# Linear fragments of Peano arithmetic

Petr Glivický

perglivicky@gmail.com

JAF 2015, New York
July 7, 2015

# Content

Section 1

# Linear arithmetics

# Linear arithmetic

Recall that:

Presburger arithmetic is the full-induction arithmetic for the language $\langle 0, 1, -, +, \leq \rangle$.

Peano arithmetic is the full-induction arithmetic for the language $\langle 0, 1, -, +, \cdot, \leq \rangle$.

We introduce:

Linear arithmetic $(\mathrm{LA})$ is the full-induction arithmetic for the language $\langle 0, 1, -, +, a\cdot, \leq \rangle$, where $a\cdot$ is a unary function of multiplication by a positive non-standard element.

# Higher order linear arithmetics

Similarly, for any cardinal $\kappa$, we introduce:

$\kappa$-linear arithmetic ($\mathrm{LA}_\kappa$) is the full-induction arithmetic for the language $\langle 0, 1, -, +, a_\alpha \cdot, \leq \rangle_{\alpha \in \kappa}$, where all "scalars" $a_\alpha$ are non-standard.

Then:

- $\mathrm{Pr} = \mathrm{LA}_0$,
- $\mathrm{LA} = \mathrm{LA}_1$.

Let $\mathcal{M} = \langle M, 0, 1, -, +, a\cdot, \leq \rangle \models \mathrm{LA}$.

Then $\mathcal{M}$ can be equipped with a structure of a (discretely ordered) module over the ring $R_a = \mathbb{Q}[a] \cap M$.

More generally, any $\mathcal{M} = \langle M, 0, 1, -, +, a_\alpha\cdot, \leq \rangle_{\alpha\in\kappa} \models \mathrm{LA}_\kappa$ can be understood as a discretely ordered module over the ring $\mathbb{Q}[a_\alpha]_{\alpha\in\kappa} \cap M$.

Section 2

# Context

We want to understand definable sets in models of linear arithmetics (certain discretely ordered modules).

There are two closely related, but simpler situations:

$\langle M, 0, -, +, r\rangle_{r\in R}$ and $\langle M, 0, 1, -, +, \leq\rangle = \langle M, 0, 1, -, +, \leq, z\rangle_{z\in\mathbb{Z}}$

(unordered) modules

discretely ordered $\mathbb{Z}$-modules
$=$ models of Presburger arithmetics

(forget the ordering)

(forget the scalars)

# Unordered modules

The following is a classical result in the theory of modules:

**Theorem (Baur-Monk)**

*Let $\mathcal{M} = \langle M, 0, -, +, r\rangle_{r\in R}$ be a (left) module over a ring (associative, with 1) $R$.*
*Every formula is in $\mathcal{M}$ equivalent to a boolean combination of primitive positive formulas, i.e. to a boolean combination of formulas of the form $(\exists \bar{z})\psi_i$, where each $\psi_i$ is a system of linear equations.*

**Remark:** A formula in a language $L$ is called primitive positive, or pp-formula, if it is of the form $(\exists \bar{z}) \bigwedge_{i<n} \chi_i$, where $\chi_i$ are atomic formulas.

# Models of Pr

For models of Presburger arithmetic, we have:

**Theorem (Presburger)**

*Every formula is in* $\mathrm{Pr}$ *equivalent to a disjunction of primitive positive formulas, i.e. to a formula of the form* $\bigvee_{i<n}(\exists \overline{z})\psi_i$*, where each* $\psi_i$ *is a system of linear inequalities.*

**Question:** Do the pp-elimination results of Baur-Monk and Presburger generalize to arithmetics $\mathrm{LA}_\kappa$ with $\kappa > 0$?

We show that the answer is "Yes" if and only if $\kappa = 1$.

The reason is that for any $\mathcal{M} = \langle M, 0, 1, -, +, a\cdot, \leq \rangle \models \mathrm{LA}$ (but not for models of $\mathrm{LA}_\kappa$ with $\kappa \geq 2$), the ring $R_a = \mathbb{Q}[a] \cap M$ is a doded.

