TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 352, Number 5, Pages 2049–2075
S 0002-9947(99)02300-4
Article electronically published on September 8, 1999

CYCLING IN PROOFS AND FEASIBILITY

A. CARBONE

Abstract. There is a common perception by which small numbers are considered more concrete and large numbers more abstract. A mathematical formalization of this idea was introduced by Parikh (1971) through an *inconsistent* theory of feasible numbers in which addition and multiplication are as usual but for which some very large number is defined to be not feasible. Parikh shows that sufficiently short proofs in this theory can only prove true statements of arithmetic. We pursue these topics in light of logical flow graphs of proofs (Buss, 1991) and show that Parikh's lower bound for concrete consistency reflects the presence of *cycles* in the logical graphs of short proofs of feasibility of large numbers. We discuss two concrete constructions which show the bound to be optimal and bring out the dynamical aspect of formal proofs.

For this paper the concept of feasible numbers has two roles, as an idea with its own life and as a vehicle for exploring general principles on the dynamics and geometry of proofs. Cycles can be seen as a measure of how complicated a proof can be. We prove that short proofs must have cycles.

1. Introduction

In philosophical works by Mannoury [21], Poincaré [27] and Wittgenstein [34] there appears already the idea that there is a genuine difference between our understanding of, say, the number 10 and the number $67^{257^{729}}$.

In [1], Bernays observes that intuitionism as well as ordinary mathematics contains a strong idealization in the facts that numbers such as 10 and $67^{257^{729}}$ are treated as objects of the same kind, even though arithmetical operations do not have a *concrete* meaning for very large numbers. He suggests strict-finitism as a conceivable position in philosophy of mathematics.[1]

Received by the editors June 13, 1997 and, in revised form, January 21, 1998.

1991 *Mathematics Subject Classification.* Primary 03F07, 03F20, 03F05, 03H15, 68R10.

Key words and phrases. Feasible numbers, cut elimination, cycles in proofs, structure of proofs, complexity of proofs.

Partially supported by the Lise-Meitner Stipendium # M00187-MAT (Austrian FWF).

[1]Strict-finitist philosophy of mathematics insists that the meanings of all terms appearing in mathematical statements must be given in relation to *constructions* which we can effectively carry out, and of our capacity to recognize *in practice* such constructions as providing proofs of those statements. No construction that is too complex or too lengthy to effect in practice, is accepted by strict-finitism. In addition to the one discussed in this introduction, several attempts to develop a formal strict-finitist theory have been made. The most developed is the *Alternative Set Theory* by Vopenka [33] and his school in Prague. See also [28], the *Predicative Arithmetic* by Nelson [22] and the theory of *Vague Predicates* by Dummett [13, 12] and Parikh [26]. For a general discussion on the strict-finitist program see [16] and for its relations with intuitionism see [31].

In [14], Esenin-Volpin declares his adherence to strict-finitism and starts some attempts in the reconstruction of mathematics along the new lines (see also [15]). He refuses the concept that natural numbers are closed under simple arithmetical operations, such as exponentiation and develops a rich and complicated theory of 'concrete' mathematical activity (leading for instance, as he claims, to a 'concrete' proof of the consistency of the Zermelo-Fraenkel set theory). He introduces the concept of a set of *feasible* numbers closed under a successor but bounded by a given natural number, say 10^{12}. The notion he introduces appears to be paradoxical and indeed it is an abstract form of the classical paradoxes of the bald man (adding one hair at a time can never turn a bald man into a non-bald man) and the heap (by removing a single grain at a time you will still have a heap).

The first precise proof-theoretical result about the 'concrete' correctness of the notion of feasible numbers was given by Parikh [25] that proves the 'concrete' consistency of the theory defined adding to Peano Arithmetic the above mentioned properties of feasible numbers. Namely, to the language of PA (defined by the symbols $0, s, +, *, <, =$) add the symbol F and to PA add the axioms[2]

(i) $F(0)$,
(ii) $F(x) \to F(s(x))$,
(iii) $F(x) \wedge F(y) \to F(x + y)$,
(iv) $F(x) \wedge F(y) \to F(x * y)$,
(v) $x = y \to (F(x) \to F(y))$,
(vi) $F(x) \wedge y < x \to F(y)$,
(vii) $\neg F(\theta)$,

where θ is some fixed variable-free term in the language of PA. Call this theory PA_F.

The theory PA_F is clearly inconsistent (because of the presence of axiom (vii)), but it can be shown that proofs with a 'small' number of formulas cannot prove inconsistency. In fact, such proofs can only establish truths in the original language of PA. Proving this, Parikh gives a precise non-elementary bound on how small a proof should be. Other bounds were computed by Gavrilenko (unpublished work) and Dragalin [11] using different techniques. We show that Parikh's bound is optimal. Moreover, we will see that from an arbitrary 'small' proof in PA_F of an F-free formula, there is no way to 'extract' a proof of it in PA. (In fact, one can formalize in PA_F 'small' proofs of true inequalities, which use axiom (vii) and where the structure of the non-feasible term θ is not taken into account. There is neither explicit nor implicit construction of the term in the proof.) This might suggest that there is some speed-up of the theory PA_F over PA (or intuitively, that thinking of large numbers induces actual shortcuts in our reasoning). In Section 6 we will indicate how such shortcuts depend on axiom (vi).

All of our results are based on a graph-theoretical analysis of formal proofs. We make use of the notion of *logical flow graph* (a directed graph defined from a proof by tracing the occurrences of logical formulas in it) introduced by Buss[3] [3] to study

[2] Note that with the symbol PA we indicate the Hilbert style formalization of Peano Arithmetic, where tautologies are axioms. In particular, observe that axiom (iii) can be derived from (iv),(vi) and the axioms of PA; axiom (v) can be derived from (ii) and (vi).

[3] The notion has been introduced by Buss to prove that the k-provability problem (i.e. given a formula A and an integer k, to determine if A has a proof with k or fewer lines) for a particular formalization of first order logic is undecidable. Concepts similar to Buss' definition of logical flow

how the influence of a formula spreads through a proof. We consider the graph-theoretical notion of *cycle* over a directed graph, we show that the logical graph of 'short' proofs of the feasibility of large numbers *must* contain cycles (Theorem 5.1), and that the elimination of these cycles corresponds to a *non-elementary* expansion of a proof in terms of the number of steps.

Although we have stressed so far the historical context of feasible numbers, the concept of *feasibility* provides a setting in which dynamics in proofs and the combinatorics of cut elimination can be seen more clearly ([4, 5]). Our examples of short proofs of the feasibility of large numbers in Section 4 and the *geometry* of the logical graphs, bring to light basic points in this respect. The concept also provides a tool for linking the combinatorics of proofs with other mathematical structures (as in [8]).

The plan of this paper is as follows: in Section 2 we review the Gentzen's Sequent Calculus and we present the sequent theory of feasible numbers; Section 3 contains the definition of logical flow graph; in Section 4 we study the graph-theoretical properties of short proofs, we discuss two constructions of large numbers and we show that Parikh's bound is optimal; in Section 5 we show that cycles are needed to have speed-up in the construction of large numbers; in Section 6 we discuss how true statements can be obtained in some implicit way through the use of axiom (vii).

The author thanks Sam Buss, Ehud Hrushovski and Rohit Parikh for comments on an earlier version of this work, and Stephen Semmes for helpful remarks and stimulating conversations on this more recent manuscript.

2. The sequent theory of feasible numbers

We work with a formalization of Peano Arithmetic in terms of sequents and we will denote such theory T. Actually, the following results hold for theories containing a rather modest portion of T. Namely, it is sufficient that the theory be strong enough to prove arithmetical equalities and inequalities.

Before introducing axioms and rules of inference for T, let us recall some basic definitions concerning the *sequent calculus*, i.e. the formulation of first order logic due to Gentzen. For a detailed exposition the reader can refer to [30, 17, 19].

The sequent calculus is formulated in a first order language (possibly containing the equality symbol) with logical symbols $\wedge, \vee, \neg, \supset, \exists$ and $\forall$; it has *free* variables denoted $a, b, c, \ldots$ and *bound* variables denoted $x, y, z, \ldots$. As usual, *terms* are formed from constant symbols, free variables and function symbols; *semi-terms* are like terms but may contain bound variables as well. *Formulas* are defined as usual with the condition that only bound variables may be quantified and only free variables may appear free. *Semi-formulas* are defined as formulas except that both bound and free variables may appear free in it; one can observe that a sub-formula of a formula is a semi-formula and in fact a sub-formula of a semi-formula is a semi-formula.

A *sequent* is a line of the form

$$A_1, \ldots, A_k \rightarrow B_1, \ldots, B_l$$

graphs have been independently and previously introduced by Jean-Yves Girard who discusses tracing the flow of formulas through Linear Logic proofs ([17]).

where the A_i's and B_j's are formulas; its intended meaning is $\bigwedge_i A_i \supset \bigvee_j B_j$. We permit k and l to be zero. A sequence of formulas separated by commas is a *cedent*; in the sequent above, $A_1, \dots, A_k$ is the *antecedent* and $B_1, \dots, B_l$ is the *succedent*. We will often refer to antecedent and succedent in a sequent using capital letters of the Greek alphabet. For instance $\Gamma \to \Delta$ denotes a sequent. In the following we will intend a sequence $A_1, \dots, A_k$ to be a *multi-set* of formulas, i.e. finite (possibly empty) set of formulas, in which repetitions of some formulas are admitted; the order of formulas in a multi-set is not essential but for every member of the multi-set the number of its occurrences is important. By the symbol Γ, Δ we denote the sum of the multi-sets Γ and Δ (i.e. the multi-set containing all formulas in Γ and Δ so that, if n_1 and n_2 are the number of occurrences of a formula A in Γ and Δ respectively, then $n_1 + n_2$ is the number of occurrences of A in the union Γ, Δ). The multi-set A, Γ is obtained from Γ by adjoining the formula A. For short we will denote Γ_1, Γ_2 as $\Gamma_{1,2}$.

It should be pointed out that a proof in the sequent calculus is intended to be a *tree* of sequents; each sequent must either be an axiom (in this case the sequent is labeling a leaf of the tree) or be derived by one of the rules of inference we will give below (the sequent is a label for an internal node of the tree). In a sequent calculus proof every occurrence of a sequent in the proof other than the end-sequent is used exactly once as a premise of an inference. Notice that a proof could be defined as sequence of sequents; but obviously any proof defined as such can be transformed into a tree-like proof by duplicating subproofs to derive intermediate results multiple times.

Let us now formalize the sequent theory of feasible numbers. To the language of arithmetic (defined by the symbols $0, s, +, *, <, =$) add the symbol F. To T we add axioms and rules of inference describing the behavior of the new predicate F. The theory will be called T_F and the formulas containing the symbol F will be referred to as F-formulas. There are five kinds of axioms

(i) *logical axioms* of the form $A, \Gamma \to \Delta, A$, where A is an F-free formula;
(ii) *special logical axioms* of the form $F(t), \Gamma \to \Delta, F(t)$, where t is some arbitrary term;
(iii) *equality axioms* of the form
- $\Gamma \to \Delta, t = t$;
- $t_1 = s_1, \dots, t_k = s_k, A(t_1, \dots, t_k), \Gamma \to \Delta, A(s_1, \dots, s_k)$ where A is an atomic F-free formula;

(iv) *PA-axioms* of the form $\Gamma \to \Delta, A$, where A is an arbitrary non-logical axiom of PA (note that the formula A is F-free); without loss of generality, we assume the axioms $\Gamma \to \Delta, \neg s(x) = x$ and $\Gamma \to \Delta, \neg s(x) < x$ be formalized as $s(x) = x, \Gamma \to \Delta$ and $s(x) < x, \Gamma \to \Delta$, respectively;
(v) a *special axiom* of the form $\Gamma \to \Delta, F(0)$.

