf{f}^{\blacksquare\square}, \\ \mathbf{h} & \stackrel{\text{def}}{=} & \mathbf{f}^{\square\blacksquare}. \end{array}$$

In this notation, we can rewrite (17) as

$$\begin{array}{rcl} \mathbf{g}(\mathbf{y},\mathbf{z}) & \leq & \mathbf{y} \\ \mathbf{h}(\mathbf{y},\mathbf{z}) & \leq & \mathbf{z}. \end{array} \tag{18}$$

Also,

$$\begin{array}{rcl} \frac{\partial \mathbf{g}}{\partial \mathbf{y}} & = & \frac{\partial \mathbf{f}}{\partial \mathbf{x}}^{\blacksquare\blacksquare}, \\ \frac{\partial \mathbf{h}}{\partial \mathbf{y}} & = & \frac{\partial \mathbf{f}}{\partial \mathbf{x}}^{\square\blacksquare}, \end{array} \qquad \begin{array}{rcl} \frac{\partial \mathbf{g}}{\partial \mathbf{z}} & = & \frac{\partial \mathbf{f}}{\partial \mathbf{x}}^{\blacksquare\square}, \\ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} & = & \frac{\partial \mathbf{f}}{\partial \mathbf{x}}^{\square\square}. \end{array}$$

Now define

$$\begin{array}{rcl} \mathbf{c}_0(\mathbf{y}) & \stackrel{\text{def}}{=} & \mathbf{h}(\mathbf{y},\mathbf{0}) \\ \mathbf{c}_{k+1}(\mathbf{y}) & \stackrel{\text{def}}{=} & \frac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^* \mathbf{c}_k(\mathbf{y}) \end{array}$$

By the induction hypthesis, there exists a P such that $\mathbf{c}_P(\mathbf{y})$ is the least solution to the system

$$\mathbf{h}(\mathbf{y},\mathbf{z}) \quad \leq \quad \mathbf{z}$$

uniformly in $\mathbf{y}$. Define

$$\begin{array}{rcl} \widehat{\mathbf{g}}(\mathbf{y}) & \stackrel{\text{def}}{=} & \mathbf{g}(\mathbf{y}, \mathbf{c}_P(\mathbf{y})) \\ \mathbf{b}_0 & \stackrel{\text{def}}{=} & \widehat{\mathbf{g}}(\mathbf{0}) \\ \mathbf{b}_{k+1} & \stackrel{\text{def}}{=} & \frac{\partial \widehat{\mathbf{g}}}{\partial \mathbf{y}}(\mathbf{b}_k)^* \mathbf{b}_k. \end{array}$$

Again by the induction hypthesis, there exists an M such that $\mathbf{b}_M$ is the least solution to the system

$$\widehat{\mathbf{g}}(\mathbf{y}) \quad \leq \quad \mathbf{y}.$$

By the uniformity of the solutions, we have that

$$\begin{array}{rcl} \mathbf{g}(\mathbf{b}_M, \mathbf{c}_P(\mathbf{b}_M)) & \leq & \mathbf{b}_M \\ \mathbf{h}(\mathbf{b}_M, \mathbf{c}_P(\mathbf{b}_M)) & \leq & \mathbf{c}_P(\mathbf{b}_M), \end{array}$$

and $\mathbf{b}_M, \mathbf{c}_P(\mathbf{b}_M)$ is the least solution to (17). Moreover, this is the least solution uniformly over all homomorphic images.

Our task now is to show that for sufficiently large N,

$$\mathbf{b}_M = \mathbf{a}_N^{\blacksquare\square}, \qquad \mathbf{c}_P(\mathbf{b}_M) = \mathbf{a}_N^{\square\blacksquare}. \tag{19}$$

The inequalities $\geq$ follow from the fact that if $\mathbf{u}$ is any solution to (17), then $\mathbf{a}_k \leq \mathbf{u}$ for all k. This can be shown by induction on k. Certainly

$$\mathbf{a}_0 \;=\; \mathbf{f}(\mathbf{0}) \;\leq\; \mathbf{f}(\mathbf{u}) \;\leq\; \mathbf{u},$$

and by (13),

$$\tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{u})\mathbf{u} \;\leq\; \mathbf{u},$$

from which it follows that

$$\begin{aligned}\mathbf{a}_{k+1} &= \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_k)^*\mathbf{a}_k \\ &\leq \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{u})^*\mathbf{u} \\ &\leq \mathbf{u}.\end{aligned}$$

Now we establish a series of inequalities from which the forward inequalities $\leq$ of (19) will follow. First,

$$\mathbf{a}_k \;\leq\; \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_k)^*\mathbf{a}_k \;=\; \mathbf{a}_{k+1},$$

thus $\mathbf{a}_i \leq \mathbf{a}_j$, $i \leq j$, and similarly for $\mathbf{b}_i$ and $\mathbf{c}_i$.