Section 3

# Dodeds and doded-modules

# Doded

An ordered integral domain $\mathrm{D} = \langle \mathrm{D}, 0, 1, +, -, \cdot, \leq \rangle$ is called a doded if it

- is discretely ordered by $\leq$, with 1 being the least positive element,
- is regularly quasi-Euclidean, i.e. the Euclidean algorithm in D is correctly defined and always stops in finitely many steps,
- has degrees, i.e. the equivalence classes of $q \sim r \Leftrightarrow (\exists n \in \mathbb{N}) q/n \leq r \leq nq$ on positive elements of $\mathrm{D}$ are ordered as an ordinal $\alpha \leq \omega$ (then there is a function $\deg : \mathrm{D} \to \mathbb{N} \cup \{-\infty\}$ such that $\mathrm{rng}(\deg)$ is an initial segment of $\mathbb{N} \cup \{-\infty\}$, and $\deg r \leq \deg q \Leftrightarrow |r| \leq n|q|$, for some $n \in \mathbb{N}$).

# Doded

**Example**

1. The ordered ring $\mathbb{Z}$ is a doded. The degree function is given by $\deg(z) = 0$ for $z \neq 0$ and $\deg(0) = -\infty$.
2. For $\mathcal{M} \models \mathrm{LA}$, the ring $R_a = \mathbb{Q}[a] \cap M$ is a doded. The degree function is the degree of polynomials.

**Remark:** Rings $R_a$ are quasi-Euclidean but not $k$-stage Euclidean for any $k \in \mathbb{N}$.

**Fact:** There are no other dodeds but those listed among the above examples.

# Doded-modules

Let $\mathbb{L}$ denotes the two-sorted language of "ordered rings-ordered modules".

A doded-module is a (two-sorted) $\mathbb{L}$-structure $\mathcal{A} = \langle \mathcal{R}, \mathcal{M}, \cdot \rangle$ such that

1) $\mathcal{R}$ is a doded,
2) $\langle \mathcal{M}, r \cdot \_ \rangle_{r \in R}$ is a discretely ordered (with 1 being the least positive element), integrally-divisible (i.e. $(\forall x)(\exists y)(\exists 0 \leq z < r \cdot 1)(x = r \cdot y + z)$ holds) $\mathcal{R}$-module.

# Doded-modules

Let $\mathbb{L}'$ denotes the extension of $\mathbb{L}$ by symbols $\frac{-1}{\mathcal{R}} : \mathcal{R}^2 \to \mathcal{R}$, $\frac{-1}{\mathcal{M}} : \mathcal{R} \times \mathcal{M} \to \mathcal{M}$ for integral division.

For a doded-module $\mathcal{A}$, let $\mathcal{A}'$ denotes the natural definable $\mathbb{L}'$-expansion of $\mathcal{A}$.

**Theorem**

*Let $\mathcal{A} = \langle \mathcal{R}, \mathcal{M}, \cdot \rangle$ be a doded-module, $\varphi(\overline{r}, \overline{y}, x)$ be an $\mathbb{L}'$-formula without scalar quantifiers, and $\overline{\rho} \in R^{l(\overline{r})}$ be scalars. Then there are finitely many $\mathbb{L}'$-terms $t_i(\overline{r}, \overline{y})$, for $i < n$, with $n \in \mathbb{N}$, such that*

$$\mathcal{A}' \models (\exists x)\varphi(\overline{\rho}, \overline{y}, x) \leftrightarrow \bigvee_{i<n} \varphi(\overline{\rho}, \overline{y}, t_i(\overline{\rho}, \overline{y})).$$

Let further on $D$ be a fixed doded and $\mathcal{M} = \langle M, 0, 1, -, +, \leq, r \rangle_{r \in D}$ be a fixed discretely ordered (with 1 being the least positive element), integrally-divisible $D$-module. Denote $\mathcal{M}'$ the expansion of $\mathcal{M}$ by definitions of functions $r^{-1}$ providing integral division by all scalars $0 < r \in D$.

