A canonical form and the distribution of values of generalized polynomials

A. Leibman
Department of Mathematics
The Ohio State University
Columbus, OH 43210, USA
e-mail: leibman@math.ohio-state.edu

January 16, 2013

Abstract

Generalized polynomials are functions obtained from conventional polynomials by applying the operations of taking the integer part, addition, and multiplication. We construct a system of "basic" generalized polynomials with the property that any bounded generalized polynomial is representable as a piecewise polynomial function of these basic ones. Such a representation is unique up to a function vanishing almost everywhere, which solves the problem of determining whether two generalized polynomials are equal a.e. The basic generalized polynomials are jointly equidistributed, thus we also obtain an effective algorithm of finding the limiting distribution of values of one or several generalized polynomials.

0. Introduction

A generalized polynomial is a function constructed from ordinary real-valued polynomials (of one or several variables) with the help of the operations of taking the integer part, addition and multiplication, like $p_1[p_2+p_3[p_4]][p_5]+p_6$, where p_i are ordinary polynomials of one or several variables. (It would be more natural to call generalized polynomials "bracket polynomials" (as it is done in [GT]), but unfortunately the term "bracket polynomial" is already reserved in knots theory.) We will call generalized polynomials *gen-polynomials*. Gen-polynomials with a common domain form an algebra, and a composition of gen-polynomials is again a gen-polynomial. The "fractional part" function $\{\!\{x\}\!\}$ and the "distance to the nearest integer" function $\langle x\rangle$ are also gen-polynomials: $\{\!\{x\}\!\}=x-[x]$, and $\langle x\rangle=\bigl(1-\bigl[\frac{1}{2}+\{\!\{x\}\!\}\bigr]\bigr)\{\!\{x\}\!\}+\bigl[\frac{1}{2}+\{\!\{x\}\!\}\bigr](1-\{\!\{x\}\!\})$.

We will mainly be interested in gen-polynomials of one or several *integer* variables, that is, gen-polynomials $\mathbb{Z}^d\longrightarrow\mathbb{R}$; we will explain later why gen-polynomials of real variables are easier to deal with.

Gen-polynomials are very natural objects, and often appear in literature in this or that form. Wide classes of gen-polynomials were studied and many nice results were obtained in [Hå1], [Hå2], [Hå3], [HåK], and [BHå]. In [GO'B], interesting questions about the uniqueness of a representation of a gen-polynomial in a certain form are asked and answered for some special classes of gen-polynomials.

It was noticed in [AGH] that if $X=G/\Gamma$, where G is the Heisenberg group $\left(\begin{smallmatrix}1&\mathbb{R}&\mathbb{R}\\0&1&\mathbb{R}\\0&0&1\end{smallmatrix}\right)$ and $\Gamma=\left(\begin{smallmatrix}1&\mathbb{Z}&\mathbb{Z}\\0&1&\mathbb{Z}\\0&0&1\end{smallmatrix}\right)\subset G$, is identified with the subset of G of matrices whose all entries are in $[0,1)$, then for any $a\in G$ the upper-right entry of the sequence $a^n \bmod \Gamma$, $n\in\mathbb{Z}$, is a gen-polynomial of the form $a_1n[a_2n]+a_3n^2+a_4n$, $a_i\in\mathbb{R}$. The space X above is a *nilmanifold*, a compact homogeneous space of a nilpotent Lie group. It was V. Bergelson who asked what gen-polynomials can be obtained in a similar way, that is, can be read off from a nilmanifold. In [BL] it is proven that *any* bounded gen-polynomial can: for any bounded gen-polynomial $u\colon\mathbb{Z}^d\longrightarrow\mathbb{R}$ there exists a nilmanifold $X=G/\Gamma$, a point $x\in X$, a piecewise polynomial function τ on X, and a polynomial mapping $P\colon\mathbb{Z}^d\longrightarrow G$ such that $\overline{P(\mathbb{Z}^d)x}=X$ and $u(n)=\tau(P(n)x)$ for all $n\in\mathbb{Z}^d$. (There is also a version of this theorem for gen-polynomials on $\mathbb{R}^d$.) Various properties of bounded gen-polynomials can be derived from this result; in particular, it follows that any bounded vector-valued gen-polynomial u has "a limiting distribution of values": the values of u are uniformly distributed on a piecewise polynomial surface. The

Supported by NSF grants DMS-0600042 and DMS-0901106.

proof of this fact given in [BL] is long and difficult; we give here a simpler proof (see Theorems 7.2 and 8.1; yet another simple proof is outlined in Section 9.)

In connection with gen-polynomials, the following two problems naturally appear:

1. "The word problem": given two gen-polynomials, determine whether they represent the same function.
2. "The distribution problem": given a (vector-valued) gen-polynomial u, determine whether the values of u are equidistributed in the range of u; more generally, find the limiting distribution of the values of u.

To solve these problems we describe a "canonical form" to which every gen-polynomial can be reduced. We introduce a system of "basic gen-polynomials", which are highly independent (jointly equidistributed, to be exact), and show that every bounded gen-polynomial is representable as a piecewise polynomial function of these basic ones. (As to the unbounded gen-polynomials, they can be represented as "polynomials" with bounded gen-polynomials as coefficients.) The independence of basic gen-polynomials implies that, for every bounded gen-polynomial, such a representation is unique up to a piecewise polynomial function vanishing a.e.; this solves the "word problem" for gen-polynomials – modulo gen-polynomials vanishing a.e. Also, the independence of basic gen-polynomials implies that the values of the composition of a function f and the basic gen-polynomials are distributed uniformly with respect to the measure $f(\mu)$, where μ is the Lebesgue measure on a unit cube, which solves the "distribution problem" for gen-polynomials.

For some technical reasons, we find it more convenient to construct gen-polynomials using the fractional part function $\{\!\!\{x\}\!\!\} = x - [x]$ instead of the integer part function $[x]$. (We will use this unusual notation for the fractional part in order not to mix it up with the set-theoretical parentheses "$\{\dots\}$".) As the reader will see, our construction of "basic gen-polynomials" is related to commutator calculus in free groups. Let $\mathcal{A}$ be a well-ordered set. We define a well-ordered "index set" $\mathcal{B} = \mathcal{B}(\mathcal{A})$ in the following way: $\mathcal{B}$ is the minimal set that contains all elements of $\mathcal{A}$ and all expressions of the form $[\gamma, m\beta]$ with $\beta, \gamma \in \mathcal{B}$, $m \in \mathbb{N}$, such that $\beta < \gamma$ and either $\gamma \in \mathcal{A}$ or $\gamma = [\lambda, k\delta]$ with $\delta < \beta$, and where the order on $\mathcal{B} \setminus \mathcal{A}$ is defined by $\mathcal{A} < (\mathcal{B} \setminus \mathcal{A})$ and $[\gamma_1, m_1\beta_1] < [\gamma_2, m_2\beta_2]$ if $(\beta_1, \gamma_1, m_1) < (\beta_2, \gamma_2, m_2)$ lexicographically. (Thus, to verify whether an expression α belongs to $\mathcal{B}$ one has to perform finitely many tests.) Less formally, $\mathcal{B}$ is constructed in the following way: we start with $\mathcal{B} = \mathcal{A}$; then, beginning with the first element of $\mathcal{A}$, for each successive element β of $\mathcal{B}$ we add at the end of $\mathcal{B}$ the set $\bigcup_{\substack{\gamma\in\mathcal{B}\\ \gamma>\beta}} \bigcup_{m=1}^{\infty} [\gamma, m\beta]$; this way we obtain $\mathcal{B} = \mathcal{A} \cup \bigcup_{\beta\in\mathcal{B}} \mathcal{E}_\beta$, where

$$\mathcal{E}_\beta = \bigcup_{\substack{\gamma\in\bigcup_{\delta<\beta}\mathcal{E}_\delta\\ \gamma>\beta}} \bigcup_{m=1}^{\infty} [\gamma, m\beta].$$

(Actually, we could finitize this definition by taking a finite $\mathcal{A}$, limiting m by a constant c, and limiting the length of elements of $\mathcal{B}$, – we then obtain a finite set $\mathcal{B}$ that would suffice for our goals. The only reason we do not do this is that we do not want to permanently keep these limits in mind; but the reader may always assume that $\mathcal{B}$ is finite.)

Example. Let $\mathcal{A} = \{a, b, c\}$, with $a < b < c$. Then the beginning elements of $\mathcal{B} = \mathcal{B}(\mathcal{A})$ are:

$$\begin{array}{rl}
 & a,\ b,\ c; \\
\mathcal{E}_a: & [b, ma],\ m\in\mathbb{N}; \quad [c, ma],\ m\in\mathbb{N}; \\
\mathcal{E}_b: & [c, mb],\ m\in\mathbb{N}; \quad [[b, ma], kb],\ k, m\in\mathbb{N}; \quad [[c, ma], kb],\ k, m\in\mathbb{N}; \\
\mathcal{E}_c: & [[b, ma], lc],\ l, m\in\mathbb{N}; \quad [[c, ma], lc],\ l, m\in\mathbb{N}; \quad [[c, mb], lc],\ l, m\in\mathbb{N}; \\
 & [[[b, ma], kb], lc],\ l, k, m\in\mathbb{N}; \quad [[[c, ma], kb], lc],\ l, k, m\in\mathbb{N}; \\
\mathcal{E}_{[b,a]}: & [[b, ma], r[b, a]],\ r, m\in\mathbb{N},\ m\ge 2; \quad [[c, ma], r[b, a]],\ r, m\in\mathbb{N}; \quad [[c, mb], r[b, a]],\ r, m\in\mathbb{N}; \\
 & [[[b, ma], kb], r[b, a]],\ r, k, m\in\mathbb{N}; \quad [[[c, ma], kb], r[b, a]],\ r, k, m\in\mathbb{N}; \\
 & [[[b, ma], lc], r[b, a]],\ r, l, m\in\mathbb{N}; \quad [[[c, ma], lc], r[b, a]],\ r, l, m\in\mathbb{N}; \quad [[[c, mb], lc], r[b, a]],\ r, l, m\in\mathbb{N}; \\
 & [[[[b, ma], kb], lc], r[b, a]],\ r, l, k, m\in\mathbb{N}; \quad [[[[c, ma], kb], lc], r[b, a]],\ r, l, k, m\in\mathbb{N}; \\
\mathcal{E}_{[b,2a]}: & [[b, ma], r[b, 2a]],\ r, m\in\mathbb{N},\ m\ge 3; \quad [[c, ma], r[b, 2a]],\ r, m\in\mathbb{N}; \quad \dots \\
\vdots &
\end{array}$$

where we assume $[b, a] = [b, 1a]$.

We define gen-polynomials $v_\alpha = v_\alpha(\mathcal{P})$, $\alpha \in \mathcal{B}$, in the variables x_δ, $\delta \in \mathcal{A}$, in the following way: we put $v_\alpha = x_\alpha$, $\alpha \in \mathcal{A}$, and for $\alpha = [\gamma, m\beta]$, $\gamma, \beta \in \mathcal{B}$, $m \in \mathbb{N}$, we put $v_\alpha = v_\gamma \{v_\beta\}^m$. We call the gen-polynomials $\{v_\alpha\}$ *basic gen-polynomials.*

Examples. 1. For $a, b, c, d, e \in \mathcal{A}$ and $m, k, r, l \in \mathbb{N}$, we have $v_{[a,mb]} = x_b\{x_a\}^m$, $v_{[[c,ma],kb]} = x_c\{x_a\}^m\{x_b\}^k$, $v_{[[[c,ma],kb],r[d,le]]} = x_c\{x_a\}^m\{x_b\}^k\{x_d\{x_e\}^l\}^r$.

2. For $\mathcal{A} = \{a, b, c\}$, assuming $a < b < c$, the beginning basic gen-polynomials are

$\{v_a\} = \{x_a\}$, $\{v_b\} = \{x_b\}$, $\{v_c\} = \{x_c\}$;

—

$\{v_{[b,ma]}\} = \{x_b\{x_a\}^m\}$, $m \in \mathbb{N}$;
$\{v_{[c,ma]}\} = \{x_c\{x_a\}^m\}$, $m \in \mathbb{N}$;

—

$\{v_{[c,mb]}\} = \{x_c\{x_b\}^m\}$, $m \in \mathbb{N}$;
$\{v_{[[b,ma],kb]}\} = \{x_b\{x_a\}^m\{x_b\}^k\}$, $k, m \in \mathbb{N}$;
$\{v_{[[c,ma],kb]}\} = \{x_c\{x_a\}^m\{x_b\}^k\}$, $k, m \in \mathbb{N}$;

—

$\{v_{[[b,ma],lc]}\} = \{x_b\{x_a\}^m\{x_c\}^l\}$, $l, m \in \mathbb{N}$;
$\{v_{[[c,ma],lc]}\} = \{x_c\{x_a\}^m\{x_c\}^l\}$, $l, m \in \mathbb{N}$;
$\{v_{[[c,mb],lc]}\} = \{x_c\{x_b\}^m\{x_c\}^l\}$, $l, m \in \mathbb{N}$;
$\{v_{[[[b,ma],kb],lc]}\} = \{x_b\{x_a\}^m\{x_b\}^k\{x_c\}^l\}$, $l, k, m \in \mathbb{N}$;
$\{v_{[[[c,ma],kb],lc]}\} = \{x_c\{x_a\}^m\{x_b\}^k\{x_c\}^l\}$, $l, k, m \in \mathbb{N}$;

—

$\{v_{[[b,ma],r[b,a]]}\} = \{x_b\{x_a\}^m\{x_b\{x_a\}\}^r\}$, $r, m \in \mathbb{N}$, $m \geq 2$;
$\{v_{[[c,ma],r[b,a]]}\} = \{x_c\{x_a\}^m\{x_b\{x_a\}\}^r\}$, $r, m \in \mathbb{N}$;
$\{v_{[[c,mb],r[b,a]]}\} = \{x_c\{x_b\}^m\{x_b\{x_a\}\}^r\}$, $r, m \in \mathbb{N}$;
$\{v_{[[[b,ma],kb],r[b,a]]}\} = \{x_b\{x_a\}^m\{x_b\}^k\{x_b\{x_a\}\}^r\}$, $r, k, m \in \mathbb{N}$;
$\{v_{[[[c,ma],kb],r[b,a]]}\} = \{x_c\{x_a\}^m\{x_b\}^k\{x_b\{x_a\}\}^r\}$, $r, k, m \in \mathbb{N}$;
$\{v_{[[[b,ma],lc],r[b,a]]}\} = \{x_b\{x_a\}^m\{x_c\}^l\{x_b\{x_a\}\}^r\}$, $r, l, m \in \mathbb{N}$;
$\{v_{[[[c,ma],lb],r[b,a]]}\} = \{x_c\{x_a\}^m\{x_c\}^l\{x_b\{x_a\}\}^r\}$, $r, l, m \in \mathbb{N}$;
$\{v_{[[[c,mb],lc],r[b,a]]}\} = \{x_c\{x_b\}^m\{x_c\}^l\{x_b\{x_a\}\}^r\}$, $r, l, m \in \mathbb{N}$;
$\{v_{[[[[b,ma],kb],lc],r[b,a]]}\} = \{x_b\{x_a\}^m\{x_b\}^k\{x_c\}^l\{x_b\{x_a\}\}^r\}$, $r, l, k, m \in \mathbb{N}$;
$\{v_{[[[[c,ma],kb],lc],r[b,a]]}\} = \{x_c\{x_a\}^m\{x_b\}^k\{x_c\}^l\{x_b\{x_a\}\}^r\}$, $r, l, k, m \in \mathbb{N}$;

—

$\{v_{[[b,ma],r[b,2a]]}\} = \{x_b\{x_a\}^m\{x_b\{x_a\}^2\}^r\}$, $r, m \in \mathbb{N}$, $m \geq 3$;
$\{v_{[[c,ma],r[b,2a]]}\} = \{x_c\{x_a\}^m\{x_b\{x_a\}^2\}^r\}$, $r, m \in \mathbb{N}$;

$\vdots$

We will now give several definitions. Let φ be a mapping from a countable abelian group H to a topological space X equipped with a probability Borel measure ν. *A Følner sequence* in H is a sequence (Φ_N) of finite subsets of H such that for every $n \in H$, $\frac{1}{|\Phi_N|}|(\Phi_N + n)\triangle\Phi_N| \longrightarrow 0$ as $N \longrightarrow \infty$. we will say that φ is (or the values of φ are) *well distributed in X with respect to ν* if for any $h \in C(X)$ and any Følner sequence (Φ_N) in H we have $\lim_{N\to\infty} \frac{1}{|\Phi_N|}\sum_{n\in\Phi_N} h(\varphi(n)) = \int_X h\,d\nu$. Let u_α, $\alpha \in \mathcal{A}$, be a family of functions $H \longrightarrow [0,1]$; we will say that the functions u_α are *jointly well distributed* if for any finite subsystem $\{\alpha_1, \ldots, \alpha_k\} \subseteq \mathcal{A}$ the function $(u_{\alpha_1}, \ldots, u_{\alpha_k})$ from H to $[0,1]^k$ is well distributed with respect to the Lebesgue measure on $[0,1]^k$.