Now we show that

$$\mathbf{g}(\mathbf{a}_k) \;\leq\; \mathbf{a}_{k+1}^{\blacksquare\square} \qquad \mathbf{h}(\mathbf{a}_k) \;\leq\; \mathbf{a}_{k+1}^{\blacksquare\blacksquare}. \tag{20}$$

By (13),

$$\begin{aligned}\mathbf{g}(\mathbf{a}_k) &= \mathbf{g}(\mathbf{0}) + \tfrac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{a}_k)\mathbf{a}_k \\ &= \mathbf{f}(\mathbf{0})^{\blacksquare} + \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_k)^{\blacksquare\blacksquare}\mathbf{a}_k \\ &= \mathbf{a}_0^{\blacksquare} + (\tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_k)\mathbf{a}_k)^{\blacksquare} \\ &\leq (\mathbf{a}_k + \tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_k)^*\mathbf{a}_k)^{\blacksquare} \\ &= \mathbf{a}_{k+1}^{\blacksquare}.\end{aligned}$$

The second inequality of (20) follows from a similar argument.

Now we show by induction on j that

$$\mathbf{c}_j(\mathbf{a}_k^{\blacksquare}) \;\leq\; \mathbf{a}_{j+k+1}^{\blacksquare}. \tag{21}$$

For the basis,

$$\begin{aligned}\mathbf{c}_0(\mathbf{a}_k^{\blacksquare}) &= \mathbf{h}(\mathbf{a}_k^{\blacksquare}, \mathbf{0}) \\ &\leq \mathbf{h}(\mathbf{a}_k) \\ &\leq \mathbf{a}_{k+1}^{\blacksquare}.\end{aligned}$$

For the induction step,

$$\begin{aligned}\mathbf{c}_{j+1}(\mathbf{a}_k^{\blacksquare}) &= \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{a}_k^{\blacksquare}, \mathbf{c}_j(\mathbf{a}_k^{\blacksquare}))^*\mathbf{c}_j(\mathbf{a}_k^{\blacksquare}) \\ &\leq \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{a}_k^{\blacksquare}, \mathbf{a}_{j+k+1}^{\blacksquare})^*\mathbf{a}_{j+k+1}^{\blacksquare} \\ &\leq \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{a}_{j+k+1})^*\mathbf{a}_{j+k+1}^{\blacksquare} \\ &= (\tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_{j+k+1})^{\blacksquare\blacksquare})^*\mathbf{a}_{j+k+1}^{\blacksquare} \\ &\leq (\tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_{j+k+1})^*)^{\blacksquare\blacksquare}\mathbf{a}_{j+k+1}^{\blacksquare} \\ &\leq (\tfrac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_{j+k+1})^*\mathbf{a}_{j+k+1})^{\blacksquare} \\ &= \mathbf{a}_{j+k+2}^{\blacksquare}.\end{aligned}$$

Now we show by induction on k that when $m = n-1$, that is, for $|z| = 1$,

$$\tfrac{\partial \mathbf{c}_{k+1}}{\partial \mathbf{y}}(\mathbf{y}) \;\leq\; \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y})). \tag{22}$$

The corresponding result for $m < n-1$ would require some specialized notation even to state. The proof for $m = n-1$ is considerably simpler, so we henceforth restrict ourselves to that case.