**Corollary**

*For any formula $\varphi(x, \overline{y})$ of $\mathcal{M}'$ there are finitely many terms (of $\mathcal{M}'$) $t_i(\overline{y})$, for $i < n$, with $n \in \mathbb{N}$, such that*

$$\mathcal{M}' \models (\exists x)\varphi(x, \overline{y}) \leftrightarrow \bigvee_{i<n} \varphi(t_i(\overline{y}), \overline{y}).$$

*Hence $\mathcal{M}'$ has quantifier elimination.*

*In $\mathcal{M}$, every formula is equivalent to a disjunction of primitive positive formulas, i.e. to a formula of the form $\bigvee_{i<n}(\exists \overline{z})\psi_i$, where each $\psi_i$ is a system of linear inequalities.*

# Two-sorted QE for doded-modules does not hold

Note that in the two-sorted case of doded-modules the quantifier elimination cannot be deduced (as the Skolem terms depend on the scalar parameters). In fact, we can prove:

**Proposition**

*Let $\mathcal{A}$ be a doded-module which is a fragment of a model of $\mathrm{PA}$. Then* **not** *every scalar-quantifier-free $\mathbb{L}'$-formula is in $\mathcal{A}'$ equivalent to a quantifier-free $\mathbb{L}'$-formula.*

# A more detailed analysis

Further on, by a term or formula, we mean always an $\mathcal{M}'$-term or an $\mathcal{M}'$-formula.
We will denote by $\mathrm{C}$ the set of all realizations of constants terms in $\mathcal{M}'$.

A term $t(\overline{x})$ is harmonic if

$$t(\overline{x}) = \sum_{i=0}^{N-1} q_i {r_i}^{-1}(x_{f(i)}) + c,$$

for some $q_i, r_i \in \mathrm{D}$, $c \in \mathrm{C}$ and $f : N \to l(\overline{x})$.

A formula is harmonic if all its maximal subterms are.

A "piecewise-term" $\tau$ is called an almost-term if it is of the form

$$\tau(\overline{x}) = \begin{cases} s(\overline{x}) + c_i \text{ if } \psi_i(\overline{x}), \ i < n, \end{cases}$$

where $s(\overline{x})$ is a term, and $c_i \in \mathrm{C}$, for $i < n$.

# Harmonic form theorem

**Theorem (Harmonic form theorem)**

1) *For every term $t(\overline{x})$, there is an open harmonic almost-term $\tau(\overline{x})$ such that $\mathcal{M}' \models t(\overline{x}) = \tau(\overline{x})$.*
2) *For every formula $\varphi(\overline{x})$, there is an open harmonic formula $\psi(\overline{x})$ such that $\mathcal{M}' \models \varphi(\overline{x}) \leftrightarrow \psi(\overline{x})$.*

# Representation of definable sets

**Corollary**

*Every set $A \subseteq M^n$ $X$-definable in $\mathcal{M}'$ (for $X \subseteq M$) can be written as*

$$A = \bigcup_{i<k} g[P_i],$$

*where $g : (K(\bar{a}) \times M)^n \to M^n$ is a "linear coordination" of $M^n$, $\bar{a} = (a_0, \ldots, a_l) \in M^l$, $l \in \mathbb{N}$, and $P_i$, for $i < k$, are finitely many polyhedra in $(K(\bar{a}) \times M)^n$ over parameters from $X$.*

"Every definable set is a finite union of linear images of polyhedra."

Section 4

# Properties of LA

## Theorem (Properties of LA)

1) LA *is model-complete.*
*Moreover: Every formula is in* LA *equivalent to a disjunction of primitive positive formulas, i.e. to a formula of the form* $\bigvee_{i<n}(\exists \overline{z})\psi_i$, *where each* $\psi_i$ *is a system of linear inequalities.*

2) *For* $\mathcal{A}, \mathcal{B} \models \mathrm{LA}$, *it is* $\mathcal{A} \equiv \mathcal{B} \Leftrightarrow a^{\mathcal{A}} \equiv a^{\mathcal{B}} \mod n$, *for all* $0 < n \in \mathbb{N}$.
$\mathrm{LA}_\tau = \mathrm{LA} + \{a \equiv \tau(p,k) \mod p^k; p \in \mathbb{N}$ *prime*, $k \in \mathbb{N}\}$, *for* $\tau \in \prod_{p\in\mathbb{P}} \mathbb{J}_p$, *are all simple complete extensions of* LA.

3) $R_\tau = \{r(a)/n; 0 \leq r(a) \in \mathbb{Z}[a],\ 0 \neq n \in \mathbb{N},\ p^k | r(\tau(p,k)) \forall p^k | n\}$ *is the unique prime model of* $\mathrm{LA}_\tau$, *for* $\tau \in \prod_{p\in\mathbb{P}} \mathbb{J}_p$.