The rules of inference are divided into three groups: the *logical rules*, the *structural rules* (note that the cut rule and the contraction rule will be the only structural rules of the calculus), and the *special rules* concerning the predicate F.

The *logical rules* are the following:

$$\neg : \textit{left} \quad \frac{\Gamma \to \Delta, A}{\neg A, \Gamma \to \Delta} \qquad \neg : \textit{right} \quad \frac{A, \Gamma \to \Delta}{\Gamma \to \Delta, \neg A},$$

$$\wedge : \textit{right} \quad \frac{\Gamma_1 \to \Delta_1, A \quad \Gamma_2 \to \Delta_2, B}{\Gamma_{1,2} \to \Delta_{1,2}, A \wedge B},$$

$$\wedge : \textit{left} \quad \frac{A, B, \Gamma \to \Delta}{A \wedge B, \Gamma \to \Delta} \qquad \frac{A, B, \Gamma \to \Delta}{B \wedge A, \Gamma \to \Delta},$$

$$\vee : \textit{left} \quad \frac{A, \Gamma_1 \to \Delta_1 \quad B, \Gamma_2 \to \Delta_2}{A \vee B, \Gamma_{1,2} \to \Delta_{1,2}},$$

$$\vee : \textit{right} \quad \frac{\Gamma \to \Delta, A, B}{\Gamma \to \Delta, A \vee B} \qquad \frac{\Gamma \to \Delta, A, B}{\Gamma \to \Delta, B \vee A},$$

$$\supset : \textit{left} \quad \frac{\Gamma_1 \to \Delta_1, A \quad B, \Gamma_2 \to \Delta_2}{A \supset B, \Gamma_{1,2} \to \Delta_{1,2}},$$

$$\supset : \textit{right} \quad \frac{A, \Gamma \to \Delta, B}{\Gamma \to \Delta, A \supset B} \qquad \frac{B, \Gamma \to \Delta, A}{\Gamma \to \Delta, B \supset A},$$

$$\exists : \textit{left} \quad \frac{A(b), \Gamma \to \Delta}{(\exists x)A(x), \Gamma \to \Delta} \qquad \exists : \textit{right} \quad \frac{\Gamma \to \Delta, A(t)}{\Gamma \to \Delta, (\exists x)A(x)},$$

$$\forall : \textit{left} \quad \frac{A(t), \Gamma \to \Delta}{(\forall x)A(x), \Gamma \to \Delta} \qquad \forall : \textit{right} \quad \frac{\Gamma \to \Delta, A(b)}{\Gamma \to \Delta, (\forall x)A(x)}.$$

In the $\exists$: *left* and $\forall$: *right* inferences the free variable b is called the *eigenvariable* and must not appear in the lower sequent. The variable x must be freely substitutable into A for all four quantifier inferences.

The *structural rules* are:

$$\textit{Cut} \quad \frac{\Gamma_1 \to \Delta_1, A \quad A, \Gamma_2 \to \Delta_2}{\Gamma_{1,2} \to \Delta_{1,2}},$$

$$\textit{Contraction} \quad \frac{\Gamma \to \Delta, A, A}{\Gamma \to \Delta, A} \qquad \frac{A, A, \Gamma \to \Delta}{A, \Gamma \to \Delta},$$

The *special rules* are:

$$F : \textit{equality} \quad \frac{\Gamma_1 \to \Delta_1, r = t \quad \Gamma_2 \to \Delta_2, F(r)}{\Gamma_{1,2} \to \Delta_{1,2}, F(t)},$$

$$F : \textit{inequality} \quad \frac{\Gamma_1 \to \Delta_1, t < r \quad \Gamma_2 \to \Delta_2, F(r)}{\Gamma_{1,2} \to \Delta_{1,2}, F(t)},$$

$$F : \textit{successor} \quad \frac{\Gamma \to \Delta, F(t)}{\Gamma \to \Delta, F(s(t))},$$

$$F : \textit{plus} \quad \frac{\Gamma_1 \to \Delta_1, F(r) \quad \Gamma_2 \to \Delta_2, F(t)}{\Gamma_{1,2} \to \Delta_{1,2}, F(r + t)},$$

$$F : \textit{times} \quad \frac{\Gamma_1 \to \Delta_1, F(r) \quad \Gamma_2 \to \Delta_2, F(t)}{\Gamma_{1,2} \to \Delta_{1,2}, F(r * t)},$$

$$F : \textit{elimination} \quad \frac{\Gamma \to \Delta, F(\theta)}{\Gamma \to \Delta}.$$

The F:*elimination* rule plays a special role in the theory, it makes the theory inconsistent.

In axioms, formulas other than those in Γ, Δ are called *distinguished*. The formula occurrence, which is introduced explicitly into the lower sequent of a given rule of inference is called the *main* formula of the inference, and the formula(s), which are distinguished in the upper sequent(s) are the *auxiliary* formula(s). For example, in the $\vee$:*left* rule, the formula $A \vee B$ is the main formula and the formulas

A and B are the auxiliary formulas. The other formulas of a sequent (i.e. the formulas in Γ, Δ) are called *side formulas*.

The cut rule has no main formula; its auxiliary formulas are also called *cut-formulas*. An application of the cut rule will simply be called a *cut*.

The auxiliary formulas of a contraction rule are also called *contraction formulas*.

We do not need the usual Weakening and Exchange rules in our calculus (as they appear in Gentzen's original formalization), because it uses multi-sets of formulas instead of sequences. Non-distinguished formulas appearing in the axioms (i.e. formulas occurring in Γ, Δ) play the role of formulas introduced by Weakening. In Section 5 we will state lemmas about addition and elimination of weak formulas in proofs.

To measure complexity of proofs we will use in the sequel is the *number of lines*, i.e. the number of nodes in the tree-like structure of the proof.

3. Logical flow graphs for T_F

In [3], Sam Buss introduces the notion of *logical flow graph* to study how the influence of a formula spreads through a proof in LK. Following his introduction, we present the notion of logical flow graph formulated for our sequent theory. For an extensive treatment of the notion of logical flow graphs for the propositional calculus, see [5]; the notion is used to give a new proof of the Craig Interpolation Theorem, derive results on the complexity of interpolation and study the dynamics of cut elimination procedures.

Let Π be a proof in T_F. An *s-formula* is an *occurrence* of a sub-formula of a formula in Π (here, '*s*-' stands for '*semi*-' or '*sub*-'). It has to be emphasized that an *s*-formula is an *occurrence* of a sub-formula in the proof as opposed to the sub-formula itself which may occur many times in the proof. A formula A is a *variant* of B if A can be obtained from B by changing some of the terms in B. The *logical flow graph* (formally defined below) is a directed graph whose nodes are *s*-formulas in Π; two *s*-formulas will be connected by an edge only if they are variants of each other; any two *s*-formulas connected by an edge will be in (distinct) sequents of some inference or will both be in an axiom on opposite sides of the sequent arrow.

We define the logical flow graph by specifying the edges.

First, in an axiom $A, \Gamma \to \Delta, A$ there is an edge directed from the left-hand A to the right-hand A. In an equality axiom $t_1 = s_1, \ldots, t_k = s_k, A(t_1, \ldots, t_k), \Gamma \to \Delta, A(s_1, \ldots, s_k)$, there is an edge directed from $A(t_1, \ldots, t_k)$ to $A(s_1, \ldots, s_k)$. There is no edge defined for a PA-axiom, a special axiom and an equality axiom $\Gamma \to \Delta, t = t$.

Second, in any logical, structural and special inferences listed above, there is an edge directed from each side formula in the antecedent Γ of the *lower* sequent to the corresponding side formula in Γ of the *upper* sequent(s). There is an edge directed from each formula in the succedent Δ in the *upper* sequent to the corresponding formula of Δ in the *lower* sequent.

Third, in any logical inference or in a contraction rule, if A (or B) is an auxiliary formula which appears in the succedent of an upper sequent of an inference, then there is an edge directed from that A (or B) to the corresponding *s*-formula in the lower sequent. If A (or B) is an auxiliary formula which appears in the antecedent of an upper sequent of an inference then there is an edge directed towards that A (or B) from the corresponding *s*-formula in the lower sequent.

Fourth, in any special rule (except the critical rule) there is an edge directed from each auxiliary F-formula in the upper sequent(s) to the main F-formula in the lower sequent. There is no outgoing edge from the equality and inequality auxiliary formulas in the upper sequent of F:*equality* and F:*inequality*, respectively.

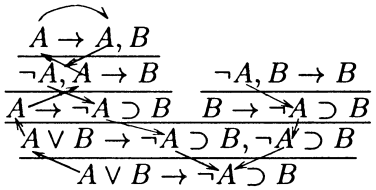
Fifth, in a cut inference there is an edge directed from the cut formula A in the succedent of the left-hand upper sequent to the occurrence of A in the antecedent of the right-hand upper sequent.

Sixth, suppose there is a directed edge from an s-formula A_1 to A_2 and suppose B_1 is a sub-formula of A_1. Since A_1 and A_2 are variants, there is a sub-formula B_2 of A_2 which corresponds to the subformula B_1 of A_1; the s-formulas B_1 and B_2 are, of course, variants. If B_1 occurs positively in A_1, then there is an edge from B_1 to B_2. If B_1 occurs negatively in A_1, then there is an edge from B_2 to B_1. Recall that B occurs positively (negatively) in A if B occurs an even (odd) number of times in the scope of a negation or in the left-hand operand of an implication. Clearly B_1 occurs positively in A_1 if and only if B_2 occurs positively in A_2. This concludes the definition of logical flow graph.

As an example, consider the following proof:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \to A, B}{\neg A, A \to B}}{A \to \neg A \supset B} \quad \dfrac{\neg A, B \to B}{B \to \neg A \supset B}}{A \vee B \to \neg A \supset B, \neg A \supset B}}{A \vee B \to \neg A \supset B}$$

with logical flow graph restricted to the formula A (edges for $\neg A$, B and $\neg A \supset B$ are not indicated)



A formula B occurs positively (negatively) in a sequent $\Gamma \to \Delta$ if B occurs negatively (positively) in a formula of Γ or positively (negatively) in a formula of Δ. If not otherwise indicated, in the following we will intend a positive or negative occurrence of a formula to be defined relatively to sequents. Notice that in a logical flow graph there are four kinds of edges: edges connecting positive occurrences, that are directed downwards; edges connecting negative occurrences, that are directed upwards; edges defined on axioms, that are directed from negative occurrences towards positive occurrences; edges defined on cut-formulas, that are directed from positive occurrences towards negative occurrences.

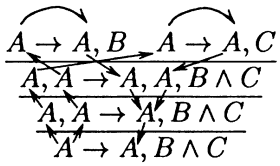
If a proof is cut-free, its logical flow graph will contain only three kinds of edges. From an easy checking of the rules of the calculus, it easily follows that contraction, F:*times* and F:*plus* rules are the only ways in which two distinct edges can be directed to/from the same node of a logical flow graph. Hence, each node of a logical flow graph, whenever labeled by a positive (negative) occurrence can have at most 2 incoming (outgoing) edges and at most 1 outgoing (incoming) edge. From

this, it follows that each node of a logical flow graph is at most of *degree* 3, where the degree of a node is the number of edges which depart or arrive at the node.

A connected subgraph of a logical flow graph in general is not a tree. Take for instance the following proof

$$\dfrac{\dfrac{\dfrac{A \to A, B \quad A \to A, C}{A, A \to A, A, B \wedge C}}{A, A \to A, B \wedge C}}{A \to A, B \wedge C}$$

and consider the connected subgraph tracing the logical relations between the occurrences of the formula A in it



We call any sequence of consecutive edges in the logical flow graph $\mathcal{L}$ of $\Pi : S$ a *path* (two consecutive edges in this sequence should meet in a vertex which is a source for one and a sink for the other). We call *F-paths* those paths whose variants are F-atomic formulas. We call any path starting and ending with two (distinct) s-formulas occurring in S a *bridge*. We call *direct path* a logical path which passes through either positive or negative occurrences *only*. Notice that a direct path cannot cross an axiom or a cut, since this would force the path through both positive and negative formula occurrences. Moreover, notice that the number of variants in a direct path is bounded by the height of Π. A *cycle* is a directed path with common starting and ending node (or equivalently, a path starting from an occurrence of a formula and going back to it).