First we note that

$$\tfrac{\partial \mathbf{c}_0}{\partial \mathbf{y}}(\mathbf{y}) \;=\; \tfrac{\partial}{\partial \mathbf{y}}(\mathbf{h}(\mathbf{y}, \mathbf{0})) \;=\; \tfrac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{y}, \mathbf{0}). \tag{23}$$

Then

$$\begin{aligned}&\tfrac{\partial \mathbf{c}_{k+1}}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad= \tfrac{\partial}{\partial \mathbf{y}}(\tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\mathbf{c}_k(\mathbf{y})) \\ &\quad= \mathbf{c}_k(\mathbf{y})\tfrac{\partial}{\partial \mathbf{y}}(\tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad= \mathbf{c}_k(\mathbf{y})\tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial}{\partial \mathbf{y}}(\tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad= \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\mathbf{c}_k(\mathbf{y})\tfrac{\partial}{\partial \mathbf{y}}(\tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad\leq \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial}{\partial \mathbf{y}}(\tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))\mathbf{c}_k(\mathbf{y})) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad\leq \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial}{\partial \mathbf{y}}(\mathbf{h}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad= \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*(\tfrac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y})) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y})) \\ &\qquad+ \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \\ &\quad= \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*(\tfrac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y})) + \tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y})),\end{aligned} \tag{24}$$

so it suffices to show

$$\tfrac{\partial \mathbf{c}_k}{\partial \mathbf{y}}(\mathbf{y}) \;\leq\; \tfrac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y}))^*\tfrac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{y}, \mathbf{c}_k(\mathbf{y})).$$

For $k = 0$, this is immediate from (23). For $k > 0$, this follows from (24) and the induction hypothesis.

Now we show by induction on k that

$$\mathbf{b}_k \quad \leq \quad \mathbf{a}^{\blacksquare}_{(k+1)(P+2)}. \tag{25}$$

For the basis,

$$\begin{aligned}
\mathbf{b}_0 &= \widehat{\mathbf{g}}(\mathbf{0}) \\
&= \mathbf{g}(\mathbf{0}, \mathbf{c}_P(\mathbf{0})) \\
&\leq \mathbf{g}(\mathbf{0}, \mathbf{c}_P(\mathbf{a}^{\blacksquare}_0)) \\
&\leq \mathbf{g}(\mathbf{0}, \mathbf{a}^{\blacksquare}_{P+1}) \\
&\leq \mathbf{g}(\mathbf{a}_{P+1}) \\
&\leq \mathbf{a}^{\blacksquare}_{P+2}.
\end{aligned}$$

For the induction step,

$$\begin{aligned}
&\mathbf{b}_{k+1} \\
&= \frac{\partial \widehat{\mathbf{g}}}{\partial \mathbf{y}}(\mathbf{b}_k)^* \mathbf{b}_k \\
&= \frac{\partial}{\partial \mathbf{y}}(\mathbf{g}(\mathbf{y}, \mathbf{c}_P(\mathbf{y})))^* \mathbf{y} \mid_{\mathbf{y}=\mathbf{b}_k} \\
&= (\frac{\partial \mathbf{g}}{\partial \mathbf{y}}(\mathbf{y}, \mathbf{c}_P(\mathbf{y})) \\
&\quad + \frac{\partial \mathbf{g}}{\partial \mathbf{z}}(\mathbf{y}, \mathbf{c}_P(\mathbf{y})) \frac{\partial \mathbf{c}_P}{\partial \mathbf{y}}(\mathbf{y}))^* \mathbf{y} \mid_{\mathbf{y}=\mathbf{b}_k} \\
&= (\frac{\partial \mathbf{g}}{\partial \mathbf{y}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)) \\
&\quad + \frac{\partial \mathbf{g}}{\partial \mathbf{z}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)) \frac{\partial \mathbf{c}_P}{\partial \mathbf{y}}(\mathbf{b}_k))^* \mathbf{b}_k \\
&\leq (\frac{\partial \mathbf{g}}{\partial \mathbf{y}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)) \\
&\quad + \frac{\partial \mathbf{g}}{\partial \mathbf{z}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)) \frac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{b}_k, \mathbf{c}_{P-1}(\mathbf{b}_k))^* \\
&\quad \cdot \frac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{b}_k, \mathbf{c}_{P-1}(\mathbf{b}_k)))^* \mathbf{b}_k \\
&\leq (\frac{\partial \mathbf{g}}{\partial \mathbf{y}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)) \\
&\quad + \frac{\partial \mathbf{g}}{\partial \mathbf{z}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)) \frac{\partial \mathbf{h}}{\partial \mathbf{z}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k))^* \\
&\quad \cdot \frac{\partial \mathbf{h}}{\partial \mathbf{y}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k)))^* \mathbf{b}_k \\
&= (\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k))^{\blacksquare\blacksquare} \\
&\quad + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k))^{\blacksquare\blacksquare} \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k))^{\blacksquare\blacksquare *} \\
&\quad \cdot \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k))^{\blacksquare\blacksquare})^* \mathbf{b}_k \\
&\leq \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{b}_k, \mathbf{c}_P(\mathbf{b}_k))^{*\blacksquare\blacksquare} \mathbf{b}_k \\
&\leq \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}^{\blacksquare}_{(k+1)(P+2)}, \mathbf{c}_P(\mathbf{a}^{\blacksquare}_{(k+1)(P+2)}))^{*\blacksquare\blacksquare} \mathbf{a}^{\blacksquare}_{(k+1)(P+2)} \\
&\leq \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}^{\blacksquare}_{(k+1)(P+2)}, \mathbf{a}^{\blacksquare}_{(k+2)(P+2)-1})^{*\blacksquare\blacksquare} \mathbf{a}^{\blacksquare}_{(k+1)(P+2)} \\
&\leq \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_{(k+2)(P+2)-1})^{*\blacksquare\blacksquare} \mathbf{a}^{\blacksquare}_{(k+2)(P+2)-1} \\
&\leq (\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{a}_{(k+2)(P+2)-1})^* \mathbf{a}_{(k+2)(P+2)-1})^{\blacksquare} \\
&= \mathbf{a}^{\blacksquare}_{(k+2)(P+2)}.
\end{aligned}$$