4) LA *is decidable.*
$\mathrm{LA}_\tau$ *is decidable if and only if* $\tau$ *is recursive.*

5) *The induction scheme in* LA *may be equivalently replaced by the scheme of integral divisibility*
$$(\exists y, z)(x = qy + z \,\&\, z < q), \text{ for all } 0 < q \in \mathbb{Z}[a].$$

**Corollary**

*Up to elementary equivalence, models of* $\mathrm{LA}$ *are exactly all ultraproducts*

$$\mathcal{Z}_{\mathcal{U}} = (\prod_{n\in\mathbb{N}} \langle \mathbb{Z}, 0, 1, +, -, \underline{n}, \leq \rangle)/\mathcal{U},$$

*where* $\mathcal{U}$ *is a non-principal ultrafilter on* $\mathbb{N}$*, i.e.* $\mathcal{U} \in \beta\mathbb{N} - \mathbb{N}$.

**Note:** Properties of $\mathrm{LA}$ and Pr are similar, but the proof for $\mathrm{LA}$ is incomparably harder (due to [in]decomposability of sets $\{qx + ry; 0 \leq x, y \in A\}$ in $\mathcal{A} \models \mathrm{Pr}[\mathrm{LA}]$).

The following corollary is on the structure of models of Peano arithmetic:

**Corollary**

*Let $\mathcal{M} = \langle M, 0, 1, -, +, \cdot, \leq \rangle$ be a saturated model of Peano arithmetic, $0 \leq a \in M - \mathbb{N}$, $c, d \in M$. Then the following are equivalent:*

1) *For every "Peano multiplication" $\circ$ on $\mathcal{M}$ with $a \cdot x = a \circ x$ for all $x$, it is $c \cdot d = c \circ d$.*
2) *$c \in \mathbb{Q}[a]$ or $d \in \mathbb{Q}[a]$.*

Recall that Presburger arithmetic is dp-minimal and thus also NIP. For $\mathrm{LA}$, this is different:

**Theorem**

$\mathrm{LA}$ *is NIP but not dp-minimal.*

**Proof sketch:** For the NIP property use the description of definable sets (in all dimensions) as linear images of finite unions of polyhedra, i.e. solutions of systems of linear inequalities. But linear inequalities have VC-dimension 1 hence are NIP.

But $\mathrm{LA}$ cannot be dp-minimal because $x \mapsto (x \operatorname{div} a, x \bmod a)$ is a definable function which is a bijection between infinite sets $[0, a^2)$ and $[0, a)^2$.

Section 5

# Properties of $LA_2$ and above

# Wild models of $LA_2$

The pp-elimination, though true for $\mathrm{Pr} = \mathrm{LA}_0$ and $\mathrm{LA} = \mathrm{LA}_1$, does not hold for $\mathrm{LA}_\kappa$ with $\kappa \geq 2$.

This is an easy consequence of the following:

**Proposition (with P. Pudlák)**

*There is a model $\mathcal{M} = \langle M, 0, 1, -, +, a\cdot, b\cdot, \leq \rangle \models \mathrm{LA}_2$ and a non-standard $I \in M$ such that $\cdot \restriction [0, I]^2$ is definable in $\mathcal{M}$ for some Peano multiplication $\cdot$ on $\mathcal{M}$ (i.e. such that $\langle M, 0, 1, +, \cdot, \leq \rangle \models \mathrm{PA}$).*

Note that no such model can exist for $\mathrm{LA}_1$ (easy consequence of model-completeness).

**Corollary**

*For $\kappa \geq 2$, the theory $\mathrm{LA}_\kappa$ does not have pp-elimination. It is not even model complete (i.e. does not have elimination to $\exists$-formulas).*

**Proof idea:**

In a saturated model of Peano arithmetic, for any $\mathbb{N} < I \in M$, we find elements $a, b$ such that the sequence $(1, 1^2, 2, 2^2, \ldots, 2I, (2I)^2)$ is encoded by the set of all numerators of convergents of the continued fraction of $a/b$. The crucial observation is that this set is definable in the language of $\mathrm{LA}_2$ with $a, b$ as the two scalars.