A logical flow graph is *acyclic* when it does not contain cycles. In [5] it is proved that cut-free proofs are acyclic (this follows from the fact that cut-free proofs contain only three types of directed edges). There it is also proved that contraction-free proofs (possibly containing cuts) are also acyclic. For an analysis of cycles in proofs see [7].

In the following, we focus on the logical relations between *atomic* formulas; thus, whenever we will refer to s-formulas we will intend them to be atomic.

4. Short proofs and cycles

Short proofs can encode large constructions. This is a basic fact which remains far from being well-understood. The explicit constructions, that we usually call *direct* or *cut-free* constructions, may be much larger than the proofs we actually make. Short proofs can be transformed effectively into direct constructions through cut elimination, for instance (this was proved by Gentzen in 1939 [17, 30]; see [9] for an introduction to the combinatorics and complexity of cut elimination). When we perform cut elimination, we are basically deforming logical paths in a proof, at times splitting these paths, as described in [5]. We never add new paths though, to the contrary we might lose some of them. This means that logical paths occurring in short proofs may *have* to wind around in complicated ways, crossing themselves and making cycles, for instance. This happens because the proof is short and at the

same time should encode all the paths occurring in a direct construction. It is this tracing of the direct construction in a small space that leads to interesting *dynamics*. In [6] this small space is investigated and it is shown that a proof containing cycles can be transformed into an acyclic proof containing *spirals* which is only elementary larger than the original proof (see also Remark 5.6).

We will see in this section how this idea of 'small space' is captured well by the construction of large numbers in arithmetic. We will start with the usual approach to proofs through an analysis which is based on logical rules, and we will describe afterwards the dynamical aspects by looking at logical paths. We will describe a way to look at a proof which cannot be captured by the usual induction arguments. A proof for us will be a graph in the plane, and its essential elements will be logical paths and not logical rules.

Let $N(\Pi)$ denote the number of sequents in a proof Π. Parikh's main result can be formulated as follows:

Theorem 4.1 (Concrete consistency of T_F)**.** *Let Π be a proof in T_F of the sequent $\rightarrow B$. The formula B does not contain F. There exists a primitive recursive function f such that if $val(\theta) > f(\Pi)$, then there is a proof $\Pi' : \rightarrow B$ in T. Here $f(\Pi)$ indicates that the arguments of f are parameters in the proof Π.*

The lower bound on the value $val(\theta)$ of the term θ is computed by Parikh using the Hilbert-Ackermann's ϵ-substitution method and by Dragalin using cut elimination. Both proofs give a non-elementary primitive recursive lower bound for $val(\theta)$. If we define the function $e_t(0, n) = n$, $e_t(i+1, n) = t^{e_t(i,n)}$ (for some natural number $t > 1$), then

- Parikh's bound (computed for a Hilbert's style formalization of T_F) can be expressed as $val(\theta) > e_2(e_2(\frac{1}{2} \cdot k \cdot (k+1), r) \cdot n, 1)$, where k is the number of axioms of the form $A(t) \supset \exists x A(x)$ occurring in the proof and such that A contains the F symbol, r is the number of quantifiers occurring in formulas A of the form described above and n is the number of occurrences of the special axioms of the form (ii) and (v) in Π;[4]
- Dragalin's bound is expressed as $val(\theta) > e_2(e_4(39 \cdot N(\Pi), 39 \cdot N(\Pi)), 1)$.

As one can see, Parikh's bound depends on the complexity of F-formulas in the deduction Π and on the number of some special axioms in it, while Dragalin's bound depends on the total number of formulas in Π.

We will show that there is no function f sensitive *only* to the number of occurrences of special axioms in the proof, and satisfying Theorem 4.1. To see this, in section 4.1 we describe how to build a proof of the sequent $\rightarrow F(n)$ containing *only three* applications of special rules, for *any* arbitrary representation of the natural number n in the language of arithmetic. In section 4.2 we give a second example to illustrate the power of nested quantifications for the construction of short proofs of the feasibility of large numbers: namely, we give a formalized proof that a term δ_n of value $e_2(n, 2)$ is feasible in $\mathcal{O}(2^n)$ number of steps by using formulas with n nested quantifiers and logical axioms defined on atomic formulas. In case logical axioms are defined on arbitrary formulas, we obtain a proof of $\mathcal{O}(n)$ lines. (Some

[4] For the theory of feasible numbers defined as PA_F but not containing axioms (iii) and (iv), Parikh improves the lower bound to $val(\theta) > n \cdot e_2(\frac{1}{2} \cdot k \cdot (k+1), r)$, while Dragalin obtains $e_2(e_4(30 \cdot N(\Pi), 30 \cdot N(\Pi)), 1)$.

considerations in this direction were developed in [2]. The power of nested quantification in proofs containing cuts was first observed by Tseitin [32]; see also Orevkov [24] and Statman [29].)

In the examples we see the role of cycles in short proofs of $F(n)$ and $F(\delta_n)$ (for all $n > 1$), and we shall see that their elimination corresponds essentially to the elimination of quantifiers of cut-formulas. We shall give a precise bound for the cost of the elimination of these cycles in terms of the number of lines in the proofs.

4.1. **A first example.** Define the term $p(0,x) = x$, $p(n+1,x) = p(n,x) * p(n,x)$ (for some term x). Clearly $val(p(n,x)) = x^{2^n}$. The aim of this section is to show that for all $n \geq 1$ there exists a T_F-proof $\Pi : \rightarrow F(p(2^n, ss0))$ containing exactly one application of the F:*times* rule and two applications of F:*successor*, and such that $N(\Pi) \leq 10 \cdot n + 7$. In other words, we show that essentially *only* one application of the F:*times* rule is enough to prove the feasibility of any number. This implies that there is no function f depending *only* on the number k of F-rules used in the T_F-proof, such that $f(k)$ bounds the value of θ in Theorem 4.1.

The proof we describe contains exactly $2 \cdot n$ cycles and it is possible to show that to eliminate these cycles one pays an *exponential* price in terms of number of lines of the proof. More precisely, for all $n \geq 1$ there exists a T_F-proof $\Pi' : \rightarrow F(p(2^n, ss0))$ containing exactly n applications of the F:*times* rule with no cycles and such that $N(\Pi)$ is $\mathcal{O}(2^n)$. Theorem 5.1 will show that one cannot do better.

Notice that, for all $n \geq 1$ there exists a T_F-proof $\Pi'' : \rightarrow F(p(2^n, ss0))$ containing exactly 2^n applications of the F:*times* rule with no cuts (and therefore no cycles) and such that $N(\Pi)$ is $\mathcal{O}(2^n)$. To see this, think of the complete binary tree (or think of the tree-like proof) of height n having as leaves the derivable sequent $\rightarrow F(ss0)$ and as internal nodes the sequents obtained by applying the F:*times* rule.

To simplify the exposition we will give a proof of $\rightarrow F(p(2^2, ss0))$ containing four cycles. It will be clear that a proof of $\rightarrow F(p(2^n, ss0))$ containing $2 \cdot n$ cycles can be derived similarly. Let us denote the terms $p(2^2,x), p(2,x), p(1,x)$ with the symbols $x^{2^{2^2}}, x^{2^2}, x^2$ to remind us of their value (i.e. x^{2^4}, x^{2^2}, x^2)

$$\dfrac{\dfrac{\begin{matrix}\Pi_1 \\ \rightarrow F(ss0)\end{matrix} \quad \begin{matrix}\Pi_2 \\ \rightarrow F(ss0) \supset F((ss0)^{2^{2^2}})\end{matrix}}{\rightarrow F(ss0) \wedge (F(ss0) \supset F((ss0)^{2^{2^2}}))} \quad \begin{matrix}\Pi_3 \\ F(ss0) \wedge (F(ss0) \supset F((ss0)^{2^{2^2}})) \rightarrow F((ss0)^{2^{2^2}}).\end{matrix}}{\rightarrow F((ss0)^{2^{2^2}})}$$

Subproofs Π_1 and Π_3 are easy to derive with 3 and 6 sequents, respectively. The main part of the proof is Π_2, which, in fact, contains four cycles. It has the following form:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{F(a) \rightarrow F(a) \quad F(a) \rightarrow F(a)}{F(a), F(a) \rightarrow F(a^2)}\, F\text{:}\, times}{F(a) \rightarrow F(a^2)}}{\rightarrow F(a) \supset F(a^2)}}{\rightarrow (\forall x) F(x) \supset F(x^2)} \quad \begin{matrix}\Pi_4 \\ (\forall x) F(x) \supset F(x^2) \rightarrow F(0) \supset F((ss0)^{2^{2^2}})\end{matrix}}{\rightarrow F(0) \supset F((ss0)^{2^{2^2}})}\, cut$$

where Π_4 has the form

$$\dfrac{\begin{array}{cc}\Pi_5 & \Pi_6\\ (\forall x)F(x)\supset F(x^2)\to(\forall x)F(x)\supset F(x^{2^2}) & (\forall x)F(x)\supset F(x^{2^2})\to F(ss0)\supset F((ss0)^{2^{2^2}})\end{array}}{(\forall x)F(x)\supset F(x^2)\to F(0)\supset F((ss0)^{2^{2^2}})}\ cut$$

and Π_5, Π_6 are built respectively as follows

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{F(b)\to F(b)\quad F(b^2)\to F(b^2)}{F(b)\supset F(b^2), F(b)\to F(b^2)}}{(\forall x)F(x)\supset F(x^2), F(b)\to F(b^2)}\quad F(b^{2^2})\to F(b^{2^2})}{(\forall x)F(x)\supset F(x^2), F(b^2)\supset F(b^{2^2}), F(b)\to F(b^{2^2})}}{(\forall x)F(x)\supset F(x^2), F(b^2)\supset F(b^{2^2})\to F(b)\supset F(b^{2^2})}}{(\forall x)F(x)\supset F(x^2), (\forall x)F(x)\supset F(x^2)\to F(b)\supset F(b^{2^2})}\ contraction}{(\forall x)F(x)\supset F(x^2)\to F(b)\supset F(b^{2^2})}}{(\forall x)F(x)\supset F(x^2)\to(\forall x)F(x)\supset F(x^{2^2})}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{F(ss0)\to F(ss0)\quad F((ss0)^{2^2})\to F((ss0)^{2^2})}{F(ss0)\supset F((ss0)^{2^2}), F(ss0)\to F((ss0)^{2^2})}}{(\forall x)F(x)\supset F(x^{2^2}), F(ss0)\to F((ss0)^{2^2})}\quad F((ss0)^{2^{2^2}})\to F((ss0)^{2^{2^2}})}{(\forall x)F(x)\supset F(x^{2^2}), F((ss0)^{2^2})\supset F((ss0)^{2^{2^2}}), F(ss0)\to F((ss0)^{2^{2^2}})}}{(\forall x)F(x)\supset F(x^{2^2}), (\forall x)F(x)\supset F(x^{2^2}), F(ss0)\to F((ss0)^{2^{2^2}})}\ contraction}{(\forall x)F(x)\supset F(x^{2^2}), F(ss0)\to F((ss0)^{2^{2^2}})}}{(\forall x)F(x)\supset F(x^{2^2})\to F(ss0)\supset F((ss0)^{2^{2^2}})}$$

Notice that in Π_2 there are two paths f_2, f_3 from $F(x)$ to $F(x^2)$ in the sequent $\to(\forall x)F(x)\supset F(x^2)$. In Π_4 there are two paths (say f_5, f_6), from $F(x^2)$ to $F(x)$ in the end-sequent $(\forall x)F(x)\supset F(x^2)\to F(0)\supset F((ss0)^{2^{2^2}})$, one passing through the axiom $F(b^2)\to F(b^2)$ (in Π_5, and we call it f_5) and the other, through the axiom $F((ss0)^{2^2})\to F((ss0)^{2^2})$ (in Π_6, and we call it f_6). Notice that the combinations f_2f_5, f_3f_5 and f_2f_6, f_3f_6 form four cycles in Π_3.