If v is a function in real variables x_α, $\alpha \in \mathcal{A}$, and $\mathcal{P}\{p_\alpha : \alpha \in \mathcal{A}\}$ is a system of real-valued functions with common domain Z, we will denote by $v(\mathcal{P})$ the function $v\big(p_\alpha(n) : \alpha \in \mathcal{A}\big)$, $n \in Z$.

By $\mathbb{Q}[n]$ we will denote the space of polynomials of one or several integer variables with rational coefficients. Our first main result is the following:

Theorem 0.1. *Let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$, be a well-ordered system of real-valued polynomials on $\mathbb{Z}^d$, $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n] + \mathbb{R}$. Then the gen-polynomials $\{\!\{v_\alpha(\mathcal{P})\}\!\}$, $\alpha \in \mathcal{B}(\mathcal{A})$, are jointly well distributed.*

Remark. The well-order on the system $\mathcal{P}$ appearing in the formulation of Theorem 0.1 and below, is arbitrary, and is simply used to construct the gen-polynomials $\{\!\{v_\alpha(\mathcal{P})\}\!\}$. This order may not be anyhow related to the degree of the polynomials in P.

We call *a piecewise polynomial function*, or *a pp-function*, any function f on $Q \subseteq \mathbb{R}^m$ such that Q can be partitioned into finitely many subsets, $Q = \bigcup_{i=1}^k Q_i$, with the property that, for each i, Q_i is defined by a system of polynomial inequalities and $f|_{Q_i}$ is a polynomial.

Example. $f(x,y) = \begin{cases} xy+1, \text{ if } y<x^2,\ x<y^2 \\ x^2+3y, \text{ if } y<x^2,\ x\geq y^2 \\ x^3+2xy^2-3, \text{ if } y\geq x^2,\ x<y^2 \\ 5, \text{ if } y\geq x^2,\ x\geq y^2 \end{cases}$ is a pp-function on $\mathbb{R}^2$.

Let u be a bounded gen-polynomial. Let R be the set of ordinary polynomials from which u is constructed (that is, the polynomials occurring in the expression defining u), let $\mathcal{R}$ be the $\mathbb{Q}$-algebra generated by R, and let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ be a system of polynomials such that $\operatorname{span}\mathcal{P} \supseteq \mathcal{R}$. We would now like to state that u can be represented as a pp-function of the gen-polynomials v_α, but there are two obstacles that prevent this from being true. The first obstacle is the problem of rational coefficients: the gen-polynomial $\{\!\{\frac{1}{2}p\}\!\}$ is not a function of the gen-polynomial $\{\!\{p\}\!\}$, so we cannot avoid dealing with $\{\!\{\frac{1}{2}p\}\!\}$, and more generally, with $\{\!\{M^{-1}p\}\!\}$ with M "divisible enough". The second obstacle is related to rational polynomials: the gen-polynomials $u(n) = \{\!\{\frac{1}{2}n\}\!\}$, $n \in \mathbb{Z}$, cannot be represented in the desirable form. We could overcome this second problem by adding rational basic gen-polynomials to our scheme, but for simplicity we have preferred to trivialize rational polynomials by passing to a suitable sublattice of $\mathbb{Z}^d$: for $u(n) = \{\!\{\frac{1}{2}n\}\!\}$, $u|_{2\mathbb{Z}} \equiv 0$ and $u|_{1+2\mathbb{Z}} \equiv 1$.

Given a system of functions $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$, we will write $M^{-1}\mathcal{P}$ for the system $\{M^{-1}p_\alpha : \alpha \in \mathcal{A}\}$. The "true" formulation of our second main result is the following:

Theorem 0.2. *Let u be a bounded gen-polynomial on $\mathbb{Z}^d$, and let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ be a well-ordered system of polynomials spanning the $\mathbb{Q}$-algebra $\mathcal{R}$ generated by the polynomials occurring in u. There exist $M \in \mathbb{N}$ and a cofinite sublattice Λ in $\mathbb{Z}^d$ such that for any translation $\Lambda' = n_0 + \Lambda$ of Λ there exist distinct indices $\alpha_1, \ldots, \alpha_l \in \mathcal{B}(\mathcal{A})$, and a pp-function f on $[0,1]^l$ such that $u|_{\Lambda'} = f\big(\{\!\{v_{\alpha_1}(M^{-1}\mathcal{P})\}\!\}, \ldots, \{\!\{v_{\alpha_l}(M^{-1}\mathcal{P})\}\!\}\big)|_{\Lambda'}$. If the system $\mathcal{P}$ is $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n] + \mathbb{R}$ and the constant M is fixed, then, for any Λ', f is defined uniquely up to a (piecewise polynomial) function vanishing a.e. in $[0,1]^l$.*

Remarks. 1. Since the gen-polynomials v_{α_j}, $j = 1, \ldots, l$, only depend on finitely many variables, we really need only a finite subset of $\mathcal{P}$, and so, only a system of polynomials that spans not the whole algebra $\mathcal{R}$ but only the space $\mathcal{R}_{\leq c} = \big\{g(q_1, \ldots, q_k),\ k \in \mathbb{N},\ q_i \text{ are polynomials occurring in } u,\ g \in \mathbb{Q}[x_1, \ldots, x_k],\ \deg g \leq c\big\}$ for some $c \in \mathbb{N}$. Moreover, the integer c can be estimated by analyzing the structure of u (it does not exceed the *complexity* of u; see below), but we do not prove such an estimate.

2. What polynomial "occur" in u depends, of course, on a concrete expression defining u, thus the algebra $\mathcal{R}$ appearing in the formulation of Theorem 0.2 is not well defined. What we mean is that we can choose any representation of u as a gen-polynomial and take $\mathcal{R}$ to be the algebra generated by the polynomials that occur in this representation.

3. The constraint that our canonical form is only defined up to a pp-function f_0 vanishing a.e. is unavoidable, since the sequence $f_0\big(\{\!\{v_{\alpha_j}(M^{-1}\mathcal{P}(n))\}\!\} : j = 1, \ldots, l\big)$, $n \in \mathbb{Z}^d$, may be identically zero for such a function f_0. For example, for the function $f_0(x,y) = \chi_{\{0\}}(x) - \chi_{\{0\}}(y)$ we have $f_0\big(\{\!\{\sqrt{2}n\}\!\}, \{\!\{\sqrt{3}n\}\!\}\big) = 0$ for all n. As a result, the gen-polynomial $u(n) = 1 + \big[-\{\!\{\sqrt{2}n\}\!\}\big] = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$ on $\mathbb{Z}$ can be represented as $\chi_{\{0\}}(\{\!\{\sqrt{2}n\}\!\})$ and also as $\chi_{\{0\}}(\{\!\{\sqrt{3}n\}\!\})$, though the polynomials $\sqrt{2}n$ and $\sqrt{3}n$ are $\mathbb{Q}$-linearly independent.

4. Note also that when a pp-function vanishes a.e. in $[0,1]^l$, it is only nonzero on a (polynomially defined) subset of a proper algebraic subvariety of $[0,1]^l$.

Examples. 1. Let $u(n)=\{\![\sqrt{2}n]\sqrt{3}n\}\!\}$, $n\in\mathbb{Z}$. The gen-polynomial u "is constructed" from the polynomials $p(n)=\sqrt{2}n$ and $q(n)=\sqrt{3}n$; we also add the polynomial $r(n)=p(n)q(n)=\sqrt{6}n^2$ from the algebra generated by p and q. We have

$$u=\{\![p]q\}\!\}=\{\!pq-q\{\!p\}\!\}\!\}=\{\!r-q\{\!p\}\!\}\!\}=\{\!\{\!r\}\!\}-\{\!q\{\!p\}\!\}\!\}\!\}=\{\!\{\!v_3\}\!\}-\{\!v_{[2,1]}\}\!\}\!\}=f\big(\{\!v_3\}\!\},\{\!v_{[2,1]}\}\!\}\big),$$

where $v_3=r$, $v_{[2,1]}=q\{\!p\}\!\}$, and f is the pp-function $f(x,y)=\{\!x-y\}\!\}=\begin{cases}x-y,\ x-y\geq 0\\ x-y+1,\ x-y<0\end{cases}$ on $[0,1]^2$. Since p,q,r are $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$, the gen-polynomials $\{\!v_3\}\!\}$ and $\{\!v_{[2,1]}\}\!\}$ are jointly well distributed; since the image by f of the Lebesgue measure on $[0,1]^2$ is the Lebesgue measure on $[0,1]$, it follows that u is also well distributed on $[0,1]$ with respect to the Lebesgue measure.

2. Let $u(n)=\{\!\sqrt{2}n\{\!\sqrt{3}n\{\!\sqrt{2}n\}\!\}\!\}-\sqrt{3}n\{\!\sqrt{2}n\}\!\}[\sqrt{2}n]\}\!\}$. Put $p=\sqrt{2}n$, $q=\sqrt{3}n$, and transform u in the following way:

$$\begin{aligned}u(n)&=\{\!p\{\!q\{\!p\}\!\}\!\}-q\{\!p\}\!\}[p]\}\!\}=\{\!\{\!p\}\!\}\{\!q\{\!p\}\!\}\!\}+[p]\{\!q\{\!p\}\!\}\!\}-q\{\!p\}\!\}[p]\}\!\}=\{\!\{\!p\}\!\}\{\!q\{\!p\}\!\}\!\}+[p](\{\!q\{\!p\}\!\}\!\}-q\{\!p\}\!\})\}\!\}\\ &=\{\!\{\!p\}\!\}\{\!q\{\!p\}\!\}\!\}-[p]\big[q\{\!p\}\!\}\big]\}\!\}=\{\!\{\!p\}\!\}\{\!q\{\!p\}\!\}\!\}\!\}=\{\!p\}\!\}\{\!q\{\!p\}\!\}\!\}=\{\!v_p\}\!\}\{\!v_{[q,p]}\}\!\}\end{aligned}$$

where $v_p=p$ and $v_{[q,p]}=q\{\!p\}\!\}$. It follows that the values of u are well distributed in $[0,1]$ with respect to the image of the Lebesgue measure on $[0,1]^2$ under the function $f(x_1,x_2)=x_1x_2$, namely, the measure $-\log t dt$.

3. To demonstrate a situation where "passing to a sublattice" is needed, let $u(n)=\{\!\sqrt{3}n\big[\frac{n}{2}\big]\}\!\}$. Then $u(n)=\{\!\frac{\sqrt{3}}{2}n^2\}\!\}$ for $n\in 2\mathbb{Z}$ and

$$u(n)=\{\!\tfrac{\sqrt{3}}{2}n^2-\tfrac{\sqrt{3}}{2}n\}\!\}=\begin{cases}\{\!\frac{\sqrt{3}}{2}n^2\}\!\}-\{\!\frac{\sqrt{3}}{2}n\}\!\} \text{ if } \{\!\frac{\sqrt{3}}{2}n^2\}\!\}-\{\!\frac{\sqrt{3}}{2}n\}\!\}\geq 0\\ \{\!\frac{\sqrt{3}}{2}n^2\}\!\}-\{\!\frac{\sqrt{3}}{2}n\}\!\}+1 \text{ if } \{\!\frac{\sqrt{3}}{2}n^2\}\!\}-\{\!\frac{\sqrt{3}}{2}n\}\!\}<0\end{cases}$$

for $n\in 2\mathbb{Z}+1$.

4. To demonstrate a situation where "passing to $M^{-1}\mathcal{A}$" is needed, let $u=\{\!\sqrt{2}n\{\!\sqrt{2}n\}\!\}\!\}$. We take $M=2$, and $p=2^{-1}(\sqrt{2}n)$. Then

$$u=\{\!2p\{\!2p\}\!\}\!\}=\begin{cases}\{\!4p\{\!p\}\!\}\!\} \text{ if } \{\!p\}\!\}<\frac{1}{2}\\ \{\!4p\{\!p\}\!\}-2p\}\!\} \text{ if } \{\!p\}\!\}\geq\frac{1}{2}\end{cases}=\begin{cases}\{\!2\{\!p\}\!\}^2\}\!\} \text{ if } \{\!p\}\!\}<\frac{1}{2}\\ \{\!2\{\!p\}\!\}^2-2\{\!p\}\!\}\}\!\} \text{ if } \{\!p\}\!\}\geq\frac{1}{2}\end{cases}=\begin{cases}2\{\!p\}\!\}^2 \text{ if } \{\!p\}\!\}<\frac{1}{2}\\ 2\{\!p\}\!\}^2-2\{\!p\}\!\}+1 \text{ if } \{\!p\}\!\}\geq\frac{1}{2},\end{cases}$$

(where we used the identity $\{\!2\sqrt{2}n\{\!\sqrt{2}n\}\!\}\!\}=\{\!-[\sqrt{2}n]^2+2n^2+\{\!\sqrt{2}n\}\!\}^2\}\!\}=\{\!\sqrt{2}n\}\!\}^2$). It follows that the values of u are well-distributed in $[0,1]$ with respect to the measure that is equal to $\frac{dx}{2\sqrt{2x}}$ on $\big[0,\frac{1}{2}\big]$ and to $\frac{dx}{2\sqrt{2x-1}}$ on $\big[\frac{1}{2},1\big]$.

In our proof of Theorems 0.1 and 0.2 we exploit the mentioned above connection between bounded gen-polynomials and nilmanifolds. Actually, the existence of a natural family of basic gen-polynomials satisfying the assertions of these theorems follows from a general argument, – these are the gen-polynomials "read off" from distinct coordinates of a free nilmanifold (see Proposition 2.5). The problem is that these "natural" basic gen-polynomials are too complicated to be explicitly written out and used; our task is therefore to "simplify" them so that they would not loose their independence. As for Theorem 0.2, its first, "existence" part is proved directly, by describing an algorithm of reducing a gen-polynomial to the desired form, and its second, "uniqueness" part is a simple corollary of Theorem 0.1.