(Note that for $\mathrm{LA}_1$ every continued fraction of scalars is finite, thus the construction above does not allow to define an infinite part of a Peano multiplication.)

The following is well known:

**Proposition**

*(Unordered) modules are stable and thus NIP.*
*Presburger arithmetic (although unstable) is NIP.*

**Question (Chernikov and Hils):** Is every ordered module NIP?

**Corollary (with P. Pudlák)**

*The model $\mathcal{M}$ from the previous proposition, considered as an ordered module over $\mathbb{Z}[a, b]/_{=_M}$, is not NIP.*

Despite not having the pp-elimination, the hierarchy for higher order linear arithmetics still collapses. This is a consequence of the following more general result:

Let $\mathcal{M} = \langle M, 0, 1, +, -, \leq, r\rangle_{r\in R}$ be an ordered $R$-module with a unit $1 > 0$. We say that $\mathcal{M}$ is i-divisible if for every $0 < r \in R$ it is $\mathcal{M} \models (\forall x)(\exists y, z)(x = ry + z \,\&\, 0 \leq z < r1)$.

**Theorem**

*Let $\mathcal{M}$ be an i-divisible ordered module. Then every formula is in $\mathcal{M}$ equivalent to a bounded formula.*

**Corollary**

*For any $\kappa$, every $\mathcal{M} \models \mathrm{LA}_\kappa$ has bounded quantifier elimination. Hence, in no model of $\mathrm{LA}_\kappa$ a Peano multiplication is definable on the whole universe.*

**Proof idea:**

The crucial part is to prove that in $\mathcal{M}$, every formula $\psi(\overline{x})$ is protoperiodic, i.e. there exists a disjunct covering $\mathcal{A}$ of $M^{l(\overline{x})}$ by finitely many convex polyhedra, such that for any $A \in \mathcal{A}$ and a direction $s = \overline{\alpha} \in R^{l(\overline{x})}$ either $A$ is bounded in the direction $s$ ($\exists 0 \leq m \in M$ s.t. $A + sm \not\subseteq A$) or there is a period $0 < P \in R$ for $\psi$ on $A$ in the direction $s$ (i.e. $(\forall \overline{u}, \overline{v} \in A)((\exists m \in M)\overline{u} = Psm + \overline{v} \rightarrow (\psi(\overline{u}) \leftrightarrow \psi(\overline{v})))$).

This is done by induction on complexity of $\psi$.

# Tame and wild linear fragments of PA

Let $\mathcal{M} \models PA$ and let $R$ be a subring of $\mathcal{M} \cup -\mathcal{M}$. Denote $\mathcal{M}_R$ the $R$-(semi)module fragment of $\mathcal{M}$. We know the following:

**Proposition**

*If $R$ is a doded then $M_R$ has pp-elimination (hence is model complete).*

But we can also prove the following:

**Proposition**

*If $R$ contains elements $a, b$ such that $a/b$ has an infinite continued fraction (in $\mathcal{M}$) then $M_R$ is not model complete (hence does not have pp-elimination).*

**Proof sketch:** The set $C$ of all pairs $(u, v)$ such that $u/v$ is a convergent of $a/b$ (in the lowest terms) is infinite and definable in $M_R$. It is easy to show that $C$ is a (graph of a) concave sequence. But then it cannot be defined by an $\exists$-formula, as each infinite $\exists$-definable set in $M_R$ contains three points lying on a line.

# Open questions and Thank you

**Open questions:**

Is it possible to decide pp-elimination/model completeness of $M_R$ also in the remaining cases (when $R$ is not a doded but does not contain $a, b$ with an infinite continued fraction)? Do such cases exist at all?

Is there a model $M_R$ which is not model complete but is still NIP?

What about full-induction modules which are not fragments of models of $\mathrm{PA}$?

In particular, is there a non-trivial (i.e. with $R = \mathbb{Z}[a, b] \cap M$ not a doded) model of $\mathrm{LA}_2$ which has pp-elimination/is model complete/is NIP?

Can the bounded quantifier elimination for models of $\mathrm{LA}_\kappa$ with $\kappa \geq 2$ be strengthened? If yes, how does it depend on $\kappa$?

Thank you for your attention.