By counting, the number of sequents occurring in the proof is $7+2\cdot 10$.

Using the same idea one builds a proof for $p(2^n, ss0)$ by modifying the subproof Π_4 of Π_2 so that the sequent $(\forall x)F(x)\supset F(x^2)\to F(0)\supset F((ss0)^{2^{2^n}})$ is derived. One does this by gluing together n copies of proofs which look like Π_5 (as Π_6) in the manner described above. More precisely, one can prove $(\forall x)F(x)\supset F(x^k)\to(\forall x)F(x)\supset F(x^{k^2})$ for any natural number k in the same manner as the proof of Π_5 above. (Basically the proof substitutes x^k into the x in x^k to get x^{k^2}.) This is our basic building block. This proof has three bridges, one from $F(x^k)$ to the $F(x)$ on the left side of the sequent, one from the $F(x)$ on the right to the $F(x)$ on the left, and one from the $F(x^k)$ on the left to the $F(x^{k^2})$ on the right. We can get a proof of $\to(\forall x)F(x)\supset F(x^{2^{2^n}})$ by connecting n of these proofs using cuts (corresponding to $k=2^{2^j}$, $j=0,\dots,n$), and then combining with the proof of $\to(\forall x)F(x)\supset F(x^2)$ (given in Π_2 above) using a cut again. It is easy to see how the logical flow graphs combine with this series of cuts, and how the $2\cdot n$ cycles arise from the connections of the bridges.

A proof containing n applications of the F:*times* rule that does not contain cycles can be obtained by applying the cut elimination procedure to the quantified formulas of the above proof until each contraction rule corresponding to a quantified cut-formula is eliminated. The elimination of these contractions will induce an exponential expansion of the number of lines in the proof. By observing that the procedure of cut elimination applied to the above proof does not increase the number of contractions on (1 of) cut-formulas, one derives that the expanded proof has $\mathcal{O}(2^n)$ lines.

Remark 4.2. Since the F:*inequality* rule establishes that all numbers smaller than a given one proved to be feasible are feasible, and since the sequent $\rightarrow F(p(2^n, ss0))$ (for all n) is provable using only one occurrence of the F:*times* rule, we derive that all natural numbers can be proved to be feasible with essentially only one use of the F:*times* rule. Moreover, notice that a similar example could have been built by using either F:*successor* or F:*plus* instead of F:*times*.

As remarked at the beginning of this section, this example is useful for 'seeing' geometry and dynamics that can occur in proofs. In this case we have cycles that go from each of our building blocks $(\forall x)F(x) \supset F(x^{2^{2^j}}) \rightarrow (\forall x)F(x) \supset F(x^{2^{2^{j+1}}})$ back to the proof of $\rightarrow (\forall x)F(x) \supset F(x^2)$ which reflect the fact that the actual multiplication takes place in $\rightarrow (\forall x)F(x) \supset F(x^2)$ (which contains the only use of the F:*times* rule). The logical flow graph arranges these cycles in a kind of linear series. The picture is more complicated in the next example, for which there is also much more compression.

4.2. **A second example.** In this section we show that Parikh's lower bound cannot be essentially improved. The example we present points out the role of nested quantification in proving (with a small number of steps) that large numbers are feasible. We use a construction due to Solovay.

The idea goes as follows. Let the multi-exponential function be denoted by $e_z(0, x) = x$, $e_z(n + 1, x) = z^{e_z(n,x)}$ (for some natural number $z > 1$). The symbol $e_z(1, x)$ will also be denoted by z^x. To prove $\rightarrow F(e_2(n, 2))$, we define the predicates $F_0, F_1, \ldots, F_n$ as follows: $F_0(x)$ as $F(x)$ and $F_i(x)$ as $(\forall z)F_{i-1}(z) \supset F_{i-1}(z^x)$, for all $i = 1, \ldots, n$. We can think of the F_i's as initial segments of the set of natural numbers, such that $F_i \supset F_{i+1}$ (for $i = 0 \ldots n - 1$) and where F_i is closed under exponentiation with respect to numbers in F_{i+1}. We do not add to the language of PA either the multi-exponential function or the predicates defined above, but we suppose to use their definition in Peano Arithmetic. Therefore, to be precise we do not prove $F(e_2(n, 2))$ but a formula of the form $(\exists y)(\text{“}y = e_2(n, 2)\text{”} \wedge F(y))$, where “$y = e_2(n, 2)$” is defined in Peano Arithmetic. A rigorous formalization of what follows should take care of this fact.

The predicates F_i's satisfy two properties:

- the sequent $F_i(x), F_i(y) \rightarrow F_i(x \cdot y)$ is provable, because by definition of F_i $F_i(x \cdot y)$ is $(\forall z)F_{i-1}(z) \supset F_{i-1}(z^{x \cdot y})$, and $F_{i-1}(z) \supset F_{i-1}(z^x)$, $F_{i-1}(z^x) \supset F_{i-1}((z^x)^y)$ by definition of $F_i(x), F_i(y)$ respectively (where $F_{i-1}((z^x)^y)$ is $F_{i-1}(z^{x \cdot y})$);
- the sequent $\rightarrow F_i(2)$ is provable, as a direct consequence of the first property and the definition of $F_i(2)$.

Using these two properties one can quickly show the sequent $\to F_0(e_2(n,2))$ (i.e. $\to F(e_2(n,2))$) by first showing

$$F_0(2), F_1(2), \ldots, F_n(2) \to F_0(e_2(n,2))$$

and then cut the antecedents of this sequent with the provable sequents $\to F_i(2)$. To simplify the exposition, let us give the formal proof of $\to F_0(e_2(2,2))$ and analyze its logical flow graph. We are interested in seeing the role of cycles in the proof. Let Π' be of the form

$$\dfrac{\dfrac{\Pi''}{F_2(2), F_1(2) \to F_1(2^2)} \quad \dfrac{F_0(2) \supset F_0(2^{2^2}) \to F_0(2) \supset F_0(2^{2^2})}{F_1(2^2) \to F_0(2) \supset F_0(2^{2^2})}\ cut}{F_2(2), F_1(2) \to F_0(2) \supset F_0(2^{2^2})} \quad \dfrac{\dfrac{F_0(2) \to F_0(2) \quad F_0(2^{2^2}) \to F_0(2^{2^2})}{F_0(2), F_0(2) \supset F_0(2^{2^2}) \to F_0(2^{2^2})}}{F_2(2), F_1(2), F_0(2) \to F_0(2^{2^2})}\ cut\ \ cut$$

with Π'' of the form

$$\dfrac{\dfrac{F_1(2) \supset F_1(2^2) \to F_1(2) \supset F_1(2^2)}{F_2(2) \to F_1(2) \supset F_1(2^2)} \quad \dfrac{F_1(2) \to F_1(2) \quad F_1(2^2) \to F_1(2^2)}{F_1(2) \supset F_1(2^2), F_1(2) \to F_1(2^2)}\ cut}{F_2(2), F_1(2) \to F_1(2^2)}\ cut$$

Then, let Π be of the form

$$\dfrac{\dfrac{\Pi_0}{\to F_0(2)} \quad \dfrac{\dfrac{\Pi_1}{\to F_1(2)} \quad \dfrac{\dfrac{\Pi_2}{\to F_2(2)} \quad \dfrac{\Pi'}{F_2(2), F_1(2), F_0(2) \to F_0(2^{2^2})}\ cut}{F_1(2), F_0(2) \to F_0(2^{2^2})}\ cut}{F_0(2) \to F_0(2^{2^2})}\ cut}{\to F_0(2^{2^2})}\ cut$$

By substituting the predicate symbols with the formulas defining them, notice that Π' is a proof of the sequent

$$(\forall x)[((\forall z)F(z)^{*1} \supset F(z^x)^{*2}) \supset ((\forall z)F(z) \supset F(z^{x^2}))], (\forall z)F(z)^{*1} \supset F(z^2)^{*2}, F(2) \to F(2^{2^2})$$

where the occurrences $F(z)$ marked by $*1$ are linked by a bridge (say f_3), as well as the occurrences $F(z^x)$ and $F(z^2)$ marked by $*2$ (call this bridge f_4). (To see this, go back to Π''.) The following proofs of $\to F_2(2)$ and $\to F_1(2)$ have bridges (say f_2 and f_1^1, f_1^2, respectively) between the formulas marked by $*1$ and $*2$ below:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{F(a) \to F(a) \quad F(a^x) \to F(a^x)}{F(a) \supset F(a^x), F(a) \to F(a^x)} \quad \dfrac{F(a^x) \to F(a^x) \quad F(a^{x^2}) \to F(a^{x^2})}{F(a^x) \supset F(a^{x^2}), F(a^x) \to F(a^{x^2})}}{(\forall z)F(z) \supset F(z^x), F(a) \to F(a^x)} \quad \dfrac{}{(\forall z)F(z) \supset F(z^x), F(a^x) \to F(a^{x^2})}\ Cut}{(\forall z)F(z) \supset F(z^x), (\forall z)F(z) \supset F(z^x), F(a) \to F(a^{x^2})}\ Contraction}{(\forall z)F(z) \supset F(z^x), F(a) \to F(a^{x^2})}}{(\forall z)F(z) \supset F(z^x) \to F(a) \supset F(a^{x^2})}}{(\forall z)F(z) \supset F(z^x) \to (\forall z)F(z) \supset F(z^{x^2})}}{\to ((\forall z)F(z) \supset F(z^x)) \supset ((\forall z)F(z) \supset F(z^{x^2}))}}{\to (\forall x)[((\forall z)F(z)^{*1} \supset F(z^x)^{*2}) \supset ((\forall z)F(z) \supset F(z^{x^2}))]}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{F(b) \to F(b) \quad F(b) \to F(b)}{F(b), F(b) \to F(b^2)}\ \mathrm{F}:times}{F(b) \to F(b^2)}}{\to F(b) \supset F(b^2)}}{\to (\forall z)F(z)^{*1} \supset F(z^2)^{*2}}$$

The compositions of the bridges $f_2 f_3 f_1^1 f_4$ and $f_2 f_3 f_1^2 f_4$ form two cycles in Π. These are the only cycles contained in the proof.

With some easy computation[5], one can show that there exists a proof of $\to F(e_2(n,2))$ with formulas having at most n nested quantifiers, containing 2^{n-1} cycles (i.e. one cycle for each quantifier occurring in the end-sequent; notice that we have 2^{n-1} such quantifiers). (The cases $n = 0, 1$ are slightly different from the general case. The proofs associated to $F_0(2)$ and $F_0(2^2)$ do not contain cycles.)

The preceding proof of $\to F_2(2)$ reflects the general pattern. For $i \geq 1$ one proves $\to F_{i+1}(2)$ in essentially the same manner. To explain the structure of bridges in this proof it is helpful to write $\to F_{i+1}(2)$ out as

$$\to \forall x[(\forall z)(F_{i-1}(z) \supset F_{i-1}(z^x)) \supset (\forall z)(F_{i-1}(z) \supset F_{i-1}(z^{x^2}))].$$

We have two types of bridges coming from our proof. The first type consists of bridges across the main occurrence of $\supset$. We call them *top-level* bridges. For this we have bridges from all of the atomic formulas in the $F_{i-1}(z)$ on the right to their counterparts in the $F_{i-1}(z)$ on the left, and bridges from all the atomic formulas in the occurrence of $F_{i-1}(z^x)$ on the left to their counterparts in $F_{i-1}(z^{x^2})$. Since F_j contains a total of 2^j atomic subformulas, we have $2 \cdot 2^{i-1} = 2^i$ bridges of this first type. The second type of bridges cross the connective $\supset$ inside $(\forall z)(F_{i-1}(z) \supset F_{i-1}(z^x))$ on the left. We call them *down-level* bridges. These bridges connect all atomic formulas inside $F_{i-1}(z^x)$ to their counterparts in $F_{i-1}(z)$, and there are 2^{i-1} of these bridges.