We focus on gen-polynomials of integer variables. Dealing with gen-polynomials of real variables is easier, since the closure of the orbit of any point in a nilmanifold under a continuous flow is always connected, whereas such a closure under the action of a discrete group may be disconnected; or, without referring to nilmanifolds, since the gen-polynomial $\{\!q(x)\}\!\}$, $x\in\mathbb{R}^d$, is equidistributed in $[0,1]$ even if a nonconstant polynomial q has all rational coefficients. For the case of gen-polynomials of real variables Theorems 0.1 and 0.2 can be adapted in the following way. Let ω be the Lebesgue measure on $\mathbb{R}^d$. *A Følner sequence* in $\mathbb{R}^d$

is a sequence (Φ_N) of measurable subsets of $\mathbb{R}^d$ such that for every $n \in \mathbb{R}^d$, $\frac{1}{\omega(\Phi_N)}\omega((\Phi_N+n)\triangle\Phi_N) \longrightarrow 0$ as $N \longrightarrow \infty$. We say that the values of a function $\varphi\colon \mathbb{R}^d \longrightarrow [0,1]^k$ are well distributed in $[0,1]^k$ if for any $h \in C\big([0,1]^k\big)$ and any Følner sequence (Φ_N) in $\mathbb{R}^d$ we have $\lim_{N\to\infty} \frac{1}{\omega(\Phi_N)} \int_{\Phi_N} h(\varphi(t))\,dt = \int_{[0,1]^k} h(x)\,dx$, and we say that a family u_α, $\alpha \in \mathcal{A}$, of functions $\mathbb{R}^d \longrightarrow [0,1]$ is jointly well distributed if for any finite subsystem $\{\alpha_1,\dots,\alpha_k\} \subseteq \mathcal{A}$ the values of the function $(u_{\alpha_1},\dots,u_{\alpha_k})$ are well distributed in $[0,1]^k$.

Theorem 0.3. *Let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$, be a well-ordered system of real-valued polynomials on $\mathbb{R}^d$, $\mathbb{Q}$-linearly independent modulo $\mathbb{R}$. Then the gen-polynomials $\{\!\{v_\alpha(\mathcal{P})\}\!\}$, $\alpha \in \mathcal{B}(\mathcal{A})$, are jointly well distributed.*

In the formulation of Theorem 0.2 there is no more need in passing to sublattices:

Theorem 0.4. *Let u be a bounded gen-polynomial on $\mathbb{R}^d$, and let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ be a well-ordered system of polynomials spanning the $\mathbb{Q}$-algebra generated by the polynomials occurring in u. There exist $M \in \mathbb{N}$, distinct indices $\alpha_1,\dots,\alpha_l \in \mathcal{B}(\mathcal{A})$, and a pp-function f on $[0,1]^l$ such that $u = f\big(\{\!\{v_{\alpha_j}(M^{-1}\mathcal{P})\}\!\} : j = 1,\dots,l\big)$. If the system $\mathcal{P}$ is $\mathbb{Q}$-linearly independent modulo $\mathbb{R}$ and M is fixed, then f is defined uniquely up to a (piecewise polynomial) function vanishing a.e. in $[0,1]^l$.*

Here is the plan of the paper. In Section 1 we discuss a relation between gen-polynomials and pp-function. In Section 2 we discuss relations between bounded gen-polynomials and nilmanifolds. In Section 3 we introduce certain technical characteristics of gen-polynomials. In Section 4 we recall the commutator calculus in free groups and introduce what we call “free coordinate gen-polynomials”. In Section 5 we establish the identities, involving the operation $\{\!\{\cdot\}\!\}$, that we will use to transform gen-polynomials. In Section 6 we prove the first part of Theorem 0.2 and describe an algorithm of reduction of any bounded gen-polynomial to its canonical form. In Section 7 we prove Theorem 0.1 and the second part of Theorem 0.2. In Section 8 we describe how the limiting distribution of values of one or several bounded gen-polynomials can be computed. In Section 9 we very briefly discuss two special classes of bounded gen-polynomials. Finally, in Section 10 and Section 11 we, also very briefly, discuss unbounded gen-polynomials and gen-polynomials of real arguments.

1. Gen-polynomials and pp-functions

We call a function f on a set $Q \subseteq \mathbb{R}^m$ *piecewise polynomial*, or *a pp-function*, if Q can be partitioned, $Q = Q_1 \cup \ldots \cup Q_k$, into finitely many subsets so that for each $j = 1,\dots,k$ the set Q_j is defined by a system of polynomial inequalities,

$$Q_j = \big\{x \in Q : \varphi_{j,1}(x),\dots,\varphi_{j,s_j}(x) > 0,\ \psi_{j,1}(x),\dots,\psi_{j,r_j}(x) \geq 0\big\} \tag{1.1}$$

where $\varphi_{j,i}$, $\psi_{j,i}$ are polynomials, and $f_{|Q_j}$ is also a polynomial. We will call the polynomials $\varphi_{j,i}$, $\psi_{j,i}$ *the conditions* of f and the polynomials $f_{|Q_j}$ *the variants* of f. Clearly, pp-functions on a set Q form an algebra, and a composition of pp-functions is a pp-function.

Since the function $\{\!\{x\}\!\}$ is piecewise polynomial (or rather, piecewise linear) on any bounded subset of $\mathbb{R}$, and since gen-polynomials are bounded on bounded sets, it is clear that the restriction of any gen-polynomial on a bounded set $Q \subset \mathbb{R}^m$ is a pp-function. The converse is also true:

Lemma 1.1. *Any pp-function $f\colon Q \longrightarrow \mathbb{R}$ on a bounded subset $Q \subset \mathbb{R}^m$ is the restriction on Q of a gen-polynomial $u\colon \mathbb{R}^k \longrightarrow \mathbb{R}$.*

Proof. Let $Q = Q_1 \cup \ldots \cup Q_k$ with Q_j defined by (1.1) so that $f_{|Q_j} = f_j$ is a polynomial. Let M be such that $|\varphi_{j,i}(x)|, |\psi_{j,i}(x)| < M$ for all j, i and all $x \in Q$. For any number t with $|t| < M$ one has $-[-t/M] = \begin{cases} 1, & t > 0 \\ 0, & t \leq 0 \end{cases}$ and $1 + [t/M] = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$. Thus, if we define a gen-polynomial u as

$$u = \sum_{j=1}^{k} \Big(\prod_{i=1}^{s_j} \big(-[-\varphi_{j,i}/M]\big)\Big)\Big(\prod_{i=1}^{r_j}\big(1 + [\psi_{j,i}/M]\big)\Big) f_j,$$

then for any $j = 1,\dots,k$ and $x \in Q_j$ we get $u(x) = f_j(x)$. ∎

It follows that the composition $f\circ u$ of a bounded gen-polynomial u and of a pp-function f is a gen-polynomial.

2. Nilmanifolds and coordinate gen-polynomials

In this section we will describe relations between bounded gen-polynomials and nilmanifolds, and explain why "basic gen-polynomials" exist.

We first list some facts about nilpotent Lie groups and nilmanifolds; for more details see [Ma]. Let G be a c-step nilpotent Lie group and let Γ be a discrete cocompact subgroup of G; we will assume that G is connected and simply-connected, which will suffice for our goals. The compact homogeneous space $X=G/\Gamma$ is called *a c-step nilmanifold*.

For any $a\in G$ there exists a unique one-parameter subgroup $\{a^t\}_{t\in\mathbb{R}}$ in G such that $a^1=a$; we are therefore allowed to raise elements of G to non-integer powers.

If $a_1,\dots,a_l\in G$ and $p_1,\dots,p_l\colon\mathbb{Z}^d\longrightarrow\mathbb{R}$ are polynomials, we call the (d-parameter) sequence $g(n)=a_1^{p_1(n)}\dots a_l^{p_l(n)}$, $n\in\mathbb{Z}^d$, *a polynomial sequence* in G.

We will denote by π the natural projection $G\longrightarrow X$. A (connected) *subnilmanifold* of X is a closed subset of X of the form $\pi(bH)$, where H is a closed connected subgroup of G and $b\in G$. The following theorem was proved in [L3]:

Theorem 2.1. *Let $g\colon\mathbb{Z}^d\longrightarrow G$ be a polynomial sequence. Then the closure of the range $\pi(g(\mathbb{Z}^d))$ of g is a finite union of subnilmanifolds $Y_1,\dots,Y_s$ of X, and there is a cofinite sublattice $\Lambda\subseteq\mathbb{Z}^d$ such that for any $n_0\in\mathbb{Z}^d$, $\overline{\pi(g(n_0+\Lambda))}=Y_i$ for some i; moreover, the sequence $g|_{n_0+\Lambda}$ is well distributed in Y_i with respect to the Haar measure.*

Let $G=G_1\supseteq G_2\supseteq\dots\supseteq G_c\supseteq G_{c+1}=\{\mathbf{1}_G\}$ be the lower central series of G, that is, $G_{j+1}=[G_j,G]$ for all j. The factor-space $T=G_2\backslash X$ is a torus, called *the maximal factor-torus* of X. The following theorem is also proved in [L3]:

Theorem 2.2. *Let $g\colon\mathbb{Z}^d\longrightarrow G$ be a polynomial sequence. Then the sequence $\pi(g(n))$, $n\in\mathbb{Z}^d$, is well distributed in X if and only if the projection of this sequence to T is dense (and then, well distributed) in T.*

For each j, G_j/G_{j+1} is a connected simply-connected commutative Lie group, and so, a finite dimensional $\mathbb{R}$-vector space. Put $r_0=0$, and for each $j=1,\dots,c$ choose elements $e_{r_{j-1}+1},\dots,e_{r_j}\in\Gamma\cap G_j$ whose classes modulo G_{j+1} generate $(\Gamma\cap G_j)/(\Gamma\cap G_{j+1})$ and form a basis in G_j/G_{j+1}; put $r=r_c$. The set $\{e_1,\dots,e_r\}$ is called *a Malcev basis* of G (compatible with Γ); it has the property that every element $a\in G$ is uniquely representable in the form $a=e_1^{x_1}\dots e_r^{x_r}$, where *the coordinates* $x_1,\dots,x_r$ of a are real numbers, and one has $a\in\Gamma$ iff $x_1,\dots,x_r\in\mathbb{Z}$. The *coordinate mapping* $\eta(a)=(x_1,\dots,x_r)$ is a diffeomorphism $G\longrightarrow\mathbb{R}^r$, with $\eta(\Gamma)=\mathbb{Z}^r$. The image of the Haar measure on G under η is, up to scaling, the Lebesgue measure on $\mathbb{R}^r$.

In coordinates, the multiplication in G is polynomial: for $a=e_1^{x_1}\dots e_r^{x_r}$ and $b=e_1^{y_1}\dots e_r^{y_r}$, we have $ab=e_1^{x_1+y_1}e_2^{x_2+y_2+P_2(x_1,y_1)}\dots e_i^{x_i+y_i+P_i(x_1,\dots,x_{i-1},y_1,\dots,y_{i-1})}\dots e_r^{x_r+y_r+P_r(x_1,\dots,x_{r-1},y_1,\dots,y_{r-1})}$. (Notice that $P_1=\dots=P_{r_1}=0$, since G/G_2 is a vector space.) We can also make an estimate of the degrees of the polynomials P_i, but we first need to introduce a notion of the "weighted degree". Given a monomial $w=x_1^{m_1}\dots x_r^{m_r}$ and nonnegative integers $d_1,\dots,d_r$, we will call the integer $\sum_{i=1}^r m_id_i$ *the weighted degree of w assuming x_i has degree d_i*, $i=1,\dots,r$; the weighted degree of a polynomial is the maximum of the weighted degrees of its monomials. Returning to the polynomial mappings that define the multiplication in G in coordinates, assume that for each j and each $i\in[r_{j-1}+1,r_j]$ the degree of x_i and of y_i is j; then for any $j\in\{1,\dots,c\}$ and any $l\in[r_{j-1}+1,r_j]$, the weighted degree of the polynomial P_l is $\leq j$. (See [L1].)

"The cube" $Q=\eta^{-1}\bigl([0,1)^r\bigr)\subset G$ is the fundamental domain for X, which means that for any $a\in G$ there exists a unique $\gamma\in\Gamma$ such that $\eta(a\gamma)\in[0,1)^r$. Indeed, put $\gamma_0=\mathbf{1}_G$, and if $\gamma_{i-1}\in\Gamma$ is such that $a\gamma_{i-1}=e_1^{u_1}\dots e_{i-1}^{u_{i-1}}e_i^{v_i}\dots e_r^{v_r}$ with $u_1,\dots,u_{i-1}\in[0,1)$, put $\gamma_i=\gamma_{i-1}e_i^{-[v_i]}$. Then $a\gamma_i=a\gamma_{i-1}e_i^{-[v_i]}=e_1^{u_1}\dots e_{i-1}^{u_{i-1}}e_i^{u_i}e_{i+1}^{v'_{i+1}}\dots e_r^{v'_r}$, where $u_i=\{\!\{v_i\}\!\}\in[0,1)$. Applying this operation r times, we get $\gamma=\gamma_r\in\Gamma$ such that $a\gamma=e_1^{u_1}\dots e_r^{u_r}$ with all $u_1,\dots,u_r\in[0,1)$. On the other hand, for any $\delta\in\Gamma$ we have $\eta(\delta)\in\mathbb{Z}^r$, so, if $\delta=e_i^{w_i}\dots e_r^{w_r}$ with $w_i\neq0$, then $a\gamma\delta=e_1^{u_1}\dots e_i^{u_i+w_i}\dots e_r^{v_r}$, and, since $u_i\in[0,1)$ and $w_i\in\mathbb{Z}$, the i-th coordinate u_i+w_i of $a\gamma\delta$ does not belong to $[0,1)$. Thus, $\pi|_Q$ is a bijection between Q and X; we call the mapping $\tau=\eta\circ\pi|_Q{}^{-1}\colon X\longrightarrow[0,1)^r$ *a coordinate mapping* on X. τ is a diffeomorphism on $\tau^{-1}\bigl((0,1)^r\bigr)$, but is discontinuous at the points of $\tau^{-1}\bigl([0,1)^r\setminus(0,1)^r\bigr)$. It is easy to see that the image under τ of the normalized Haar measure on X coincides with the Lebesgue measure on $[0,1)^r$.

For $a=e_1^{x_1}\dots e_r^{x_r}\in G$, let $\gamma\in\Gamma$ be such that $a\gamma\in Q$, and let $u_1,\dots,u_r$ be the coordinates of $a\gamma$, considered as functions of the coordinates $x_1,\dots,x_r$ of a; the mapping $\hat\pi(x_1,\dots,x_r)=\bigl(u_1(x_1,\dots,x_r),\dots,u_r(x_1,\dots,x_r)\bigr)$ from $\mathbb{R}^r$ to $[0,1)^r$ is the coordinate representation of the projection π from G to X. It is easy to see, by induction on i, that the functions $u_1,\dots,u_r$ are gen-polynomials; we will call them *coordinate gen-polynomials* (associated with G, Γ, and a Malcev basis in G).

Example. For the Heisenberg group $G=\mathbb{R}^3$ with multiplication given by $(x_1,x_2,x_3)(y_1,y_2,y_3)=\bigl(x_1+y_1,x_2+y_2,x_3+y_3+x_1y_2\bigr)$, $\Gamma=\mathbb{Z}^3\subset G$, and the standard Malcev basis $\bigl\{(1,0,0),(0,1,0),(0,0,1)\bigr\}$, the coordinate gen-polynomials are $u_1=\{\!\{x_1\}\!\}$, $u_2=\{\!\{x_2\}\!\}$, and $u_3=\bigl\{\!\bigl\{x_3-x_1[x_2]\bigr\}\!\bigr\}$.