It follows from (21) and (25) that

$$\begin{aligned}
\mathbf{b}_M &\leq \mathbf{a}^{\blacksquare}_{(M+1)(P+2)}, \\
\mathbf{c}_P(\mathbf{b}_M) &\leq \mathbf{a}^{\blacksquare}_{(M+2)(P+2)-1}.
\end{aligned}$$

Taking $m = n-1$, Theorem 4.1 says that $P = 1$ suffices. Thus the N in the statement of the theorem is bounded by $(M+2)(P+2) - 1 = 3M + 5$. This gives the following recurrence for N as a function of n:

$$\begin{aligned}
N(1) &= 1 \\
N(n+1) &= 3N(n) + 5
\end{aligned}$$

with solution $N(n) = (7 \cdot 3^n - 5)/2$. □

Acknowledgements

We thank the anonymous referees for their valuable criticism. The support of the National Science Foundation under grant CCR-9708915 is gratefully acknowledged.

References

[1] J. H. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall, London, U.K., 1971.

[2] K. Iwano and K. Steiglitz. A semiring on convex polygons and zero-sum cycle problems. *SIAM J. Comput.*, 19(5):883–901, 1990.

[3] S. C. Kleene. Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, pages 3–41. Princeton University Press, Princeton, N.J., 1956.

[4] D. Kozen. *The Design and Analysis of Algorithms*. Springer-Verlag, New York, 1991.

[5] D. Kozen. A completeness theorem for Kleene algebras and the algebra of regular events. *Infor. and Comput.*, 110(2):366–390, May 1994.

[6] D. Kozen. Kleene algebra with tests. *Transactions on Programming Languages and Systems*, pages 427–443, May 1997.

[7] D. Kozen. On the complexity of reasoning in Kleene algebra. In *Proc. 12th Symp. Logic in Comput. Sci.*, pages 195–202, Los Alamitos, Ca., June 1997. IEEE.

[8] W. Kuich. The Kleene and Parikh theorem in complete semirings. In T. Ottmann, editor, *Proc. 14th Colloq. Automata, Languages, and Programming*, volume 267 of *Lecture Notes in Computer Science*, pages 212–225, New York, 1987. EATCS, Springer-Verlag.

[9] R. J. Parikh. On context-free languages. *J. Assoc. Comput. Mach.*, 13(4):570–581, 1966.

[10] D. L. Pilling. Commutative regular equations and Parikh's theorem. *J. London Math. Soc.*, 6(4):663–666, June 1973.

[11] I. Stewart. *Galois Theory*. Chapman and Hall, London, 1973.

[12] B. L. van der Waerden. *Algebra*, volume 1. Frederick Ungar, 1970.