Note that the proofs of $\to F_i(2)$ are slightly different when $i = 1, 0$. In the proof of $\to F_1(2)$ (given above) one has the analogues of the top-level bridges, but there are no bottom-level bridges. The proof of $\to F_0(2)$ has no bridges at all.

In the proof as a whole the 2^{i-1} bottom-level bridges in $\to F_{i+1}(2)$ are combined with the 2^{i-1} top-level bridges in $\to F_i(2)$ to make 2^{i-1} cycles when $i \geq 1$. These bridges are connected through the rest of the proof. To understand this it is helpful to think about the other kind of building block used in the proof, namely proofs of $F_i(w), F_{i-1}(2) \to F_{i-1}(2^w)$. This sequent is very easy to prove using the definitions and some logical steps (as for Π''). If we write out $F_i(w)$ as $(\forall u)(F_{i-1}(u) \supset F_{i-1}(u^w))$, then all of the atomic formulas in $F_{i-1}(u^w)$ inside $F_i(w)$ are connected by bridges to their counterparts in $F_{i-1}(2^w)$, and all of the atomic formulas in $F_{i-1}(2)$ are connected to their counterparts in $F_{i-1}(u)$ inside $F_i(w)$. These building blocks are combined using cuts to get $F_0(2), F_1(2), \ldots, F_n(2) \to F_0(e_2(n,2))$, and the whole proof is obtained by combining this with $\to F_i(2)$, $i = 0, 1, \ldots, n$ using cuts. One can check that the bridges match up to give cycles as described above.

To make the nesting of cycles in the global proof more clear, it is helpful to use a different but equivalent organization of the proof. Think of proving $\to F_i(e_2(n-i,2))$ successively for $i = n, n-1, \ldots, 0$. We start with $\to F_n(2)$ as the $i = n$ case. Suppose now that $\to F_i(e_2(n-i,2))$ has been proved for some i, and we want to reduce this to $i-1$. The proof actually gives a structure for the bridges of $\to F_i(e_2(n-i,2))$ which is just like the structure of bridges for $\to F_i(2)$ discussed above. This is automatic when $i = n$, and in general it comes from

[5]Notice that the subproofs of $\Pi : \to F_0(e_2(n,2))$ with end-sequents $\to F_k(2)$ (for $n > k > 1$) and $\to F_1(2)$, contain respectively 2^{k-1} and 2 bridges. Such bridges contribute to form distinct cycles in Π. Since $2^{n-2} + 2^{n-3} + \ldots + 2^2 + 2 + 1 + 1 = 2^{n-1} - 1 + 1 = 2^{n-1}$, the claim follows.

the induction. Thus we suppose that we have a proof of $\rightarrow F_i(e_2(n-i,2))$ with the same structure of bridges as for $\rightarrow F_i(2)$, and we want to produce a proof of $\rightarrow F_{i-1}(e_2(n-i+1,2))$ with the same structure of bridges as for $\rightarrow F_{i-1}(2)$.

To prove $\rightarrow F_{i-1}(e_2(n-i+1,2))$ we first combine $\rightarrow F_i(e_2(n-i,2))$ with $F_i(e_2(n-i,2)), F_{i-1}(2) \rightarrow F_{i-1}(2^{e_2(n-i,2)})$ using a cut. In saying that $\rightarrow F_i(e_2(n-i,2))$ has the same kind of bridges as $\rightarrow F_i(2)$ we have that there are two levels of bridges, one level that goes across the main occurrence of $\supset$ in $\rightarrow F_i(e_2(n-i,2))$, the other level going across the left occurrence of $\supset$ within $\rightarrow F_i(e_2(n-i,2))$. (This should be modified slightly when $i=1$.) When we use the cut rule to get $F_{i-1}(2) \rightarrow F_{i-1}(2^{e_2(n-i,2)})$ the top level of bridges in $\rightarrow F_i(e_2(n-i,2))$ leads to bridges which connect all atomic formulas in $F_{i-1}(2)$ to all atomic formulas in $F_{i-1}(2^{e_2(n-i,2)})$. Therefore, when we combine this with $\rightarrow F_{i-1}(2)$ we get a proof of $\rightarrow F_{i-1}(2^{e_2(n-i,2)})$ with the same structure of bridges as in $\rightarrow F_{i-1}(2)$. The bottom level of bridges in $\rightarrow F_i(e_2(n-i,2))$ (for $i \geq 2$) leads to extra bridges in $F_{i-1}(2) \rightarrow F_{i-1}(2^{e_2(n-i,2)})$ between the atomic formulas inside the two halves of $F_{i-1}(2)$. When we combine this with $\rightarrow F_{i-1}(2)$ using the cut, these bridges combine with the top-level bridges from $\rightarrow F_{i-1}(2)$ to make cycles, as indicated above.

Thus we get a proof of $\rightarrow F_{i-1}(e_2(n-i+1,2))$ with the right kind of bridges, and so the argument can be repeated. Notice that the number of bridges slowly disappears, until we get to the proof of $\rightarrow F_0(e_2(n,2))$, which has no bridges, only a single atomic formula.

Remark 4.3. The number of steps in the proof of $\rightarrow F(e_2(n,2))$ is $\mathcal{O}(2^n)$. In fact, notice that the number of steps in a proof of $\rightarrow F_k(2)$ is $\mathcal{O}(2^k)$ (in particular, the sequent $\rightarrow F_k(2)$ has an exponential number of atomic sub-formulas $\mathcal{O}(2^k)$) and of $F_0(2), F_1(2), \ldots, F_n(2) \rightarrow F_0(e_2(n,2))$ is $\mathcal{O}(n)$. By using logical axioms which are defined on formulas of arbitrary logical complexity other than atomic, we would be able to obtain proofs of $\rightarrow F_k(2)$ of $\mathcal{O}(k)$ steps and therefore a proof of $\rightarrow F(e_2(n,2))$ of $\mathcal{O}(n)$ lines.

Remark 4.4. The construction proposed above explains why the bound found by Parikh is sensitive to the number of quantified F-formulas *and* to the number of nesting of quantifiers in them. From the construction, it follows that such bound *cannot* be essentially improved (for Peano Arithmetic) by reducing it to a fixed iterated exponential function f in the number of quantified F-formulas, the number of nesting of quantifiers in them and the number of F-formulas applied in the proof. In fact, the proof of $\rightarrow F(e_2(n,2))$ contains $\mathcal{O}(2^n)$ quantified F-formulas, at most n nesting of quantifiers in them and only one F:*times* application. In the construction we do not require the exponential function to be total, therefore the bound holds for all theories in which such a function can be defined. As for the example discussed in section 4.1, a similar construction would have lead to the same results in case F:*successor* or F:*plus* rules were used instead of F: *times*.

Remark 4.5. As mentioned already, the construction described above leads precisely to a proof of a formula $(\exists y)(\text{“}y = e_2(n,2)\text{”} \wedge F(y))$, where “$y = e_2(n,2)$” is defined in Peano Arithmetic. In Peano Arithmetic, one can show the equality “$e_2(n,2) = \delta$”, for some term δ expressed in the language of arithmetic. Hence, by using the F:*equality* rule one can derive $F(\delta)$ from $(\exists y)(\text{“}y = e_2(n,2)\text{”} \wedge F(y))$ and “$e_2(n,2) = \delta$” in a constant number of steps. The number of steps involving the F-formulas remains essentially the same as for the proof of $(\exists y)(\text{“}y = e_2(n,2)\text{”} \wedge F(y))$.

Remark 4.6. Parikh's bound depends only on the number of steps in proofs involving F-formulas. Dragalin's bound takes into account the number of steps formalized in Peano Arithmetic as well. It remains open if such a bound is in fact optimal or not.

Intuitively the reader might already be convinced that cycles are necessary for the compression of proofs of feasibility of very large numbers. We will formally prove that short proofs *need* to have cycles in the next section (Theorem 5.1).

To conclude, let us say that this example is again particularly useful for seeing nontrivial dynamics and geometry in proofs in a concrete way. The structure is much more complicated here, and this reflects the nesting of a large number of quantifiers. The intricate structure indicates how short proofs can reflect 'internal symmetry' in their graph (this is investigated in [10]).

5. Shortening needs cycles

We show here that the logical flow graph of short proofs of large numbers has to contain cycles. Let T_F^- be defined as T_F by dropping the F:*elimination* rule.

Theorem 5.1. *Let* $\Pi : \rightarrow F(m)$ *be a proof in* T_F^- *without cycles. Then there is a proof of at most* $\mathcal{O}(N(\Pi)^2)$ *lines in* T*, of* $\rightarrow m < t \vee m = t$ *where* t *is a term in the language of arithmetic* $0, s, +, *$ *which contains at most* $2^{\mathcal{O}(N(\Pi))}$ *symbols. In particular, there is a constant* $c > 0$ *so that any proof of* $\rightarrow F(m)$ *in* T_F (T_F^*) *with* $\leq c \cdot \log \log m$ *lines must contain a cycle.*

First we need to show a few lemmas on the structure of proofs in the theory T_F^- that will be used all along the paper. We say that a *weak occurrence* of a formula in a proof Π is a formula A whose direct paths all go to weak formulas A' which are variants of A occurring in axioms of Π. We can always add new weak formulas to an end-sequent without augmenting the complexity of the proof. In other words, the weakening rule can be simulated by the system.

Lemma 5.2 (Addition of weak occurrences)**.** *Let* $\Pi : \Gamma \rightarrow \Delta$ *be a proof and* Λ, Θ *be multi-sets. A proof* $\Pi' : \Gamma, \Lambda \rightarrow \Delta, \Theta$ *can be constructed such that* $N(\Pi') = N(\Pi)$.

Proof. By induction on the height of Π. If Π is an axiom, then Π' is an axiom as well. If Π ends with a rule of inference R, then apply the induction hypothesis to one of the immediate subproofs of Π whenever R is binary, or to the only one immediate subproof whenever R is unary; apply again the rule R to the result. (Special rules behave the same way as ordinary logical rules.) Notice that we may have to rename variables to handle the cases $\exists$:*left* and $\forall$: *right* (but this can be done as described in [30]). Since the structure of Π' is essentially the same as the structure of Π (the only difference lies in the presence of weak occurrences Λ, Θ in Π'), we have that $N(\Pi') = N(\Pi)$. □

Given some weak occurrence in the end-sequent, we can always eliminate it. This is the content of the next lemma.

Lemma 5.3 (Elimination of weak occurrences)**.** *Let* $\Pi : S$ *be a proof and* A *a weak occurrence of a formula in* S*. Let* S' *be the result of omitting* A *from* S*. Then a proof* $\Pi' : S'$ *can be constructed such that* $N(\Pi') = N(\Pi)$.