Now let $g(n)$, $n\in\mathbb{Z}^d$, be a polynomial sequence in G. Then g can be "written in coordinates" as $g(n)=e_1^{p_1(n)}\dots e_r^{p_r(n)}$, $n\in\mathbb{Z}^d$, where p_i are polynomials $\mathbb{Z}^d\longrightarrow\mathbb{R}$. Taking the ith coordinate $\tau_i(\pi(g(n))$ of the sequence $\pi(g(n))$ in X, we get the gen-polynomial $u_i\bigl(p_1(n),\dots,p_r(n)\bigr)$, $n\in\mathbb{Z}^d$, where u_i is the i-th coordinate gen-polynomial of G. It is proven in [BL] that *any* bounded gen-polynomial can be obtained this way:

Theorem 2.3. *Let u be a gen-polynomial $\mathbb{Z}^d\longrightarrow[0,1)$. There exists a nilmanifold $X=G/\Gamma$, a polynomial sequence $g\colon\mathbb{Z}^d\longrightarrow G$, and a coordinate τ_i on X such that $u(n)=\tau_i(\pi(g(n)))$. Thus, there exists a coordinate gen-polynomial u_i and polynomials $p_1,\dots,p_r\colon\mathbb{Z}^d\longrightarrow\mathbb{R}$ such that $u(n)=u_i\bigl(p_1(n),\dots,p_r(n)\bigr)$, $n\in\mathbb{Z}^d$.*

(We will not use this result in this paper.)

Now let $u_1,\dots,u_r$ be the coordinate gen-polynomials associated with G, Γ, and a Malcev basis in G compatible with Γ. Let $r_1=\dim G/G_2$, then $u_i(x_1,\dots,x_r)=\{\!\{x_i\}\!\}$, $1\le i\le r_1$. Let $g(n)=e_1^{p_1(n)}\dots e_r^{p_r(n)}$, $n\in\mathbb{Z}^d$, be a polynomial sequence in G. The projection of $g(n)$ to the maximal torus $T\simeq\mathbb{R}^{r_1}/\mathbb{Z}^{r_1}$ of X is the sequence $\mathbf{t}(n)=\bigl(\{\!\{p_1\}\!\},\dots,\{\!\{p_{r_1}\}\!\}\bigr)$. If the polynomials $p_1,\dots,p_{r_1}$ are linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$, then $\mathbf{t}(n)$ is well distributed in T, thus, by Theorem 2.2, the sequence $\pi(g(n))$ is well distributed in X, and so, the sequence $\mathbf{u}(n)=\bigl(u_1(p_1(n),\dots,p_r(n)),\dots,u_r(p_1(n),\dots,p_r(n))\bigr)$ is well distributed in $[0,1)^r$ (with respect to the Lebesgue measure). We obtain the following fact:

Proposition 2.4. *Let $u_1,\dots,u_r$ be coordinate gen-polynomials associated with a nilpotent Lie group G, a discrete cocompact subgroup Γ of G, and a Malcev basis in G, and let $r_1=\dim G/G_2$. Let $p_1,\dots,p_r$ be a system of polynomials on $\mathbb{Z}^d$ such that the polynomials $p_1,\dots,p_{r_1}$ are linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$. Then the gen-polynomials $u_1(p_1(n),\dots,p_r(n))$, $\dots$, $u_r(p_1(n),\dots,p_r(n))$ are jointly well distributed.*

(A similar argument was used in [Hå1].)

Let $\mathcal{F}$ be the free product of m copies of $\mathbb{R}$, and let $\mathcal{F}=\mathcal{F}_1\supset\mathcal{F}_2\supset\dots$ be the lower central series of $\mathcal{F}$. Let $c\in\mathbb{N}$, and $F=\mathcal{F}/\mathcal{F}_{c+1}$; then F is a c-step nilpotent Lie group. We will refer to F as to *the free c-step nilpotent Lie group* (with m generators); clearly, any connected c-step nilpotent Lie group with m generators is a factor-group of F. The image Δ in F of the free product of m copies of $\mathbb{Z}$ is a cocompact subgroup of F; we call the nilmanifold $W=F/\Delta$ *the free c-step nilmanifold* (with m generators). Clearly, any connected nilmanifold $X=G/\Gamma$ is a factor of a free nilmanifold $W=F/\Delta$. We will call the coordinate gen-polynomials associated with a free nilpotent group *free coordinate gen-polynomials.*

Homomorphisms of nilpotent Lie groups are given, in coordinates, by polynomial mappings: if $\varphi\colon F\longrightarrow G$ is such a homomorphism and $\eta_F\colon F\longrightarrow\mathbb{R}^s$ and $\eta\colon G\longrightarrow\mathbb{R}^r$ are coordinate mappings, then the composition $\hat\varphi\colon\eta\circ\varphi\circ\eta_F^{-1}\colon\mathbb{R}^s\longrightarrow\mathbb{R}^r$ is a polynomial mapping. If Δ and Γ are discrete cocompact subgroups of F and G respectively with $\varphi(\Delta)\subseteq\Gamma$, then φ induces a mapping ψ from the nilmanifold $W=F/\Delta$ to the nilmanifold $X=G/\Gamma$, which we also call a *homomorphism.* If ψ is surjective, it turns X into a factor of W. In coordinates, ψ is given by gen-polynomials; indeed, if τ_W and τ are coordinate mappings of W and X respectively, then the mapping $\hat\psi=\tau\circ\psi\circ\tau_W^{-1}$ from $[0,1)^s$ to $[0,1)^r$ can be represented as the composition of the restriction on $[0,1)^s$ of the polynomial mapping $\hat\varphi\colon\mathbb{R}^s\longrightarrow\mathbb{R}^r$ and of the "projection" mapping $\hat\pi\colon\mathbb{R}^r\longrightarrow[0,1)^r$ given by gen-polynomials. Since $\hat\psi$ is defined on a bounded set, it is a pp-function.

Now let u be a bounded generalized polynomial. By Theorem 2.3, there exists a nilmanifold $X=G/\Gamma$ with projections $\pi\colon G\longrightarrow X$, a polynomial sequence $g(n)$, $n\in\mathbb{Z}^d$, in G, and a coordinate τ_i on X such that $u(n)=\tau_i(\pi(g(n)))$, $n\in\mathbb{Z}^d$. Let $W=F/\Delta$ be a free nilmanifold, for which surjective homomorphisms $\varphi\colon F\longrightarrow G$ and $\psi\colon W\longrightarrow X$ commuting with the projections $\pi_F\colon F\longrightarrow W$ and $\pi\colon G\longrightarrow X$ exist:

$$\begin{array}{ccccc} F\xrightarrow{\varphi}G & & \xrightarrow{\eta_F,\eta} & & \mathbb{R}^s\xrightarrow{\hat\varphi}\mathbb{R}^r \\ {\scriptstyle\pi_F}\downarrow\quad\ \downarrow{\scriptstyle\pi} & & & & {\scriptstyle\hat\pi_F}\downarrow\qquad\downarrow{\scriptstyle\hat\pi} \\ W\xrightarrow{\psi}X & & \xrightarrow{\tau_F,\tau} & & [0,1)^s\xrightarrow{\hat\psi}[0,1)^r \end{array}$$

Lift g to a polynomial sequence h in F, so that $g(n)=\varphi(h(n))$, and so, $\pi(g(n))=\psi(\pi_F(h(n)))$, $n\in\mathbb{Z}^d$. Let $\eta_F(h(n))=(p_1(n),\ldots,p_s(n))$, $n\in\mathbb{Z}^d$. Then

$$u(n)=\tau_i(\pi(g(n))=\tau_i(\psi(\pi_F(h(n))))=\hat\psi_i(\hat\pi_F(p_1,\ldots,p_s))=\hat\psi_i\big(u_1(p_1,\ldots,p_s),\ldots,u_s(p_1,\ldots,p_s)\big),\ n\in\mathbb{Z}^d,$$

where $u_1,\ldots,u_s$ are free coordinate gen-polynomials. Since $\hat\psi_i$ is a pp-function, we obtain:

Proposition 2.5. *Any bounded gen-polynomial u on $\mathbb{Z}^d$ can be written in the form*

$$u(n)=f\big(u_1(p_1(n),\ldots,p_s(n)),\ldots,u_s(p_1(n),\ldots,p_s(n))\big),\ n\in\mathbb{Z}^d,$$

where $u_1,\ldots,u_s$ are free coordinate gen-polynomials, $p_1,\ldots,p_s$ are polynomials, and f is a pp-function.

Notice that we derived this proposition from results in [BL]; we will reprove this proposition below (see Theorem 7.2) using different methods.

So, we see that the free coordinate gen-polynomials can play the role of basic gen-polynomials. However, it is hard to deal with these gen-polynomials because they are too cumbersome, and it is practically impossible to write them out explicitly; what we will use as basic gen-polynomials are, roughly speaking, the "principal monomials" of free coordinate gen-polynomials. Our goal is to show that these "monomials" also satisfy Proposition 2.4 and Proposition 2.5.

3. Gen-monomials and the complexity of gen-polynomials

In this section we will introduce some technical notions and notation related to gen-polynomials and used below. Actually, under the term "generalized polynomial" we understand two different things: it is either a function, or a formal expression, "a word", constructed with the help of the "$\{\cdot\}$" brackets and the "+" and "−" signs from symbols standing for ordinary polynomials and constants, like $\lambda_1p_1\{\lambda_2p_2\}\{p_3\{p_4\}+\lambda_3\{p_1\}\}$. (We will also use, for shortening, the raising-to-a-power expressions, like w^k.) We should distinguish between gen-polynomials as "functions" and gen-polynomials as formal expressions, but this would make our exposition unreasonably formalistic. We will call a gen-polynomial-formal expression that represents a gen-polynomial-function u *a representation* of u. In Section 5 below we will establish certain "identities" that will allow us to change a representation of u while preserving u as a function, and, in Sections 6 and 7, we will use these identities to reduce the representation of u to a desired form.

The parameters of a gen-polynomial u that we introduce below are related to representations of u; when we treat u as a function, we assume that a representation of u has been fixed.

We will call a gen-polynomial u *closed* if it has the form $u=\{u'\}$ for some gen-polynomial u'.

We will call *gen-monomials* the gen-polynomials constructed from ordinary polynomials without using the operation of addition, that is, the gen-polynomials whose formal expression does not contain the "+" and "−" signs.

Example. $p_1\{p_2\}\{p_3\{2p_4\}\}$, $\{p_1\}\{p_2\}\{p_3\{2p_4\}\}$, $\{p_1\{p_2\}\{p_3\{2p_4\}\}\}$ are gen-monomials, $p_1\{p_2+\{p_3\}p_4\}$ is not.

Sub-gen-monomials of a gen-polynomial u are the gen-monomials that can be "extracted" from u. Formally, we define two sets of gen-monomials, $\mathrm{Mon}(u)$ and $\mathrm{Moc}(u)$, with $\mathrm{Moc}(u) \subseteq \mathrm{Mon}(u)$, inductively on the construction of u:

$\mathrm{Mon}(u) = u$ and $\mathrm{Moc}(u) = \emptyset$ if u is a polynomial;
$\mathrm{Mon}(\{u\}) = \{\mathrm{Mon}(u)\}$ and $\mathrm{Moc}(\{u\}) = \{\mathrm{Mon}(u)\}$;
$\mathrm{Mon}(u_1 + u_2) = \mathrm{Mon}(u_1) \cup \mathrm{Mon}(u_2)$ and $\mathrm{Moc}(u_1 + u_2) = \mathrm{Moc}(u_1) \cup \mathrm{Moc}(u_2)$;
$\mathrm{Mon}(u_1 u_2) = \mathrm{Mon}(u_1)\,\mathrm{Mon}(u_2) \cup \mathrm{Mon}(u_1) \cup \mathrm{Mon}(u_2)$ and $\mathrm{Moc}(u_1 u_2) = \mathrm{Moc}(u_1) \cup \mathrm{Moc}(u_2)$.

We will call the elements of $\mathrm{Mon}(u)$ *sub-gen-monomials of* u. Notice that the sub-gen-monomials of u from $\mathrm{Moc}(u)$ are always closed, and the sub-gen-monomials of u from $\mathrm{Mon}(u)$ are closed iff u is closed.

Example. The sub-gen-monomials of $u = p_1\{p_2 + p_3\{p_4\}\}\{p_5\} + p_6$ are p_1, $\{p_2\}$, $\{p_3\}$, $\{p_4\}$, $\{p_3\{p_4\}\}$, $p_1\{p_2\}$, $p_1\{p_3\}$, $p_1\{p_4\}$, $p_1\{p_3\{p_4\}\}$, $\{p_5\}$, $\{p_2\}\{p_5\}$, $\{p_3\}\{p_5\}$, $\{p_4\}\{p_5\}$, $\{p_3\{p_4\}\}\{p_5\}$, $p_1\{p_5\}$, $p_1\{p_2\}\{p_5\}$, $p_1\{p_3\}\{p_5\}$, $p_1\{p_4\}\{p_5\}$, $p_1\{p_3\{p_4\}\}\{p_5\}$, and p_6. The set $\mathrm{Moc}(u)$ consists of $\{p_2\}$, $\{p_3\}$, $\{p_4\}$, $\{p_3\{p_4\}\}$, and $\{p_5\}$.

Lemma 3.1. *Any gen-polynomial u is a piecewise linear function of its sub-gen-monomials, whose variants are piecewise linear functions, with integer coefficients, of* $\mathrm{Mon}(u)$ *and conditions are piecewise linear functions of* $\mathrm{Moc}(u)$.

Proof. This is clear by induction on the construction of u. If the assertion of the lemma is true for gen-polynomials u_1 and u_2, then $u_1 + u_2$ is a piecewise linear function whose variants are piecewise linear functions, with integer coefficients, of $\mathrm{Mon}(u_1) \cup \mathrm{Mon}(u_2)$, and conditions are piecewise linear functions of $\mathrm{Moc}(u_1) \cup \mathrm{Moc}(u_2)$; $u_1 u_2$ is a piecewise linear function whose variants are piecewise linear functions, with integer coefficients, of $\mathrm{Mon}(u_1)\,\mathrm{Mon}(u_2) \cup \mathrm{Mon}(u_1) \cup \mathrm{Mon}(u_2)$ and conditions are piecewise linear functions of $\mathrm{Moc}(u_1) \cup \mathrm{Moc}(u_2)$, so the assertion of the lemma holds for $u_1 + u_2$ and for $u_1 u_2$.

For any $f = \lambda_1 f_1 + \ldots + \lambda_k f_k$, with $\lambda_1, \ldots, \lambda_k \in \mathbb{Z}$, we have $\{f\} = \lambda_1\{f_1\} + \ldots + \lambda_k\{f_k\} + C$, where C is an integer-valued piecewise linear function, with integer coefficients, of $\{f_i\}$, $i = 1, \ldots, k$. Thus if the assertion of the lemma holds for a gen-polynomial u, then $\{u\}$ is a piecewise linear function whose variants are piecewise linear functions, with integer coefficients, of $\{\mathrm{Mon}(u)\}$, and whose conditions are piecewise linear functions of $\mathrm{Moc}(u) \cup \{\mathrm{Mon}(u)\}$. ■

The complexity $\mathrm{cmp}(u)$ of a gen-polynomial u is the maximal number of pairs of brackets $\{\}$ appearing in the expressions for the sub-gen-monomials of u; we define it formally in the following way:

$\mathrm{cmp}(u) = 0$ if u is a polynomial;
$\mathrm{cmp}(\{u\}) = \mathrm{cmp}(u) + 1$;
$\mathrm{cmp}(u_1 u_2) = \mathrm{cmp}(u_1) + \mathrm{cmp}(u_2)$;
$\mathrm{cmp}(u_1 + u_2) = \max\{\mathrm{cmp}(u_1), \mathrm{cmp}(u_2)\}$.