Proof. By induction on the height of Π. If S is an axiom, then S' is an axiom as well (since A is a non-distinguished occurrence). The induction step is proved deleting the weak occurrence A from a premise of the last rule of inference R of Π and applying R again to the resulting proof(s). If R is a logical rule, then A cannot be its main formula and we can simply apply the induction hypothesis to the immediate subproofs of Π where A occurs. In case R is a contraction rule, then we can apply the induction hypothesis twice to the auxiliary formulas and derive the claim. If R is a special rule, we apply the induction hypothesis to its immediate subproofs. It follows that the tree-like structure of Π' is the same as the tree-like structure of Π, then $N(\Pi') = N(\Pi)$. □

Lemma 5.4. *Let $\Pi : S$ be proof of k lines. Then there is a proof $\Pi' : S$ of at most k lines such that it never happens that an auxiliary formula of a binary rule is weak, or that all of the auxiliary formulas of a unary logical rule are weak, or that an auxiliary formula of a contraction rule is weak. Furthermore, if Π is cut-free then Π' is cut-free.*

Proof. This is proved by induction on the height of the proof Π. Suppose that the last rule of inference of Π is a binary rule where at least one of the auxiliary formulas is weak. Suppose a $\wedge$:*right* rule (similarly for $\vee$:*left*, cut, F:*equality*, F:*inequality*, F:*plus* and F:*times* rules) is applied to the subproofs $\Pi_1 : \Gamma_1 \to \Delta_1, A$ and $\Pi_2 : \Gamma_2 \to \Delta_2, B$, where A is a weak occurrence in Π_1. By eliminating A from Π_1 (using Lemma 5.3) and adding (using Lemma 5.2) to the resulting proof weak occurrences $\Gamma_2, \Delta_2, A \wedge B$, we obtain a proof with the same number of lines as Π_1 of the sequent $\Gamma_{1,2} \to \Delta_{1,2}, A \wedge B$. Let Π' be such a proof.

Now suppose that the last rule of inference of Π is a unary logical rule applied to two weak auxiliary formulas. Suppose that a $\wedge$:*left* rule (similarly for $\vee$:*right*) is applied to the subproof $\Pi_1 : A, B, \Gamma_1 \to \Delta_1$, where A, B are weak occurrences in Π_1. We first eliminate A, B from Π_1 (using Lemma 5.3) and then we add the weak formula $A \wedge B$ (using Lemma 5.2). The proof we obtain has the same number of lines as Π_1. Let it be Π'.

If the last rule of inference is unary and has only one auxiliary formula which is weak, the treatment is similar. This is the case for $\neg$:*right*, $\neg$:*left*, $\forall$:*right*, $\forall$:*left*, $\exists$:*right*, $\exists$:*left* and F:*successor* rules. We proceed in a similar manner when the rule is a contraction. □

We are now ready to prove Theorem 5.1.

Proof. (Theorem 5.1). Let $\Pi : \to F(m)$ be a proof in T_F^- without cycles. We will show that we can build a term t in the language of arithmetic satisfying the conditions.

Without loss of generality, we begin by applying Lemma 5.4 to Π. (This is not crucial for the construction itself but it will be in counting the complexity of the terms at the end.) We obtain a proof which will still have no cyclic F-paths. Let us call this proof Π again. We want to consider F-paths and we want to substitute all occurrences $F(t)$ (where t is an arbitrary term) in Π with formulas of the form $t = u \vee t < u$ where u is some *closed* term. We will perform these substitutions by induction on the height of the subproofs of Π. During the construction we will use some auxiliary variables $x_1, x_2, \ldots$ and the symbol $\leq$ which do not belong to the current language. The symbol $\leq$ is used to simplify the exposition but the auxiliary variables play a more important role. In fact, we cannot a priori perform

substitutions in special logical axioms of the form $F(t), \Gamma \to \Delta, F(t)$ until we know which formula to substitute. The auxiliary variables will be used to delay the substitution until all constraints imposed by the structure of the proof are known. To simplify the construction we also extend our proof system by adding the axiom

$$\Gamma \to \Delta, 0 \leq 0$$

and the following set of new rules of inference to T_F^*:

$$\leq\text{: } \textit{equality} \qquad \frac{\Gamma_1 \to \Delta_1, u = t \quad \Gamma_2 \to \Delta_2, t \leq s}{\Gamma_{1,2} \to \Delta_{1,2}, u \leq s},$$

$$\leq\text{: } \textit{inequality} \qquad \frac{\Gamma_1 \to \Delta_1, u \leq t \quad \Gamma_2 \to \Delta_2, t \leq s}{\Gamma_{1,2} \to \Delta_{1,2}, u \leq s},$$

$$\leq\text{: } \textit{successor} \qquad \frac{\Gamma \to \Delta, u \leq t}{\Gamma \to \Delta, s(u) \leq s(t)},$$

$$\leq\text{: } \textit{plus} \qquad \frac{\Gamma_1 \to \Delta_1, u \leq t \quad \Gamma_2 \to \Delta_2, s \leq r}{\Gamma_{1,2} \to \Delta_{1,2}, u + s \leq t + r},$$

$$\leq\text{: } \textit{times} \qquad \frac{\Gamma_1 \to \Delta_1, u \leq t \quad \Gamma_2 \to \Delta_2, s \leq r}{\Gamma_{1,2} \to \Delta_{1,2}, u * s \leq t * r},$$

$$\leq\text{: } \textit{weakening} \qquad \frac{\Gamma \to \Delta, u \leq t}{\Gamma \to \Delta, u \leq t + r} \qquad \frac{\Gamma \to \Delta, u \leq t}{\Gamma \to \Delta, u \leq r + t.}$$

The axiom and the rules are consistent with Peano Arithmetic. The axiom and the first five rules are the obvious counterparts of the special axiom and special rules of T_F^-. The last rule will be used to handle contractions on F-formulas. Notice that in the $\leq$:*weakening* rule r is any arbitrary term.

Our argument will proceed in two stages. In the first stage we introduce the auxiliary variables x_i. Each atomic F-formula $F(t)$ will be substituted with $t \leq u$ where u is either a closed term or a term whose only variables are the auxiliary variables x_i's. We will perform the substitutions by induction on the height of the subproofs of Π'. At each step we will consider a certain subproof Π_*, we will apply the induction hypothesis to its immediate subproofs and recombine them again either with the same rule of inference used as the last rule of Π_* or with one of the rules for $\leq$ introduced above. If the last rule of inference in Π_* is either a contraction on F-formulas or a cut rule, the treatment will be a bit more complicated. Some constraints are imposed by these rules on the substitutions, thus at the end of the induction we only obtain a *proof structure* instead of a *proof*. These constraints will be resolved in the second part of our argument with the special treatment of the cut rule. The argument there will entail global considerations, and the absence of cycles will be crucial. During the construction there will be no conflict with quantified variables. In fact for all inequalities $t \leq u$ substituted for some $F(t)$, the term u will contain only variables x_i's which will not be quantified. An important point is that we do not create new cycles in the first part of the construction.

Let us now begin to describe the induction argument. First suppose that Π_* is a special axiom of the form $\Gamma \to \Delta, F(0)$. We replace it with the sequent $\Gamma^* \to \Delta^*, 0 \leq 0$, where Γ^*, Δ^* are obtained from Γ, Δ by replacing weak occurrences $F(t)$ with $t \leq x_i$, using a different variable x_i for each occurrence. We also ask the auxiliary variables to be new for the construction.

We will substitute special logical axioms of the form $F(t), \Gamma \to \Delta, F(t)$ with axioms $t \leq x_i, \Gamma^* \to \Delta^*, t \leq x_i$, where the symbol x_i has not yet been used in the construction, and Γ^*, Δ^* are defined as above. For all other axioms in Π_* we substitute the side formulas Γ, Δ with Γ^*, Δ^* as defined above (their distinguished occurrences are F-free).

If the last rule of inference R of our subproof Π_* is a logical rule, then we apply to the immediate subproof(s) the induction hypothesis and the rule R again. If R is one of the special rules F:*equality*, F:*inequality*, F:*successor*, F:*plus*, F:*times*, then we will apply the corresponding $\leq$:*equality*, $\leq$:*inequality*, $\leq$:*successor*, $\leq$:*plus*, $\leq$:*times* rules instead.

Before we continue the inductive construction let us make some comments about the way that the auxiliary variables x_i are used. When an auxiliary variable x_i appears *negatively* in our proof, it appears in an atomic formula of the form $t \leq x_i$, where t is a term in our original language. In *positive* occurrences we can have $t \leq u$, where t is a term in our original language and u is a term which can involve constants and auxiliary variables, but not variables from the original language. It is easy to see that these restrictions on the way that auxiliary variables appear are consistent with the part of the construction described above, and they will be consistent also with the remainder of our inductive procedure.

Continuing the argument we would like to say that there is at most *one* inequality which contains x_i and occurs negatively in each sequent in the proof structure that we construct. That is, compatible with the part of the construction that we have discussed already. For the argument relative to contractions which we will discuss next, this will not quite work. It is almost true however. We shall see that when we modify a particular proof, each x_i can be contained in exactly one inequality occurring negatively in the end-sequent and that any two inequalities containing x_i and occurring negatively in a sequent of the new proof structure will eventually go (by direct paths) into the same occurrence in a contraction formula.

Let us now continue our argument and suppose the last rule of inference in Π_* to be a *contraction* rule. In case the contraction formulas are F-free we apply the induction hypothesis to the immediate subproof of Π_* and the contraction rule again.

If the contraction formulas are atomic formulas $F(t)$, then we apply the induction hypothesis and obtain a pair of formulas $t \leq u$ and $t \leq s$ which correspond to the pair of auxiliary contraction formulas $F(t)$. Here we need to make a case distinction. The occurrences $F(t)$ (respectively, the inequalities substituted for them) might appear positively or negatively in the sequent and we will treat these cases in different ways. If $F(t)$ appears in the right-hand side of the sequent arrow (i.e. positively), then we apply $\leq$:*weakening* rules to both inequalities to obtain a pair of inequalities $t \leq u + s$ and afterwards we contract them. If $F(t)$ appears in the left-hand side of the sequent arrow (i.e. negatively), notice that by construction both $F(t)$'s should have been replaced by inequalities $t \leq x_i$ and $t \leq x_j$ for some $i \neq j$. We replace all x_j's in the proof with x_i and we contract the inequalities afterwards.

The case where the contraction formulas are non-atomic F-formulas needs more attention. Each pair of corresponding atomic F-formulas occurring in the contraction formulas will be substituted by induction hypothesis with a pair of corresponding inequalities, say $t \leq u$ and $t \leq s$. If they appear positively in the sequent, then we should like to have them be an inequality of the form $t \leq u + s$ instead. To

do this we look at the direct paths going up in the proof from these occurrences of $t \leq u$ and $t \leq s$ towards the axioms. If we start with an occurrence of $t \leq u$ (or $t \leq s$), then we will reach occurrences of the form $t' \leq u$ ($t' \leq s$) until we possibly go through rules for $\leq$, with t' differing from t only through substitutions that result from the use of quantifier rules. Given a path, consider the first moment that it reaches an atomic formula in the proof. If there is no such moment, then it means that the path goes all the way up to a non-atomic formula in an axiom, and this non-atomic formula cannot be the distinguished formula in the axiom. (In the original proof, a distinguished F-formula in an axiom must always be atomic.) Then we simply replace $t' \leq u$ ($t' \leq s$) with $t' \leq u + s$ in the non-distinguished formula in the axiom, and also in its successors in later stages of the proof. If we reach an atomic formula, and if that formula either came from the main formula of a rule (either a contraction rule or a $\leq$-rule) or from a distinguished occurrence in an axiom, then we apply the $\leq$:*weakening* rule to it to obtain $t' \leq u + s$, and afterwards follow the earlier proof. If that first atomic formula did not come from a rule or a distinguished occurrence of an axiom, then it comes from a single weak occurrence in an axiom, and we can make a substitution there as before. After this transformation we obtain two formulas which are the same and we apply a contraction to them. Notice that these changes cannot interact with the quantifier rules in a dangerous way, because u and s involve only the new auxiliary variables and none of the original variables. If we started with negative occurrences, of the form $t \leq x_i$ and $t \leq x_j$, say, with $i \neq j$, then we replace all occurrences of x_i with x_j as in the preceding paragraph, and then we perform the contraction.

Notice that, given any pair of contraction formulas, the number of inequalities to which we apply $\leq$:*weakening* rules is bounded by the number of lines in Π_*. Indeed, when the contraction formula is atomic we applied $\leq$:*weakening* rules twice. The non-atomic case is also not hard to check, because we applied $\leq$:*weakening* rules only to formulas that came from the main formula of a rule, or came straight from a distinguished positive occurrences in an axiom. Moreover, notice that our treatment of the contraction rules is compatible with the restrictions on the negative occurrences of the auxiliary variables described above.

If the last rule of inference R is a *cut* rule, we apply the induction hypothesis to the immediate subproofs and again the cut rule. If the cut-formulas are F-formulas then our construction will lead to a pair of new formulas with the same logical structure but with sub-formulas $t \leq u$ and $t \leq s$ which are supposed to correspond to each other but which may have $u \neq s$. If this is the case then we do not get a *proof* (since the cut-formulas ought to be identical) but only a *proof-structure*, i.e., a labeled tree of sequents in which a proper substitution of variables may lead to a proof.

This concludes the first stage of our construction. We have obtained a proof structure and we would like to replace all occurrences of variables x_i with closed terms to get a proof.

Before we continue let us make some observations about the geometry of this proof structure. We can define a logical flow graph for it in the same way as usual. The graph that we get is practically the same as for the original proof. It *is* the same for the parts that do not involve F-formulas. For the parts that involve F-formulas we have, in effect, only added vertices to some of the old edges. These new vertices correspond to the addition of $\leq$:*weakening* rules. Except for these additions we did not change the logical structure of the proof or its associated graph. We conclude

that the part of the logical flow graph involving the inequalities (i.e. the part that corresponds to the old F-formulas) has no cycles either. In the end this lack of cycles will permit us to make substitutions to reconcile the cut-formulas and get a proof.

Next, we need to notice that all inequalities which appear negatively and contain x_i (for some i) will have the form $t \leq x_i$ for some t and will have a direct path starting from some cut-formula. There is only *one* such cut-formula and there is exactly *one* inequality containing x_i occurring in it. This is not hard to check, from the construction. It uses the fact that the auxiliary variables were all chosen to be different initially, and they were relabeled to be the same only in contractions, and then only in the case of *negative* occurrences in contraction formulas. In other words, if two negative occurrences were relabeled to use the same auxiliary variable x_i, then they were connected below in a contraction.

We should also observe that an inequality containing x_i and occurring negatively in some cut-formula will have direct paths which either end into non-distinguished formulas or will turn over some special logical axioms and will go down to some cut-formula, by construction. Some of these paths might be reconnected through binary rules for $\leq$, but we do not mind. Some others might go through contractions, and in the first stage of our construction we might have introduced new occurrences of the variable x_i going up from these contractions. These new occurrences are all positive, and they do not cross any axioms (the 'backtracking' of positive occurrences of x_i from a contraction do not provide any new connections to cut-formulas). This describes completely the subgraph of inequalities in our proof structure which contain x_i. It is a connected subgraph. Of course we are using here the fact that paths of inequalities containing x_i cannot cross a cut, because of our construction with different auxiliary formulas chosen originally. (If an inequality containing x_i occurs only positively in the proof structure, one can again analyze the corresponding part of the logical flow graph, but we shall not need this.)

These special properties of logical paths in our proof structure will permit us to substitute variables x_i to bridge the gaps across the cuts without having conflicts. Here the hypothesis of acyclicity of F-paths plays a crucial role. In fact, we want to substitute closed terms for the auxiliary variables x_i to match cut-formulas properl, and we will have that positive occurrences of inequalities will establish the substitution over negative occurrences. We know that once an inequality occurring negatively is considered for substitution, there is no way it will be considered again. Let us describe with some detail how this substitution of auxiliary variables is achieved.

Let us call an auxiliary variable x_i *critical* if it has a negative occurrence in the proof structure that we have constructed, and *relaxed* if it has only positive occurrences. We shall first substitute critical variables with terms which contain only relaxed variables, and then we shall substitute closed terms for relaxed variables.

Take two corresponding inequalities in a pair of cut-formulas in the proof structure such that the negative occurrence contains a critical variable x_i. As above, the negative occurrence will be of the form $t \leq x_i$ and it is the unique negative occurrence of x_i in a cut-formula. (Our substitutions will not change these facts, since critical variables will be removed but not added.) Suppose that the corresponding inequality which occurs positively is of the form $t \leq u$, where the term u contains only relaxed variables if any. Substitute all occurrences of x_i with u. Do this repeatedly until all critical variables have been replaced.

It is precisely here that we need our assumption concerning the absence of cycles. This ensures that at each stage there is a critical variable in a cut-formula whose counterpart u in the dual formula contains only relaxed variables. This observation uses our earlier discussion of critical variables, which ensures that any time a critical variable occurs positively in a cut-formula there was an oriented path to that occurrence from a negative occurrence in another cut-formula. If there were ever a time in which all the u's contained critical variables, one could find a cycle by stringing together a sequence of such paths. The argument is easier to imagine than to read, and we omit the details.

Thus we can get rid of all the critical variables by substitution. Once we do this our proof structure becomes a proof, because the cut-formulas match up. We then replace the relaxed variables with arbitrary closed terms, like 0.

This concludes our construction. To build a proof Π' in T we only need to substitute all occurrences of formulas $t \leq u$ with formulas of the form $t = u \vee t < u$. Both axioms and rules for $\leq$ can be substituted with proper subproofs of T. It is easy to check that all such subproofs have a constant number of steps. Moreover, for each contraction rule applied to F-formulas we needed at most $N(\Pi')$ applications of $\leq$:*weakening*. Since there are at most $N(\Pi')$ such contractions, the bound on the complexity of Π' is $\mathcal{O}(N(\Pi)^2)$.

To conclude the proof of our statement, we only need to argue that the number of symbols in t is at most $2^{\mathcal{O}(N(\Pi))}$. Let us recall here the main points behind the proof of the bounds on t. At the base of the construction of t from Π are the special axioms and positive weak occurrences. With the special axioms one starts with the term 0, while positive weak occurrences may be associated to more complicated terms, because of substitutions that we made in the proof while treating contractions. All symbols which appear in t are introduced through the treatment of F-rules, positive contractions (i.e. those contractions affecting F-atomic formulas occurring positively), special axioms and positive weak occurrences, and so are controlled by the number of lines of Π. (See the next paragraph below concerning positive weak occurrences.) Cuts and negative contractions (i.e. those contractions affecting F-atomic formulas occurring negatively) lead to substitutions and duplication of symbols that we should account for. However, one can check that the total number of duplications is controlled by the number of auxiliary variables that were introduced. In fact, only auxiliary variables which have both negative and positive occurrences matter for this part (the others do not contribute), and so the total number is bounded by the number of axioms in Π. Duplications lead to exponential growth which is bounded as claimed above, because the proof has no cyclic F-paths.

These are the most important points. Let us also mention that the application of Lemma 5.4 at the beginning of the construction plays a crucial role in the counting of the positive weak occurrences. Because of Lemma 5.4 we know that positive weak formulas in Π are never used as auxiliary formulas for F-rules. As a practical matter this ensures that they do not really participate in the construction of t in a relevant way. Indeed, we had to make some 'backtracking' substitutions from positive contractions in order to make sure that we had a proof, but the part of the argument that affected the auxiliary terms was much simpler. It is not too hard to check that suitable estimates are obtained. (This part is relatively minor compared to the exponential effects of the substitutions through cuts and negative contractions.) This shows the first part of our statement.

Notice now that all terms representing m or values larger than m in the language of arithmetic should have at least $b \cdot \log\, m$ symbols for some constant $b > 0$. This is not hard to check, since multiplication is the most efficient operation for making large numbers. Using this one can derive easily the second part of the theorem from the first. In fact, the term t is supposed to represent m or larger values, and so it should have at least $b \cdot \log m$ symbols. If there are no cycles, then it has at most $2^{\mathcal{O}(N(\Pi))}$ symbols. If $N(\Pi)$ is small compared to $\log \log\, m$, then $2^{\mathcal{O}(N(\Pi))}$ will be bounded by a small power of $\log\, m$. This will be smaller than $b \log m$ when m is large enough, as soon as the small power is $1/2$, say. On the other hand, if m is not too large and we choose c small enough, then $N(\Pi)$ would have to be < 1, which is impossible. □

The construction in Theorem 5.1 can be only carried out under the hypothesis that there is no cycling of F-paths in Π. In fact cycles code in the proof a potentially *infinite* process of substitution and this would not allow a replacement of the auxiliary variables. One should notice that the procedure of cut elimination applied to F-formulas would eliminate cycles (as a consequence of the elimination of F-formulas from the proof) at the cost of an expansion which is finite but in the worse case is non-elementary in the number of lines of the original proof. This 'finiteness', which is coded in the structure of the proof, depends on the fact that a proof describes both a 'potentially infinite' process and the 'input' for this process.

Remark 5.5. In Section 4 we have given two examples which illustrate how quantified cut-formulas can efficiently codify large numbers. In general, one can ask what is the coding power of the quantifiers in a proof. Can we measure it? The construction used to show Theorem 5.1 furnishes a way to measure the coding power of quantified cut-formulas when cycles are not present. In fact, arithmetical terms can explicitly be associated to each node of an F-path in a proof and their complexity turns out to be elementary bounded on the complexity of the proof. (It is important to notice here that an F-path can pass several times through the quantified cut-formulas in the proof and each node of the F-path crossing the cut-formula might be affected by the action of a different quantifier. This is the case for our examples of feasibility.) In the case of cyclic paths, the complexity of terms that can be built following the construction in Theorem 5.1 is *unbounded*. This is because terms lying along cycles feed themselves back and a finite measuring is therefore impossible. One might be interested to associate an infinite measure to the cycles or to estimate the rate of growth of terms along cyclic paths, but this point goes behind the aim of this paper even though it deserves investigation.

Remark 5.6. In [6] it is shown that cycles can be transformed into spirals by slightly modifying the LK calculus. For this purpose a new sequent calculus for classical logic, ALK ('acyclic LK') has been introduced. It is close to *Linear Logic* [18] in spirit (the applicability of its rules depends on the life of the formulas in the proof, their past and their future), enjoys cut-elimination, the logical graphs of its proofs have no cyclic paths and its proofs are just *elementary* larger than proofs in LK. Precise bounds on the complexity of the transformation between the two systems are given. The proofs in the new calculus can be obtained by a *small perturbation* of proofs in LK. The transformation of a proof Π into an ALK-proof can be seen as the effect of going from a space to its covering. In this view a proof in LK can

be seen as a *projection* of a proof in ALK, where the *identification* of some of the formulas forces the presence of cycles.

Lower bounds on the complexity of the *elimination* of spirals from a proof can be derived from the complexity of the elimination of cycles. Notice that Theorem 5.1 shows that the complexity of the elimination of cycles *is* non-elementary. This is because acyclic proofs of large feasible numbers must be long.