Examples. $\mathrm{cmp}(p_1\{p_2\}) = 1$, $\mathrm{cmp}(p_1^2\{p_2\}) = 1$, $\mathrm{cmp}(p_1\{p_2\}^2) = 2$, $\mathrm{cmp}(p_1\{p_2\{p_3\}\}) = 2$, $\mathrm{cmp}\big(p_1\{p_2\{p_3\}\}\{p_4\}\big) = 3$, $\mathrm{cmp}\big(p_1\{p_2\{p_3\} + p_4\}\{p_5\} + \{p_6\}\big) = 3$.

We will call any sub-gen-monomial w of a gen-polynomial u with $\mathrm{cmp}(w) = \mathrm{cmp}(u)$ *a principal sub-gen-monomial* of u.

Example. The only principal sub-gen-monomial of the gen-polynomial $p_1\{p_2 + p_3\{p_4\}\}\{p_5\} + p_6$ is $p_1\{p_3\{p_4\}\}\{p_5\}$.

From the definition of complexity, we get:

Lemma 3.2. *If $u_1, \ldots, u_k$ are gen-polynomials and $u = u_1^{m_1} \ldots u_k^{m_k}$, then $\mathrm{cmp}(u) = \sum_{i=1}^k m_i \,\mathrm{cmp}(u_i)$. Thus, if $q(x_1, \ldots, x_k)$ is a monomial of weighted degree d assuming that the degree of the variable x_i is d_i, $i = 1, \ldots, k$, and u_i are gen-polynomials with $\mathrm{cmp}(u_i) = d_i$, $i = 1, \ldots, k$, then $\mathrm{cmp}(q(u_1, \ldots, u_k)) = d$; if $\mathrm{cmp}(u_1) \le d_1 - l$, $\mathrm{cmp}(u_i) \le d_i$, $i = 2, \ldots, k$, and q has degree m_1 with respect to x_1, then $\mathrm{cmp}(q(u_1, \ldots, u_k)) \le d - m_1 l$.*

The parameter *depth*, $\mathrm{dep}(u)$, of a gen-polynomial u is defined in the following way:

$$\begin{aligned}&\mathrm{dep}(u)=0 \text{ if } u \text{ is a polynomial;}\\ &\mathrm{dep}(\{\!\{u\}\!\})=\mathrm{dep}(u)+1;\\ &\mathrm{dep}(u_1u_2)=\mathrm{dep}(u_1+u_2)=\max\bigl\{\mathrm{dep}(u_1),\mathrm{dep}(u_2)\bigr\}.\end{aligned}$$

Examples. $\mathrm{dep}(p_1\{\!\{p_2\}\!\})=1$, $\mathrm{dep}(p_1\{\!\{p_2\}\!\}^2)=1$, $\mathrm{dep}(p_1\{\!\{p_2\{\!\{p_3\}\!\}\}\!\})=2$, $\mathrm{dep}\bigl(p_1\{\!\{p_2\{\!\{p_3\}\!\}\}\!\}\{\!\{p_4\}\!\}\bigr)=2$, $\mathrm{dep}\bigl(p_1\{\!\{p_2\{\!\{p_3\{\!\{p_4\}\!\}\}\!\}\}\!\}\{\!\{p_5\}\!\}\bigr)=3$, $\mathrm{dep}\bigl(p_1\{\!\{p_2\{\!\{p_3\}\!\}+p_4\}\!\}\{\!\{p_5\}\!\}+\{\!\{p_6\}\!\}\bigr)=2$.

We will say that a gen-polynomial v *occurs* in u if the expression for u contains the expression for v as a subword; we will call such a subword *an occurrence* of v in u.

Let $\mathcal{S}$ be a set of gen-polynomials; we will say that u *is constructed* from elements of $\mathcal{S}$ if u is obtained from elements of $\mathcal{S}$ by applying the operations of addition, multiplication, and $\{\!\{\cdot\}\!\}$.

Examples. Any gen-polynomial is constructed from the (ordinary) polynomials occurring in it; for instance, $u=p_1\{\!\{p_2+p_3\{\!\{p_4\}\!\}\}\!\}\{\!\{p_5\}\!\}+p_6$ is constructed from the polynomials $p_1,\dots,p_6$. Also, this u is constructed from the sub-gen-monomials p_1, p_2, $p_3\{\!\{p_4\}\!\}$, $\{\!\{p_5\}\!\}$, and p_6; or from the gen-polynomials $p_1\{\!\{p_2+p_3\{\!\{p_4\}\!\}\}\!\}$, p_5 and p_6.

4. Commutator calculus in free groups and free coordinate gen-polynomials

In this section, we will find the principal sub-gen-monomials of the free coordinate gen-polynomials. Our computations will be based on commutator calculus in free groups; see [MKS] for details.

Let $\mathcal{A}$ be a well-ordered set. Let $\mathcal{F}$ be the free product of $|\mathcal{A}|$ copies of $\mathbb{R}$, generated by the elements of $\mathcal{A}$. ($\mathcal{F}$ consists of the empty word e, playing the role of a neutral element, and words of the form $\alpha_1^{x_1}\dots\alpha_m^{x_m}$, where $m\in\mathbb{N}$, $\alpha_1,\dots,\alpha_m\in\mathcal{A}$, $x_1,\dots,x_m\in\mathbb{R}$, with the relations $\alpha^0=e$ and $\alpha^x\alpha^y=\alpha^{x+y}$ for all $\alpha\in\mathcal{A}$ and $x,y\in\mathbb{R}$; the operation on $\mathcal{F}$ is concatenation of words.) For $a,b\in\mathcal{F}$ let $[a,b]=[a,1b]=a^{-1}b^{-1}ab$ and $[a,(m+1)b]=[[a,mb],b]$, $m\in\mathbb{N}$.

Let $a,b\in\mathcal{F}$ and $x,t\in\mathbb{R}$. Using commutator calculus and induction, or with the help of the Campbell-Hausdorff formula, one can establish the following formula:

$$b^xa^t=$$
$$a^tb^x\prod_{m\ge1}[b,ma]^{r_m(x,t)}\prod_{m,k\ge1}[[b,ma],kb]^{r'_{m,k}(x,t)}\prod_{\substack{m\ge2\\ l\ge1}}\bigl[[b,ma],l[b,a]\bigr]^{r''_{m,l}(x,t)}\prod_{\substack{m,l\ge1\\ k\ge0}}\bigl[[[b,ma],kb],l[b,a]\bigr]^{r'''_{m,k,l}(x,t)}\dots \tag{4.1}$$

where the expression in the right hand side is the product of "basic commutators" in the variables a,b, that is, elements of the set $\mathcal{B}(\{a,b\})$ defined in Introduction, raised to polynomials degrees. The polynomials $r_{\dots},r'_{\dots},\dots$ appearing in (4.1) have rational coefficients and zero constant term. Moreover, the polynomial $r_{\dots},r'_{\dots},\dots$ arising in the exponent of a commutator expression having k occurrences of a and l occurrences of b (counting each sub-commutator m times if it appears with coefficient m) has degree $\le k$ with respect to t and $\le l$ with respect to x.

Example. The initial piece of (4.1) is

$$b^xa^t=a^tb^x\Bigl(\prod_{m\ge1}[b,ma]^{x\binom{t}{m}}\Bigr)\Bigl(\prod_{m\ge1}\prod_{k\ge1}[[b,ma],kb]^{\binom{x}{k+1}\binom{t}{m}}\Bigr)\bigl[[b,2a],[b,a]\bigr]^{r''_{2,1}(x,t)}\dots\bigl[[[b,a],b],[b,a]\bigr]^{r'''_{1,1,1}(x,t)}\dots \tag{4.2}$$

where $r''_{2,1}(x,t)=\binom{x}{2}t\binom{t}{2}+x\binom{t}{3}$ and $r'''_{1,1,1}(x,t)=t^2\frac{x(x-1)(2x-1)}{6}$.

The set $\mathcal{B} = \mathcal{B}(\mathcal{A})$ of *basic commutators* in $\mathcal{F}$ is defined in the following way: $\mathcal{B}$ is the minimal set that contains all elements of $\mathcal{A}$ and all elements of $\mathcal{F}$ of the form $[\gamma, m\beta]$ with $\beta, \gamma \in \mathcal{B}$, $m \in \mathbb{N}$, such that $\beta < \gamma$ and either $\gamma \in \mathcal{A}$ or $\gamma = [\lambda, k\delta]$ with $\delta < \beta$, and where the order on $\mathcal{B} \setminus \mathcal{A}$ is defined by $\mathcal{A} < (\mathcal{B} \setminus \mathcal{A})$, and $[\gamma_1, m_1\beta_1] < [\gamma_2, m_2\beta_2]$ if $(\beta_1, \gamma_1, m_1) < (\beta_2, \gamma_2, m_2)$ lexicographically. Thus, $\mathcal{B} = \mathcal{A} \cup \bigcup_{\beta \in \mathcal{B}} \mathcal{E}_\beta$ where

$$\mathcal{E}_\beta = \bigcup_{\substack{\gamma \in \mathcal{A} \cup \bigcup_{\delta < \beta} \mathcal{E}_\delta \\ \gamma > \beta}} \bigcup_{m=1}^{\infty} [\gamma, m\beta]$$

Notice that this construction of $\mathcal{B}(\mathcal{A})$ is the same as the construction of the "abstract" set $\mathcal{B}(\mathcal{A})$ in Introduction, we now only define $\mathcal{B}(\mathcal{A})$ as a subset of $\mathcal{F}$. For $\alpha \in \mathcal{B}$ we define $|\alpha|$ inductively: we put $|\alpha| = 1$ for $\alpha \in \mathcal{A}$, and $|\alpha| = |\gamma| + m|\beta|$ for $\alpha = [\gamma, m\beta]$.

Let $\mathcal{F} = \mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \ldots$ be the lower central series of $\mathcal{F}$, that is, $\mathcal{F}_{i+1} = [\mathcal{F}_i, \mathcal{F}]$, $i \in \mathbb{N}$.

Theorem 4.1. (See [MKS].) *For every $i \in \mathbb{N}$, the basic commutators $\alpha \in \mathcal{B}$ with $|\alpha| = i$ form a basis in the vector space $\mathcal{F}_i/\mathcal{F}_{i+1}$. Thus, given $c \in \mathbb{N}$, every element $a \in \mathcal{F}$ is uniquely representable modulo $\mathcal{F}_{c+1}$ as a product $a = \prod_{\substack{\alpha \in \mathcal{B} \\ |\alpha| \leq c}} \alpha^{x_\alpha}$, with $x_\alpha \in \mathbb{R}$, where multiplication is performed in accordance with the order on $\mathcal{B}$.*

(Actually, the fact that any element a of $\mathcal{F}$ is representable modulo $\mathcal{F}_{c+1}$ in this form is clear: let $a = \alpha_1^{x_1} \ldots \alpha_m^{x_m}$, with $\alpha_i \in \mathcal{A}$, and, with the help of formula (4.1), start pushing the elements of $\mathcal{A}$, and then the nascent elements of $\mathcal{B}$, in accordance with the order on $\mathcal{B}$, to their places; then the basic commutators are exactly the elements that arise during this process.)

In the representation $a = \prod_{\alpha \in \mathcal{B}} \alpha^{x_\alpha}$, the numbers x_α play the role of coordinates of a in the basis $\mathcal{B}$ of $\mathcal{F}$; if we factorize $\mathcal{F}$ by $\mathcal{F}_{c+1}$ for some c to get the free c-step nilpotent Lie group $F = \mathcal{F}/\mathcal{F}_{c+1}$, then the set $\mathcal{B}_c = \{\alpha \in \mathcal{B} : |\alpha| \leq c\}$ is a Malcev basis in F, and x_α with $\alpha \in \mathcal{B}_c$ are the coordinates of a in this basis. It these coordinates, the multiplication in $\mathcal{F}$ is given by polynomials: if $a = \prod_{\alpha \in \mathcal{B}} \alpha^{x_\alpha}$ and $b = \prod_{\alpha \in \mathcal{B}} \alpha^{y_\alpha}$, then $ab = \prod_{\alpha \in \mathcal{B}} \alpha^{x_\alpha + y_\alpha + z_\alpha}$, where the exponents z_α are polynomial functions in the variables x_β, y_β with $\beta < \alpha$, and the weighted degree of z_α is $|\alpha|$, assuming the degree of x_β and of y_β is $|\beta|$.

For $\beta \in \mathcal{B}$, define $\mathcal{B}_\beta = \mathcal{A} \cup \bigcup_{\delta < \beta} \mathcal{E}_\delta$ and $\mathcal{B}'_\beta = \bigcup_{\delta > \beta} \mathcal{E}_\delta$. Let $\beta \in \mathcal{B}$ and $t \in \mathbb{R}$. Then, applying formula (4.1) to push β^t to its place in the expression $\big(\prod_{\alpha \in \mathcal{B}} \alpha^{x_\alpha}\big)$, we get $\big(\prod_{\alpha \in \mathcal{B}} \alpha^{x_\alpha}\big)\beta^t = \prod_{\alpha \in \mathcal{B}} \alpha^{y_\alpha}$, where

$$y_\alpha = \begin{cases} x_\alpha \text{ for } \alpha < \beta; \\ x_\beta + t \text{ for } \alpha = \beta; \\ x_\alpha \text{ for } \alpha \in \mathcal{B}_\beta,\ \alpha > \beta; \\ x_\alpha + x_\gamma \binom{t}{m} + \sum_{k=1}^{m-1} x_{[\gamma,k\beta]} \binom{t}{m-k} \text{ for } \alpha \in \mathcal{E}_\beta,\ \alpha = [\gamma, m\beta]; \\ x_\alpha + q_\alpha(t, x_\gamma :\ \beta < \gamma < \alpha) \text{ for } \alpha \in \mathcal{B}'_\beta, \text{ where } q_\alpha \text{ is a polynomial over } \mathbb{Q}. \end{cases} \tag{4.3}$$

Moreover, for any $\alpha \in \mathcal{B}'_\beta$, any principal (that is, of weighted degree $|\alpha|$) monomial of the polynomial q_α either contains at least two factors x_γ with $\gamma \in \mathcal{B}_\beta$, $\gamma > \beta$, or a factor x_γ with $\gamma \in \mathcal{E}_\beta \cup \mathcal{B}'_\beta$ (or both).

We will now compute the principal sub-gen-monomials of the free coordinate gen-polynomials corresponding to the basis $\mathcal{B}$ in $\mathcal{F}$. We start with the product $\prod_{\alpha \in \mathcal{A}} \alpha^{x_\alpha}$ and multiply it from the right by a (uniquely defined) product $\prod_{\alpha \in \mathcal{B}} \alpha^{n_\alpha}$ with $n_\alpha \in \mathbb{Z}$, $\alpha \in \mathcal{B}$, to get $\prod_{\alpha \in \mathcal{B}} \alpha^{\{\!\{u_\alpha\}\!\}}$; the gen-polynomials $\{\!\{u_\alpha\}\!\}$ in variables x_α that we obtain this way are the free coordinate gen-polynomials we introduced earlier. Our goal is to show the following:

Proposition 4.2. *$u_\alpha = x_\alpha$ for $\alpha \in \mathcal{A}$, and for any $\alpha = [\gamma, m\beta] \in \mathcal{B} \setminus \mathcal{A}$, $\operatorname{cmp}(u_\alpha) = |\alpha| - 1$ and $u_\alpha = \frac{1}{m!} u_\gamma \{\!\{u_\beta\}\!\}^m + w_\alpha$ with $\operatorname{cmp}(w_\alpha) \leq |\alpha| - 2$.*

Proof. We will multiply the product $\prod_{\alpha \in \mathcal{A}} \alpha^{x_\alpha}$ from the right successively by elements β^{n_β}, $\beta \in \mathcal{B}$, with $n_\beta \in \mathbb{Z}$. Right before the step β we have the expression $\prod_{\alpha \in \mathcal{B}} \alpha^{u_{\alpha,\beta^-}}$ where we assume by induction that

$u_{\alpha,\beta^-} = \{\!\{u_\alpha\}\!\}$ and $\operatorname{cmp}(\{\!\{u_\alpha\}\!\}) = |\alpha|$ for $\alpha < \beta$;
$u_{\beta,\beta^-} = u_\beta$ and $\operatorname{cmp}(u_\beta) = |\beta| - 1$;
$u_{\alpha,\beta^-} = u_\alpha$ and $\operatorname{cmp}(u_\alpha) = |\alpha| - 1$ for $\alpha \in \mathcal{B}_\beta$, $\alpha > \beta$;
$\operatorname{cmp}(u_{\alpha,\beta^-}) \leq |\alpha| - 2$ for $\alpha \in \mathcal{E}_\beta \cup \mathcal{B}'_\beta$.

Making the step β we multiply this expression from the right by $\beta^{-[u_\beta]} = \beta^{\{\!\{u_\beta\}\!\}-u_\beta}$ and obtain $\prod_{\alpha\in\mathcal{B}} \alpha^{u_{\alpha,\beta}}$; using (4.3), we see that

$$
\begin{aligned}
&u_{\alpha,\beta} = \{\!\{u_\alpha\}\!\} \text{ for } \alpha < \beta;\\
&u_{\beta,\beta} = \{\!\{u_\beta\}\!\};\\
&u_{\alpha,\beta} = u_{\alpha,\beta^-} = u_\alpha \text{ for } \alpha \in \mathcal{B}_\beta,\ \alpha > \beta;\\
&u_{\alpha,\beta} = u_{\alpha,\beta^-} + u_\gamma\tbinom{\{\!\{u_\beta\}\!\}-u_\beta}{m} + \textstyle\sum_{k=1}^{m-1} u_{[\gamma,k\beta]}\tbinom{\{\!\{u_\beta\}\!\}-u_\beta}{m-k} \text{ for } \alpha \in \mathcal{E}_\beta,\ \alpha = [\gamma, m\beta];\\
&u_{\alpha,\beta} = u_{\alpha,\beta^-} + q_{\alpha,\beta} \text{ for } \alpha \in \mathcal{B}'_\beta, \text{ where } q_{\alpha,\beta} \text{ is a polynomial over } \mathbb{Q}\\
&\qquad\qquad\qquad\qquad \text{in the variables } [u_\beta] \text{ and } u_{\gamma,\beta^-},\ \beta < \gamma < \alpha.
\end{aligned}
$$

For $\alpha \in \mathcal{E}_\beta$, $\alpha = [\gamma, m\beta]$, $u_{\alpha,\beta}$ will remain unchanged until the very α step, thus

$$u_\alpha = u_{\alpha,\beta} = u_{\alpha,\beta^-} + u_\gamma\tbinom{\{\!\{u_\beta\}\!\}-u_\beta}{m} + \textstyle\sum_{k=1}^{m-1} u_{[\gamma,k\beta]}\tbinom{\{\!\{u_\beta\}\!\}-u_\beta}{m-k} = \frac{1}{m!}u_\gamma\{\!\{u_\beta\}\!\}^m + w_\alpha,$$

where w_α is the remaining part of u_α. Since, by induction hypothesis, $\mathrm{cmp}(u_{\alpha,\beta^-}) \le |\alpha|-2$, $\mathrm{cmp}(u_\gamma) = |\gamma|-1$, $\mathrm{cmp}(u_{[\gamma,k\beta]}) \le |\gamma|+k|\beta|-2$ for $k \in \mathbb{N}$, and $\mathrm{cmp}(\{\!\{u_\beta\}\!\}) = |\beta|$, we have $\mathrm{cmp}\big(u_\gamma\{\!\{u_\beta\}\!\}^m\big) = |\gamma|-1+m|\beta| = |\alpha|-1$ and the complexity of w_α is strictly less than this.

For $\alpha \in \mathcal{B}'_\beta$, $q_{\alpha,\beta}$ is a polynomial of weighted degree $\le |\alpha|$ in variables u_β and $\{\!\{u_\beta\}\!\}$, assumed to be of degree $|\beta|$, and in variables u_{γ,β^-} with $\beta < \gamma < \alpha$, assumed to be of degree $|\gamma|$; moreover, any principal monomial of the polynomial $q_{\alpha,\beta}$ either contains at least two factors u_{γ,β^-} with $\gamma \in \mathcal{B}_\beta$, $\gamma > \beta$, or a factor u_{γ,β^-} with $\gamma \in \mathcal{E}_\beta \cup \mathcal{B}'_\beta$, or both; any non-principal monomial contains at least one factor u_{γ,β^-} with $\gamma > \beta$. Since $\mathrm{cmp}(u_{\gamma,\beta^-}) \le |\gamma|-1$ for each $\gamma \in \mathcal{B}_\beta$, $\gamma > \beta$, and $\mathrm{cmp}(u_{\gamma,\beta^-}) \le |\gamma|-2$ for each $\gamma \in \mathcal{E}_\beta \cup \mathcal{B}'_\beta$, we get $\mathrm{cmp}(q_{\alpha,\beta}) \le |\alpha|-2$ by Lemma 3.2. Hence, $\mathrm{cmp}(u_{\alpha,\beta}) \le |\alpha|-2$ for $\alpha \in \mathcal{B}'_\beta$, which justifies the induction step. ■

Example. Even the simplest examples are very cumbersome if one tries to make explicit computations. Let us take $\mathcal{A} = \{a,b\}$ and confine ourselves to commutators from $\mathcal{B}_5$ only, ignoring all commutators of depth ≥ 6; that is, let us make our computations in the group $F = \mathcal{F}/\mathcal{F}_6$. The basic commutators of depth ≤ 5 are

a, b, $[b,a]$, $[b,2a]$, $[b,3a]$, $[b,4a]$, $[[b,a],b]$, $[[b,a],2b]$, $[[b,a],3b]$, $[[b,2a],b]$, $[[b,2a],2b]$, $[[b,3a],b]$, $\big[[b,2a],[b,a]\big]$, $\big[[[b,a],b],[b,a]\big]$.

Let $g_0(x,y) = a^x b^y$. (For convenience, we will use the variables x and y instead of x_1 and x_2.)

Step a**.** We multiply g_0 by $a^{\{\!\{x\}\!\}-x}$ and, by (4.2), get

$$
\begin{aligned}
g_a(x,y) = g(0)(x,y)a^{\{\!\{x\}\!\}-x} &= a^{\{\!\{x\}\!\}}b^y[b,a]^{(\{\!\{x\}\!\}-x)y}[b,2a]^{\binom{\{\!\{x\}\!\}-x}{2}y}[b,3a]^{\binom{\{\!\{x\}\!\}-x}{3}y}[b,4a]^{\binom{\{\!\{x\}\!\}-x}{4}y}\\
&\cdot[[b,a],b]^{(\{\!\{x\}\!\}-x)\binom{y}{2}}[[b,a],2b]^{(\{\!\{x\}\!\}-x)\binom{y}{3}}[[b,a],3b]^{(\{\!\{x\}\!\}-x)\binom{y}{4}}[[b,2a],b]^{\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}}[[b,2a],2b]^{\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{3}}\\
&\cdot[[b,3a],b]^{\binom{\{\!\{x\}\!\}-x}{3}\binom{y}{2}}\big[[b,2a],[b,a]\big]^{(\{\!\{x\}\!\}-x)\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}+y\binom{\{\!\{x\}\!\}-x}{3}}\big[[[b,a],b],[b,a]\big]^{(\{\!\{x\}\!\}-x)^2\frac{y(y-1)(2y-1)}{6}}.
\end{aligned}
$$

Step b**.** We now multiply g_a by $b^{\{\!\{y\}\!\}-y}$ and get

$$
\begin{aligned}
g_b(x,y) = g_a(x,y)b^{\{\!\{y\}\!\}-y} &= a^{\{\!\{x\}\!\}}b^{\{\!\{y\}\!\}}[b,a]^{(\{\!\{x\}\!\}-x)y}[[b,a],b]^{(\{\!\{x\}\!\}-x)y(\{\!\{y\}\!\}-y)}[[b,a],2b]^{(\{\!\{x\}\!\}-x)y\binom{\{\!\{y\}\!\}-y}{2}}\\
&\cdot[[b,a],3b]^{(\{\!\{x\}\!\}-x)y\binom{\{\!\{y\}\!\}-y}{3}}\big[[[b,a],b],[b,a]\big]^{\binom{(\{\!\{x\}\!\}-x)y}{2}(\{\!\{y\}\!\}-y)}[b,2a]^{\binom{\{\!\{x\}\!\}-x}{2}y}[[b,2a],b]^{\binom{\{\!\{x\}\!\}-x}{2}y(\{\!\{y\}\!\}-y)}\\
&\cdot[[b,2a],2b]^{\binom{\{\!\{x\}\!\}-x}{2}y\binom{\{\!\{y\}\!\}-y}{2}}[b,3a]^{\binom{\{\!\{x\}\!\}-x}{3}y}[[b,3a],b]^{\binom{\{\!\{x\}\!\}-x}{3}y(\{\!\{y\}\!\}-y)}[b,4a]^{\binom{\{\!\{x\}\!\}-x}{4}y}[[b,a],b]^{(\{\!\{x\}\!\}-x)\binom{y}{2}}\\
&\cdot[[b,a],2b]^{(\{\!\{x\}\!\}-x)\binom{y}{2}(\{\!\{y\}\!\}-y)}[[b,a],3b]^{(\{\!\{x\}\!\}-x)\binom{y}{2}\binom{\{\!\{y\}\!\}-y}{2}}[[b,a],2b]^{(\{\!\{x\}\!\}-x)\binom{y}{3}}[[b,a],3b]^{(\{\!\{x\}\!\}-x)\binom{y}{3}(\{\!\{y\}\!\}-y)}\\
&\cdot[[b,a],3b]^{(\{\!\{x\}\!\}-x)\binom{y}{4}}[[b,2a],b]^{\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}}[[b,2a],2b]^{\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}(\{\!\{y\}\!\}-y)}[[b,2a],2b]^{\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{3}}[[b,3a],b]^{\binom{\{\!\{x\}\!\}-x}{3}\binom{y}{2}}\\
&\cdot\big[[b,2a],[b,a]\big]^{(\{\!\{x\}\!\}-x)\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}+y\binom{\{\!\{x\}\!\}-x}{3}}\big[[[b,a],b],[b,a]\big]^{(\{\!\{x\}\!\}-x)^2\frac{y(y-1)(2y-1)}{6}}.
\end{aligned}
$$

Since in our group F any two commutators of depth ≥ 3 commute, we can "collect similar terms" and rewrite g_b as

$$
\begin{aligned}
g_b(x,y) = g_a(x,y)b^{\{\!\{y\}\!\}-y} &= a^{\{\!\{x\}\!\}}b^{\{\!\{y\}\!\}}[b,a]^{(\{\!\{x\}\!\}-x)y}[b,2a]^{\binom{\{\!\{x\}\!\}-x}{2}y}[b,3a]^{\binom{\{\!\{x\}\!\}-x}{3}y}[b,4a]^{\binom{\{\!\{x\}\!\}-x}{4}y}\\
&\cdot[[b,a],b]^{(\{\!\{x\}\!\}-x)y(\{\!\{y\}\!\}-y)+(\{\!\{x\}\!\}-x)\binom{y}{2}}[[b,a],2b]^{(\{\!\{x\}\!\}-x)y\binom{\{\!\{y\}\!\}-y}{2}+(\{\!\{x\}\!\}-x)\binom{y}{2}(\{\!\{y\}\!\}-y)+(\{\!\{x\}\!\}-x)\binom{y}{3}}\\
&\cdot[[b,a],3b]^{(\{\!\{x\}\!\}-x)y\binom{\{\!\{y\}\!\}-y}{3}+(\{\!\{x\}\!\}-x)\binom{y}{2}\binom{\{\!\{y\}\!\}-y}{2}+(\{\!\{x\}\!\}-x)\binom{y}{3}(\{\!\{y\}\!\}-y)+(\{\!\{x\}\!\}-x)\binom{y}{4}}[[b,2a],b]^{\binom{\{\!\{x\}\!\}-x}{2}y(\{\!\{y\}\!\}-y)+\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}}\\
&\cdot[[b,2a],2b]^{\binom{\{\!\{x\}\!\}-x}{2}y\binom{\{\!\{y\}\!\}-y}{2}+\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}(\{\!\{y\}\!\}-y)+\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{3}}[[b,3a],b]^{\binom{\{\!\{x\}\!\}-x}{3}y(\{\!\{y\}\!\}-y)+\binom{\{\!\{x\}\!\}-x}{3}\binom{y}{2}}\\
&\cdot\big[[b,2a],[b,a]\big]^{(\{\!\{x\}\!\}-x)\binom{\{\!\{x\}\!\}-x}{2}\binom{y}{2}+y\binom{\{\!\{x\}\!\}-x}{3}}\big[[[b,a],b],[b,a]\big]^{(\{\!\{x\}\!\}-x)^2\frac{y(y-1)(2y-1)}{6}+\binom{(\{\!\{x\}\!\}-x)y}{2}(\{\!\{y\}\!\}-y)}.
\end{aligned}
$$

Step $[b,a]$**.** Finally, we multiply g_a by $[b,a]^{\{(\{x\}-x)y\}-(\{x\}-x)y}$ and, after permutation of (commuting in F) terms get

$$
\begin{aligned}
g_b(x,y) = g_a(x,y)b^{\{y\}-y} &= a^{\{x\}}b^{\{y\}}[b,a]^{\{(\{x\}-x)y\}}[b,2a]^{\binom{\{x\}-x}{2}y}[b,3a]^{\binom{\{x\}-x}{3}y}[b,4a]^{\binom{\{x\}-x}{4}y} \\
&\cdot[[b,a],b]^{(\{x\}-x)y(\{y\}-y)+(\{x\}-x)\binom{y}{2}}[[b,a],2b]^{(\{x\}-x)y\binom{\{y\}-y}{2}+(\{x\}-x)\binom{y}{2}(\{y\}-y)+(\{x\}-x)\binom{y}{3}} \\
&\cdot[[b,a],3b]^{(\{x\}-x)y\binom{\{y\}-y}{3}+(\{x\}-x)\binom{y}{2}\binom{\{y\}-y}{2}+(\{x\}-x)\binom{y}{3}(\{y\}-y)+(\{x\}-x)\binom{y}{4}}[[b,2a],b]^{\binom{\{x\}-x}{2}y(\{y\}-y)+\binom{\{x\}-x}{2}\binom{y}{2}} \\
&\cdot[[b,2a],2b]^{\binom{\{x\}-x}{2}y\binom{\{y\}-y}{2}+\binom{\{x\}-x}{2}\binom{y}{2}(\{y\}-y)+\binom{\{x\}-x}{2}\binom{y}{3}}[[b,3a],b]^{\binom{\{x\}-x}{3}y(\{y\}-y)+\binom{\{x\}-x}{3}\binom{y}{2}} \\
&\cdot\big[[b,2a],[b,a]\big]^{(\{x\}-x)\binom{\{x\}-x}{2}\binom{y}{2}+y\binom{\{x\}-x}{3}+\binom{\{x\}-x}{2}y(\{(\{x\}-x)y\}-(\{x\}-x)y)} \\
&\cdot\big[[[b,a],b],[b,a]\big]^{(\{x\}-x)^2\frac{y(y-1)(2y-1)}{6}+\binom{(\{x\}-x)y}{2}(\{y\}-y)+((\{x\}-x)y(\{y\}-y)+(\{x\}-x)\binom{y}{2})(\{(\{x\}-x)y\}-(\{x\}-x)y)}.
\end{aligned}
$$