Remark 5.7. A nonstandard theory asserting the *existence* of a non-feasible number can be defined from PA_F by substituting the axiom $\neg F(\theta)$ with the axiom $(\exists x)\neg F(x)$. In the formalism of the sequent calculus, this is analogous to ask that the F:*elimination* rule in T_F be substituted with the following:

$$F : \textit{elimination}^* \qquad \frac{\Gamma \to \Delta, (\forall x)F(x)}{\Gamma \to \Delta}\ .$$

Call T_F^* the new sequent theory (the superscript $*$ indicates that the special rule does not introduce inconsistency into the basic system T). The fact that T^* is a consistent extension of T is easily proved using nonstandard models of PA and interpreting $F(x)$ as 'x is a standard number' [25]. This result also follows as a corollary of the First ϵ-Theorem or the Cut Elimination Theorem applied to F-formulas. Here we want to show the following

Theorem 5.8. *Let $\Pi : S$ be a proof in T_F^*. The sequent S does not contain F and Π does not contain F-paths which are cyclic. Then there is a proof $\Pi' : S$ in T such that $N(\Pi')$ is $\mathcal{O}(N(\Pi)^2)$.*

This result can be proved neither through usual model-theoretic means nor through the Cut Elimination Theorem or the First ϵ-Theorem. The proof is a direct consequence of the construction used to show Theorem 5.1 where the coding power of the quantifiers in an acyclic proof is measured through the expansion of terms lying in its F-paths (see Remark 5.5). Here we make *use* of those expanded terms to transform acyclic proofs in the theory T_F^* into arithmetical proofs. When cycles are present one cannot infer the existence of such terms, and a proof of a general statement (where the acyclicity assumption is dropped) should lie on a different idea.

Proof (Theorem 5.8). The construction developed to show Theorem 5.1 can be extended to the theory T_F^*. We only need to handle, in the first stage of the construction, the presence of F:*elimination** rules occurring in the proof $\Pi :\to F(m)$. (We use here the same notation employed in the proof of Theorem 5.1.) Suppose that the last rule of inference in Π_* is a F: *elimination** rule. Then by induction hypothesis the auxiliary formula will be transformed into a formula $(\forall x)x \leq u$ where u might either be a closed term or a term whose only variables are the x_i's. We will cut this formula with a sequent of the form $(\forall x)x \leq u \to$, which can be derived in two steps from the sequent $s(u) \leq u \to$. In particular, the sequent $su < u \vee su = u \to$ is provable in T in a constant number of steps and this assures that the proof Π' has bounded complexity.

The construction in the proof of Theorem 5.1 did not use any assumption on the form of the end-sequent. In particular, it did not use the fact that an F-formula should appear in the end-sequent at all. Only the acyclicity condition played a role. Therefore we can apply the construction in Theorem 5.1 (extended to T_F^* as indicated above) to our proof Π and derive the claim. □

6. The neck of the bottle

The theory T_F has historical interest for the study of feasible numbers even though it does not fit as well with the idea of a proof of feasibility of $F(t)$ as providing a 'description' of t. If one is interested to 'measure' the *effort* in building large terms, then the theory defined by taking T_F and dropping the F:*inequality* rule seems to be the most appealing context in which this aim might be carried out. In fact, in the theory T_F it is enough to approach a large value from above (for instance with a certain power of a power of 2 which is larger than the value) and by the property of closure from below (expressed by the F:*inequality* rule) it would be possible to derive the feasibility of the large value in one step. It is clear though that in between two large powers of 2 there might be numbers represented by terms which are very difficult to build. (To be precise, it might still be complicated to show that the large term is smaller than the large power of 2. The inequality could be proved in arithmetic using induction for instance, but in this case we would not be able to trace the 'effort' that we want to measure). This choice of dropping the F:*inequality* rule is also supported in [8] where the idea of feasibility is applied to mathematical contexts other than arithmetic for studying the 'construction' of mathematical objects other than natural numbers. One might also think to drop the F:*equality* rule since it might help to obtain tricky representations of large numbers using arithmetical induction, but we shall not discuss this matter here.

Let $\widetilde{T}_F$ be the theory defined as T_F but without F:*inequality* and F:*equality* rules. It seems plausible that given a proof $\Pi : S$ in $\widetilde{T}_F$, where S is a F-free sequent, then either there is a T-proof $\Pi' : S$ of $\mathcal{O}(N(\Pi)^2)$ lines, or there is a $\widetilde{T}_F$-proof $\Pi' : \rightarrow$ of $\mathcal{O}(N(\Pi)^2)$ lines. The result would say in essence that either we find in Π a proof of the contradiction (i.e. of the empty sequent $\rightarrow$) or Π 'contains' a proof of S in the language of arithmetic. (Notice that when the size of the proof Π is small compared to the size of the non-feasible term, both Parikh and Dragalin establish the existence of a proof in T of the F-free sequent S which turns out to be of very large complexity. This is due to the fact that both the Hilbert First ϵ-Theorem and the Cut Elimination Theorem induce in the worse case a non-elementary growth of the complexity of the proof.)

As a consequence we would know that short proofs of large numbers (described in Section 4), cannot influence in some *essential* way the complexity of proofs of theorems in the arithmetical theory $\widetilde{T}_F$. In fact, either inconsistency would be explicitly derived in a proof, or the proof already would know how to derive quickly, in the language of arithmetic, a true statement. The next example points out that for the theory T_F instead, true atomic statements can be obtained in some *implicit* way.

Let Π be of the form

$$\dfrac{\dfrac{\dfrac{\theta < p(2^n, ss0) \rightarrow \theta < p(2^n, ss0) \quad \overset{\Pi'}{\rightarrow F(p(2^n, ss0))}}{\theta < p(2^n, ss0) \rightarrow F(\theta)}}{\theta < p(2^n, ss0) \rightarrow}}{\rightarrow \neg\theta < p(2^n, ss0)}$$

where $val(p(2^n, ss0)) = 2^{2^n}$, $val(\theta) > e_2(e_4(39 \cdot N(\Pi), 39 \cdot N(\Pi)), 1)$ (for some suitable choice of n) and Π' is the proof described in section 4.1.

Notice that with $\mathcal{O}(n)$ lines we proved that the non-feasible number θ is larger than $p(2^n, ss0)$. It is clear that such proof does not contain enough information on which any proof in Peano Arithmetic of the inequality $\neg\theta < p(2^n, ss0)$ can be based (in fact, no property of the structure of θ has been taken into account).

References

1. P. Bernays. Sur le platonisme dans les mathématiques. *L'Enseignement Mathématique*, 34:52–69, 1935.
2. G. Boolos. Don't eliminate cut. *Journal of Philosophical Logic*, 13:373–378, 1984. MR **85i:**03029
3. S. Buss. The undecidability of k-provability. *Annals of Pure and Applied Logic*, 53:72–102, 1991. MR **92g:**03071
4. A. Carbone. On Logical Flow Graphs. Ph.D. Dissertation, Graduate School of The City University of New York, May 1993.
5. A. Carbone. Interpolants, Cut Elimination and Flow Graphs for the Propositional Calculus. In *Annals of Pure and Applied Logic*, 83:249–299, 1997. MR **98i:**03073
6. A. Carbone. Turning cycles into spirals. *Annals of Pure and Applied Logic*, 96:57–73, 1997.
7. A. Carbone. Proofs and Groups with distorted Length Function. Manuscript, 1996.
8. A. Carbone and S. Semmes. Looking from the inside and from the outside. I.H.E.S. preprint M/96/44, 1996 (Bures-sur-Yvette, France).
9. A. Carbone and S. Semmes. Making proofs without Modus Ponens: An introduction to the combinatorics and complexity of cut elimination. In *Bulletin of the American Mathematical Society*, 34:131–159, 1997. MR **98c:**03112
10. A. Carbone and S. Semmes. *Geometry, Symmetry and Implicit Representations of Combinatorial Structures*. Book in preparation, 1997.
11. A.G. Dragalin. Correctness of inconsistent theories with notion of feasibility. In *LNCS 208*, pages 58–79. Springer Verlag, 1985. MR **87e:**03150
12. M.E. Dummett. Wang's paradox. In *Truth and Other Enigmas*, pages 248–268. Harvard University Press, Cambridge, 1978. Also appeared in *Synthèse* 30:301–324, 1975.
13. M.E. Dummett. Wittgenstein's philosophy of mathematics. In *Truth and Other Enigmas*, pages 166–185. Harvard University Press, Cambridge, 1978. Also appeared in *The Philosophical Review*, Cornell University, 1959.
14. A.S. Esenin-Volpin. Le programme ultra-intuitionniste des fondements des mathématiques. In *Infinitistic Methods, Proceedings of the Symposium on the Foundations of Mathematics*, pages 201–223. PWN. Warsaw, 1961. MR **26:**4905
15. A.S. Esenin-Volpin. The ultra-intuitionistic criticism and the antitraditional program for foundations of mathematics. In A. Kino, J. Myhill, and R.E. Vesley, editors, *Intuitionism and Proof Theory*, pages 3–45. North-Holland, 1970. MR **45:**4938
16. R.O. Gandy. Limitation to mathematical knowledge. In D. van Dalen, D. Lascar, and J. Smiley, editors, *Logic Colloquium 80*, pages 129–146. North-Holland Publishing Company, 1982. MR **84d:**00010
17. J.Y. Girard. *Proof Theory and Logical Complexity*, volume 1 of *Studies in Proof Theory, Monographs*. Bibliopolis, 1987. MR **89a:**03113
18. J.Y. Girard. Linear Logic. *Theoretical Computer Science*, 50:1–102, 1987. MR **89m:**03057
19. S.C. Kleene. *Introduction to Metamathematics*, volume 1 of *Bibliotheca Mathematica, Monographs on Pure and Applied Mathematics*. Groningen, Walters-Noordhoff, 1988.
20. G. Kreisel. Axiomatizations of nonstandard analysis that are conservative extensions of formal systems for classical standard analysis. In *Application of Model Theory*, W.A.J. Luxemburg, editor, pages 93–106. Holt, Rinehart and Winston, New York, 1969. MR **39:**50
21. G. Mannoury. *Methodologisches und Philosophisches zur Elementar-Mathematik*. Visser, Haarlem, 1909.
22. E. Nelson. *Predicative Arithmetic*. Princeton University Press, Princeton, New Jersey, 1986. MR **88c:**03061
23. R. Parikh. A conservation result. In W.A.J. Luxemburg, editor, *Application of Model Theory*, pages 107–108. Holt, Rinehart and Winston, New York, 1969. MR **38:**3145
24. V. P. Orevkov. Lower Bounds for increasing complexity of derivations after cut elimination. *Journal of Soviet Mathematics*, 20(4), 1982.

25. R. Parikh. Existence and feasibility in arithmetic. *Journal of Symbolic Logic*, 36:494–508, 1971. MR **46:**3287
26. R. Parikh. The problem of vague predicates. In *Language, Logic and Method*, R.S. Cohen and M.W. Wartofsky, editors, pages 241–261. 1983.
27. H. Poincaré. *The Foundations of Science*. Science Press, 1913.
28. V.Yu. Sazonov. On feasible numbers. In *Proceedings of The Kleene's Conference on Mathematical Logic, Varna*, 1990.
29. R. Statman. Bounds for proof-search and speed-up in predicate calculus. In *Annals of Mathematical Logic*, 15:225–287. 1978. MR **81c:**03049
30. G. Takeuti. *Proof Theory*, volume 81 of *Studies in Logic and the Foundations of Mathematics*. North Holland, 1975. MR **58:**27366a
31. A.S. Troelstra. Remarks on intuitionism and the philosophy of mathematics (revised version). In *Proceedings of the Congresso Internazionale "Nuovi Problemi della Logica e della Filosofia della Scienza"*, Viareggio 8-13 January, 1990.
32. G.S. Tseitin. The complexity of a deduction in the propositional predicate calculus. In *Zap. Nauchn. Sem. Leningrad Otd. Mat. Inst. Akad. Nauk SSSR*, 8:234–259. 1968. MR **43:**6098
33. P. Vopenka. *Mathematics in the Alternative Set Theory*. Teubner, Leipzig, 1979. MR **83f:**03051
34. L. Wittgenstein. In *Remarks on the Foundations of Mathematics*. R. Rhees, G.H. von Wright and G.E.M. Anscombe, editors, The MIT Press, Cambridge, 1978. Revised Version. MR **80f:**00009

Mathématiques/Informatique, Université de Paris XII, 61 Avenue du Général de Gaulle, 94010 Créteil Cedex, France

E-mail address: ale@gauss.math.jussieu.fr, ale@univ-paris12.fr