Further steps do not affect the gen-polynomials that we have already obtained in the exponents (except that they conclude them into the $\{\cdot\}$ brackets), and we obtain that

$$
\begin{aligned}
&u_a = x, \\
&u_b = y, \\
&u_{[b,a]} = (\{x\}-x)y = \{x\}y - xy = u_b\{u_a\} + \dots, \\
&u_{[b,2a]} = \binom{\{x\}-x}{2}y = \tfrac{1}{2}\{x\}^2 y + \dots = \tfrac{1}{2}u_b\{u_a\}^2 + \dots, \\
&u_{[b,3a]} = \binom{\{x\}-x}{3}y = \tfrac{1}{3!}\{x\}^3 y + \dots = \tfrac{1}{3!}u_b\{u_a\}^3 + \dots, \\
&u_{[b,4a]} = \binom{\{x\}-x}{4}y = \tfrac{1}{4!}\{x\}^4 y + \dots = \tfrac{1}{4!}u_b\{u_a\}^4 + \dots, \\
&u_{[[b,a],b]} = (\{x\}-x)y(\{y\}-y) + (\{x\}-x)\binom{y}{2} = (\{x\}-x)y\{y\} + \dots = u_{[b,a]}\{u_b\} + \dots, \\
&u_{[[b,a],2b]} = (\{x\}-x)y\binom{\{y\}-y}{2}+(\{x\}-x)\binom{y}{2}(\{y\}-y)+(\{x\}-x)\binom{y}{3} = \tfrac{1}{2}(\{x\}-x)y\{y\}^2+\dots = \tfrac{1}{2}u_{[b,a]}\{u_b\}^2+\dots, \\
&u_{[[b,a],3b]} = (\{x\}-x)y\binom{\{y\}-y}{3} + (\{x\}-x)\binom{y}{2}\binom{\{y\}-y}{2} + (\{x\}-x)\binom{y}{3}(\{y\}-y) + (\{x\}-x)\binom{y}{4} \\
&\qquad\qquad = \tfrac{1}{3!}(\{x\}-x)y\{y\}^3 + \dots = \tfrac{1}{3!}u_{[b,a]}\{u_b\}^3 + \dots, \\
&u_{[[b,2a],b]} = \binom{\{x\}-x}{2}y(\{y\}-y) + \binom{\{x\}-x}{2}\binom{y}{2} = \binom{\{x\}-x}{2}y\{y\} + \dots = u_{[b,2a]}\{u_b\} + \dots, \\
&u_{[[b,2a],2b]} = \binom{\{x\}-x}{2}y\binom{\{y\}-y}{2} + \binom{\{x\}-x}{2}\binom{y}{2}(\{y\}-y) + \binom{\{x\}-x}{2}\binom{y}{3} = \tfrac{1}{2}\binom{\{x\}-x}{2}y\{y\}^2 + \dots = \tfrac{1}{2}u_{[b,2a]}\{u_b\}^2 + \dots, \\
&u_{[[b,3a],b]} = \binom{\{x\}-x}{3}y(\{y\}-y) + \binom{\{x\}-x}{3}\binom{y}{2} = \binom{\{x\}-x}{3}y\{y\} + \dots = u_{[b,3a]}\{u_b\} + \dots, \\
&u_{[[b,2a],[b,a]]} = (\{x\}-x)\binom{\{x\}-x}{2}\binom{y}{2} + y\binom{\{x\}-x}{3} + \binom{\{x\}-x}{2}y(\{(\{x\}-x)y\} - (\{x\}-x)y) \\
&\qquad\qquad = \binom{\{x\}-x}{2}y\{(\{x\}-x)y\} + \dots = u_{[b,2a]}\{u_{[b,a]}\} + \dots, \\
&u_{[[[b,a],b],[b,a]]} = \\
&(\{x\}-x)^2\,\frac{y(y-1)(2y-1)}{6} + \binom{(\{x\}-x)y}{2}(\{y\}-y) + ((\{x\}-x)y(\{y\}-y) + (\{x\}-x)\binom{y}{2})(\{(\{x\}-x)y\} - (\{x\}-x)y) \\
&\qquad\qquad = ((\{x\}-x)y(\{y\}-y) + (\{x\}-x)\binom{y}{2})\{(\{x\}-x)y\} + \dots = u_{[[b,a],b]}\{u_{[b,a]}\} + \dots,
\end{aligned}
$$

where "…" stands for summands of smaller complexity.

5. Identities

In this section we will treat gen-polynomials as "functions" (see the discussion in the beginning of Section 3).

Let $u, u_1, \dots, u_k$ be any real numbers or functions. In subsequent sections we will pass from one representation of a given gen-polynomial to another with the help of the following identities:

$$
\{u_1+u_2\} = \begin{cases} \{u_1\} + \{u_2\} \text{ if } \{u_1\} + \{u_2\} < 1 \\ \{u_1\} + \{u_2\} - 1 \text{ if } \{u_1\} + \{u_2\} \geq 1 \end{cases} \tag{5.1}
$$

For any positive $\lambda \in \mathbb{R}$,

$$
\{\lambda u\} = \begin{cases} \lambda\{u\} \text{ if } \{u\} < \frac{1}{\lambda} \\ \lambda\{u\} - 1 \text{ if } \frac{1}{\lambda} \leq \{u\} < \frac{2}{\lambda} \\ \qquad\vdots \\ \lambda\{u\} - [\lambda] \text{ if } \frac{[\lambda]}{\lambda} \leq \{u\} \end{cases} \tag{5.2}
$$

and

$$\{-u\} = \begin{cases} 1 - \{u\} \text{ if } \{u\} > 0 \\ 0 \text{ if } \{u\} = 0 \end{cases} \tag{5.3}$$

For any $k \in \mathbb{N}$,

$$\Big\{ \prod_{i=1}^{k} \{u_i\} \Big\} = \prod_{i=1}^{k} \{u_i\} \tag{5.4}$$

Next, for any $k \in \mathbb{N}$, $k \geq 2$,

$$\prod_{i=1}^{k} [u_i] = \prod_{i=1}^{k} (u_i - \{u_i\}) = (-1)^k \prod_{i=1}^{k} \{u_i\} + (-1)^{k-1} \sum_{j=1}^{k} u_j \prod_{i \neq j} \{u_i\} + \sum_{l=2}^{k} (-1)^{k-l} \sum_{\substack{S \subset \{1,\ldots,k\} \\ |S|=l}} \prod_{i \in S} u_i \prod_{i \notin S} \{u_i\} \tag{5.5}$$

Thus,

$$\Big\{ u_1 \prod_{i=2}^{k} \{u_i\} \Big\} = \Big\{ \prod_{i=1}^{k} \{u_i\} - \sum_{j=2}^{k} u_j \prod_{i \neq j} \{u_i\} + \sum_{l=2}^{k} \sum_{\substack{S \subset \{1,\ldots,k\} \\ |S|=l}} q_S \prod_{i \notin S} \{u_i\} \Big\}, \tag{5.6}$$

where, for each S, $q_S = \pm \prod_{i \in S} u_i$. (For $k = 2$ the formula is

$$\{u_1 \{u_2\}\} = \{\{u_1\}\{u_2\} - u_2\{u_1\} + u_1 u_2\},$$

for $k = 3$ it is

$$\{u_1\{u_2\}\{u_3\}\}$$
$$= \{\{u_1\}\{u_2\}\{u_3\} - u_2\{u_1\}\{u_3\} - u_3\{u_1\}\{u_2\} + u_1u_2\{u_3\} + u_1u_3\{u_2\} + u_2u_3\{u_1\} - u_1u_2u_3\}.)$$

For any $m \leq k$, taking $u_1 = u_2 = \ldots = u_m$ in (5.5), we get

$$u_1\{u_1\}^{m-1} \prod_{i=m+1}^{k} \{u_i\} = \frac{1}{m}\Big(\prod_{i=1}^{k} \{u_i\} - \sum_{j=m+1}^{k} u_j \prod_{i \neq j} \{u_i\} + \sum_{l=2}^{k} \sum_{\substack{S \subset \{1,\ldots,k\} \\ |S|=l}} q_S \prod_{i \notin S} \{u_i\} \Big) \pm \frac{1}{m} \prod_{i=1}^{k} [u_i],$$

and so,

$$\Big\{ \lambda u_1 \{u_1\}^{m-1} \prod_{i=m+1}^{k} \{u_i\} \Big\} = \Big\{ \frac{\lambda}{m} \prod_{i=1}^{k} \{u_i\} - \frac{\lambda}{m} \sum_{j=m+1}^{k} u_j \prod_{i \neq j} \{u_i\} + \frac{\lambda}{m} \sum_{l=2}^{k} \sum_{\substack{S \subset \{1,\ldots,k\} \\ |S|=l}} q_S \prod_{i \notin S} \{u_i\} \Big\} \tag{5.7}$$

for any $\lambda \in \mathbb{Z}$ divisible by m.

Proving Theorem 0.2, we will show that with the help of identities (5.1) – (5.7) every gen-polynomial can be reduced to its unique (up to the choice of a basis in a space of polynomials, and up to a gen-polynomial vanishing on a set of zero density) canonical form. This means that, in some sense, all relations between gen-polynomials can be derived from these identities. We do not try to formalize this statement.

6. The canonical representation of bounded gen-polynomials

In this section we will prove the "existence" part of Theorem 0.2, namely, that any bounded gen-polynomial can be represented as a pp-function of the basic ones; we will do this by describing an algorithm that transforms (any representation of) any gen-polynomial into this form. In this and the next section we will consider gen-polynomials both as functions and as formal expressions: given a gen-polynomial u "as a formal expression", using the identities from Section 5 we will change the representation of u "as a function".

Let $\mathcal{A}$ be a well-ordered set and let $\mathcal{B} = \mathcal{B}(\mathcal{A})$. Recall that the gen-polynomials v_α, $\alpha \in \mathcal{B}$, in the variables x_α, $\alpha \in \mathcal{A}$, are defined in the following way: $v_\alpha = x_\alpha$ for $\alpha \in \mathcal{A}$, and for $\alpha = [\gamma, m\beta]$, $v_\alpha = v_\gamma \{\!\{v_\beta\}\!\}^m$. We call the gen-polynomials v_α *(open) basic gen-polynomials*, and the gen-polynomials $\{\!\{v_\alpha\}\!\}$ *(closed) basic gen-polynomials*. By induction, we have $\mathrm{cmp}(v_\alpha) = |\alpha| - 1$.

If $f(x_1, \ldots, x_k)$ is a pp-function with conditions φ_i and variants f_j, and $u_1, \ldots, u_k$ are bounded gen-polynomials, under the complexity, $\mathrm{cmp}\big(f(u_1, \ldots, u_k)\big)$, of $f(u_1, \ldots, u_k)$ we will understand the maximum of $\mathrm{cmp}\big(\varphi_i(u_1, \ldots, u_k)\big)$ and $\mathrm{cmp}\big(f_j(u_1, \ldots, u_k)\big)$ for all i, j. (So defined complexity may differ from the complexity of $f_j(u_1, \ldots, u_k)$ if it is interpreted as a gen-polynomial. The depth, $\mathrm{dep}(f_j(u_1, \ldots, u_k))$, of $f_j(u_1, \ldots, u_k)$ is equal to the maximum of $\mathrm{dep}(u_1), \ldots, \mathrm{dep}(u_k)$, and is the same if $f_j(u_1, \ldots, u_k)$ is interpreted as a gen-polynomial.)

We will say that a statement, depending on M, holds *for M divisible enough* if there exists $M_0 \in \mathbb{N}$ such that the statement holds whenever $M \in \mathbb{N}$ is divisible by M_0.

The first part of Theorem 0.2 follows from the following theorem:

Theorem 6.1. *Let u be a closed gen-polynomial over $\mathbb{Z}^d$. Let $\mathcal{R}$ be the $\mathbb{Q}$-algebra generated by the polynomials occurring in u and let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ be a system of polynomials such that $\mathrm{span}_{\mathbb{Q}} \mathcal{P} + \mathbb{Q}[n] + \mathbb{R} \supseteq \mathcal{R}$. If $M \in \mathbb{N}$ is divisible enough, then there exists a sublattice Λ in $\mathbb{Z}^d$ such that for any translate $\Lambda' = n_0 + \Lambda$ of Λ one has $u|_{\Lambda'} = f\big(\{\!\{v_\alpha(M^{-1}\mathcal{P})\}\!\} : \alpha \in \mathcal{B}\big)|_{\Lambda'}$ where f is a pp-function, $\mathrm{cmp}\big(f\big(\{\!\{v_\alpha\}\!\} : \alpha \in \mathcal{B}\big)\big) \leq \mathrm{cmp}(u)$ and $\mathrm{dep}\big(f\big(\{\!\{v_\alpha\}\!\} : \alpha \in \mathcal{B}\big)\big) \leq \mathrm{dep}(u)$.*

(Under a pp-function of infinitely many arguments we understand a pp-function that only depends on finitely many of these arguments.)

Remarks. 1. Instead of passing to sublattices of $\mathbb{Z}^d$, we could add one more argument to f and write u as $u(n) = f\big(n \bmod \Lambda, \{\!\{v_\alpha(M^{-1}\mathcal{P})\}\!\} : \alpha \in \mathcal{B}\big)$.
2. We, actually, do not need $\mathcal{P}$ to span the whole algebra $\mathcal{R}$; it suffices if $\mathrm{span}_{\mathbb{Q}} \mathcal{P} + \mathbb{Q}[n] + \mathbb{R}$ contains the products of any c polynomials occurring in u, where $c = \mathrm{cmp}(u)$. In particular, we may assume that $\mathcal{P}$ (and so, $\mathcal{A}$) is a finite set.

To deduce Theorem 0.2 from Theorem 6.1, given a non-closed bounded gen-polynomial u, after proper scaling we may assume that $0 \leq u < 1$, and replace u by $\{\!\{u\}\!\}$; this makes u closed but increases $\mathrm{cmp}(u)$ by 1. Another possibility is to notice that any non-closed bounded gen-polynomial u is representable in the form $u = P(u_1, \ldots, u_k) + w$, where P is a polynomial, u_i are closed gen-polynomials, and $w(n)$ is a bounded gen-polynomial vanishing for all n but a set of zero density (see Section 10); thus, if we ignore a set of n of zero density, we may assume that u is closed.

For technical needs, we will now introduce one more notion. Let $\mathbb{Z}^d = \bigcup_{j=1}^k Z_j$ be a finite partition of $\mathbb{Z}^d$ where each Z_j is defined by a finite system of inequalities $w_{j,i} > 0$ or $w_{j,i} \geq 0$, $i = 1, \ldots, l_j$, where $w_{j,i}$ are bounded gen-polynomials, and let φ be a function on $\mathbb{Z}^d$ such that for every j, $\varphi|_{Z_j}$ is a bounded gen-polynomial. We call φ constructed this way *a piecewise gen-polynomial*, and call the polynomials $w_{j,i}$ *the conditions* of φ and the gen-polynomials $\varphi|_{Z_j}$, $j = 1, \ldots, k$, *the variants* of φ. We define $\mathrm{cmp}(\varphi) = \max_j \mathrm{cmp}(\varphi|_{Z_j})$ and $\mathrm{cmc}(\varphi) = \max_{j,i} \mathrm{cmp}(w_{j,i})$. Of course, any piecewise gen-polynomial is a pp-function of its variants and conditions, and so, a gen-polynomial; however, if a piecewise gen-polynomial is interpreted this way, the parameter cmp thereof may be different. It is clear that piecewise gen-polynomials form an algebra, and that a composition of piecewise gen-polynomials is a piecewise gen-polynomial. For two piecewise gen-polynomials φ_1 and φ_2, $\mathrm{cmp}(\varphi_1 + \varphi_2) \leq \max\{\mathrm{cmp}(\varphi_1), \mathrm{cmp}(\varphi_2)\}$, $\mathrm{cmp}(\varphi_1 \varphi_2) \leq \mathrm{cmp}(\varphi_1) + \mathrm{cmp}(\varphi_2)$, and $\mathrm{cmc}(\varphi_1 + \varphi_2), \mathrm{cmc}(\varphi_1 \varphi_2) \leq \max\{\mathrm{cmc}(\varphi_1), \mathrm{cmc}(\varphi_2)\}$.

Proof of Theorem 6.1. We will use induction on $c = \mathrm{cmp}(u)$. All the gen-polynomials we will deal with will have complexity $\le c$, thus the index set $\mathcal{B}$ below may be replaced by $\{\alpha \in \mathcal{B} : |\alpha| \le c\}$ and assumed to be finite.

Step 1. Represent all the polynomials occurring in u as linear combinations with rational coefficients of polynomials p_α, polynomials from $\mathbb{Q}[n]$, and constants. Then the sub-gen-monomials of u are constructed from rational multiples of polynomials p_α, polynomials from $\mathbb{Q}[n]$, and constants; moreover, no products p_1p_2 of two polynomials occur in u.

If a rational polynomial $p \in \mathbb{Q}[n]$ occurs in u and is closed in u, then $\{\!\{p\}\!\}$ takes on only finitely many values; after passing to any translate of a suitable sublattice of $\mathbb{Z}^d$ we convert $\{\!\{p\}\!\}$ to a rational constant. If there is a non-closed occurrence of a polynomial $p \in \mathbb{Q}[n]$ in u, we apply the transformation (5.6) to close p, and then eliminate it. (The second term in (5.6) has excessive complexity, but an application of (5.1) and (5.4) reduces it by 1.) Thus, we now assume that no rational polynomials occur in u.

If a constant occurs in u, by applying the identities (5.2), (5.3), and (5.4) we reduce the complexity of u and use induction. So, assume that no constants occur in u, except rational coefficients before polynomials p_α. Now, u is constructed from polynomials λp_α, with coefficients $\lambda \in \mathbb{Q}$. Let r_0 be the least common multiple of the denominators of the coefficients λ.

By Lemma 3.1, u is a piecewise linear function of its monomials, so if the assertion of the theorem holds for the sub-gen-monomials of u, it holds for u; thus, we may assume that u is a gen-monomial.

Take $M \in \mathbb{N}$ divisible enough, to be specified. We will write v_α, $\alpha \in \mathcal{B}$, for $v_\alpha(M^{-1}\mathcal{P})$. Since $v_\alpha(M^{-1}\mathcal{P}) = M^{-1}p_\alpha$, $\alpha \in \mathcal{A}$, u is constructed from the gen-polynomials λv_α, $\alpha \in \mathcal{A}$, with coefficients $\lambda \in Mr_0^{-1}\mathbb{Z}$; assuming that M is divisible by r_0, we will have all $\lambda \in \mathbb{Z}$.

We will eliminate occurrences of gen-polynomials v_β, $\beta \in \mathcal{B}$, successively from the expression for u; since $\mathcal{B}$ is assumed to be finite, this will be a finite process. Assume by induction on β that u is constructed from gen-polynomials λv_α with $\alpha \in \mathcal{B}_\beta$, $\alpha \ge \beta$, and coefficient $\lambda \in M^{|\alpha|}r^{-1}\mathbb{Z}$, where r depends only on r_0, $c = \mathrm{cmp}(u)$ and β; assuming that M is divisible by r, we will have all $\lambda \in \mathbb{Z}$. We will now eliminate v_β from u.

Using (5.2) and (5.3), and taking only the principal sub-gen-monomial of the obtained expression, we replace each occurrence of $\{\!\{\lambda v_\beta\}\!\}$ with $\lambda \ne 1$ by $\lambda\{\!\{v_\beta\}\!\}$. If there is a non-closed occurrence of v_β in u, it appears in an expression $A = \{\!\{\lambda' v_\beta \lambda'' \{\!\{v_\beta\}\!\}^{m-1}\{\!\{u_1\}\!\}\dots\{\!\{u_l\}\!\}\}\!\}$ with $m \ge 1$, $l \ge 0$, and $u_i \ne v_\beta$, $i = 1,\dots,l$. We apply the transformation (5.7) to close v_β; we assume here that M is divisible enough for $\lambda = \lambda'\lambda''$ to be divisible by m. In each principal sub-gen-monomial of the new representation of u, the expression A is replaced either by one of the expressions $A_j = \{\!\{\frac{-\lambda}{m}u_j\{\!\{v_\beta\}\!\}^m \prod_{i\ne j}\{\!\{u_i\}\!\}\}\!\}$, or by $A_0 = \{\!\{\frac{\lambda}{m}\{\!\{v_\beta\}\!\}^m \prod_{i=1}^{l}\{\!\{u_i\}\!\}\}\!\}$; the complexity of A_0 is larger by 1 than the complexity of A, but an application of (5.2), (5.3) and (5.4) fixes this problem. We now represent u as a piecewise linear function of its sub-gen-monomials, and deal with the principal sub-gen-monomials of u separately. Notice also that $m \le c$, so that if $\lambda' \in M^{|\beta|}r^{-1}\mathbb{Z}$ and $\lambda'' \in M^{(m-1)|\beta|}r^{-m}\mathbb{Z}$, then the new coefficient $\frac{\pm\lambda}{m}$ before $\{\!\{v_\beta\}\!\}^m$ is contained in $M^{m|\beta|}r^{-c}(c!)^{-1}\mathbb{Z}$.

Now assume that all occurrences of v_β in u are closed. If $\{\!\{v_\beta\}\!\}$ occurs on the upper level of u, that is, if $u = \lambda\{\!\{v_\beta\}\!\}^m\{\!\{u_1\}\!\}\dots\{\!\{u_l\}\!\}$ with $m \ge 1$, $l \ge 0$, then we replace u by the gen-polynomials $\{\!\{u_1\}\!\}$, …, $\{\!\{u_l\}\!\}$ and use induction on $\mathrm{cmp}(u)$. Otherwise, v_β occurs in an expression $\{\!\{\lambda' v_\gamma \lambda''\{\!\{v_\beta\}\!\}^m\{\!\{u_1\}\!\}\dots\{\!\{u_l\}\!\}\}\!\}$ with $\gamma \in \mathcal{B}_\beta$, $\gamma > \beta$, $m \ge 1$, $l \ge 0$, and $u_i \ne v_\beta$, $i = 1,\dots,l$. We then replace $\lambda' v_\gamma \lambda''\{\!\{v_\beta\}\!\}^m$ by λv_α where $\alpha = [\beta, m\gamma] \in \mathcal{E}_\gamma$, eliminating thereby m occurrences of $\{\!\{v_\beta\}\!\}$. Notice also that if $\lambda' \in M^{|\gamma|}r^{-1}\mathbb{Z}$ and $\lambda'' \in M^{m|\beta|}r^{-m}\mathbb{Z}$, then $\lambda = \lambda'\lambda'' \in M^{|\alpha|}r^{-c-1}\mathbb{Z}$. Thus, after the "closing" of all gen-monomials v_β in u, the coefficient λ before every gen-monomial v_α is contained in $M^{|\alpha|}r^{-1}\mathbb{Z}$, where this (new) r only depends on r_0, c and β.

Step 2. In the process of closing the gen-monomial v_β in u, we have got, in addition to the principal sub-gen-monomials of the new representation of u, a new gen-polynomial w of smaller complexity. We cannot apply our induction assumption to w immediately since w can contain products of non-closed sub-gen-monomials, and so, we have to start all the process of reduction of w from the beginning. However, the coefficients with which the polynomials p_α occur in w may now contain the factor M^{-1}, thus the canonical form of w will be a pp-function in the variables $v_\alpha(M_1^{-1}\mathcal{P})$ with M_1 depending on M. So, what we can state now is only that u is representable as a pp-function in the variables $v_\alpha(M_j^{-1}\mathcal{P})$, $\alpha \in \mathcal{B}$, $j \in \mathbb{N}$. If we want to have a universal M, we have to reconstruct w from polynomials λp_α with $\lambda \in r^{-1}\mathbb{Z}$ with r not depending on M.

The gen-polynomial w is constructed from gen-monomials $\lambda v_\alpha(M^{-1}\mathcal{P})$ with $\lambda \in M^{|\alpha|}r^{-1}\mathbb{Z}$, where r

does not depend on M. We will decompose these gen-monomials back to an expression in polynomials p_α, $\alpha \in \mathcal{A}$, with the help of the identity (5.2), from which it follows that for any $\alpha = [\gamma, m\beta]$,

$$M^{|\alpha|} v_\alpha = M^{|\gamma|+m|\beta|} v_\gamma \{\!\!\{ v_\beta \}\!\!\}^m = M^{|\gamma|} v_\gamma \big(\{\!\!\{ M^{|\beta|} v_\beta \}\!\!\} + C\big)^m, \tag{6.1}$$

where C is an integer-valued piecewise linear function of $\{\!\!\{v_\beta\}\!\!\}$. Applying this transformation inductively, and taking into account that $Mv_\alpha(M^{-1}\mathcal{P}) = MM^{-1}p_\alpha = p_\alpha$ for $\alpha \in \mathcal{A}$, we represent w as a piecewise gen-polynomial whose conditions are constructed from gen-polynomials $\{\!\!\{v_\beta(M^{-1}\mathcal{P})\}\!\!\}$, $\beta \in \mathcal{B}$, and variants from the polynomials λp_α, $\alpha \in \mathcal{A}$, $\lambda \in r^{-1}\mathbb{Z}$. We can now apply the induction hypothesis to the variants of w and also represent each of them as a pp-function of $\{\!\!\{v_\beta(M^{-1}\mathcal{P})\}\!\!\}$, $\beta \in \mathcal{B}$; this gives us a representation of w as a pp-function in the variables $\{\!\!\{v_\beta(M^{-1}\mathcal{P})\}\!\!\}$, $\beta \in \mathcal{B}$.

Notice also that during the whole process we increased neither the complexity nor the depth of our gen-polynomials, so the complexity and the depth of the resulting gen-polynomial will be $\leq \mathrm{cmp}(u)$ and $\leq \mathrm{dep}(u)$ respectively. ■

Examples. 1. Let $u = \{\!\!\{p\{\!\!\{q\{\!\!\{p\{\!\!\{q\}\!\!\}\}\!\!\}\}\!\!\}\}\!\!\}$; to demonstrate the method, we will only find the principal sub-gen-monomials of the canonical form of u and ignore all gen-monomials of complexity less than 4. Let us use the same symbols p and q for indices, so that $\mathcal{A} = \{p, q\}$. We will assume $p < q$ (although the canonical form would be found easier if we assumed that $q < p$). Take $M = 2$, put $v_p = v_p\{\!\!\{2^{-1}\mathcal{A}\}\!\!\} = \frac{1}{2}p$ and $v_q = v_q\{\!\!\{2^{-1}\mathcal{A}\}\!\!\} = \frac{1}{2}q$. Rewrite u as $\{\!\!\{2v_p\{\!\!\{2v_q\{\!\!\{2v_p\{\!\!\{2v_q\}\!\!\}\}\!\!\}\}\!\!\}\}\!\!\}$.

We use (5.6) to close "the deeper" occurrence of v_p in u, then the principal monomials of the new expression for u will be

$$u_1 = \{\!\!\{2v_p\{\!\!\{4v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{2v_q\}\!\!\}\}\!\!\}\}\!\!\} \text{ and } u_2 = \{\!\!\{2v_p\{\!\!\{2v_q\{\!\!\{4v_q\{\!\!\{v_p\}\!\!\}\}\!\!\}\}\!\!\}\}\!\!\}.$$

Next, we transform u_1 into a combination of

$$u_3 = \{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\}\}\!\!\} = \{\!\!\{v_p\}\!\!\}\{\!\!\{v_{[[q,p],q]}\}\!\!\} \text{ and } u_4 = \{\!\!\{v_q\{\!\!\{v_p\}\!\!\}^2\{\!\!\{v_q\}\!\!\}\}\!\!\} = \{\!\!\{v_{[[q,2p],q]}\}\!\!\},$$

and u_2 to a combination of

$$u_5 = \{\!\!\{v_p\}\!\!\}\{\!\!\{2v_q\{\!\!\{4v_q\{\!\!\{v_p\}\!\!\}\}\!\!\}\}\!\!\} = \{\!\!\{v_p\}\!\!\}\{\!\!\{2v_q\{\!\!\{4v_{[q,p]}\}\!\!\}\}\!\!\} \text{ and } u_6 = \{\!\!\{4v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{4v_q\{\!\!\{v_p\}\!\!\}\}\!\!\}\}\!\!\} = \{\!\!\{4v_{[q,p]}\{\!\!\{4v_{[q,p]}\}\!\!\}\}\!\!\}.$$

The gen-polynomials u_3 and u_4 are already in the final form. In u_5, closing v_q, we transform $\{\!\!\{2v_q\{\!\!\{4v_q\{\!\!\{v_p\}\!\!\}\}\!\!\}$ into a combination of

$$\{\!\!\{v_q\}\!\!\}\{\!\!\{v_q\{\!\!\{v_p\}\!\!\}\}\!\!\} = \{\!\!\{v_q\}\!\!\}\{\!\!\{v_{[q,p]}\}\!\!\} \text{ and } \{\!\!\{v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\}\}\!\!\} = \{\!\!\{v_{[[q,p],q]}\}\!\!\}.$$

Finally, using (5.7), we transform u_6 into $8\{\!\!\{v_{[q,p]}\}\!\!\}^2$.

2. We will now transform, in whole details, the gen-polynomial $u = \{\!\!\{q\{\!\!\{p\{\!\!\{q\}\!\!\}\}\!\!\}\}\!\!\}$ into its canonical form. In addition to p and q we will need the polynomials $r = pq$, $s = q^2$, and $t = rq = pq^2$. We will assume that the polynomials p, q, r, s, t are $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n] + \mathbb{R}$; if this is not so, the canonical form will change. We take $\mathcal{A} = \{p, q, r, s, t\}$ with $p < q < r < s < t$ (although, again, this order is not optimal), and put $v_p = p$, $v_q = q$, $v_r = r$, $v_s = s$, and $v_t = t$ to start with, so that $u = \{\!\!\{v_q\{\!\!\{v_p\{\!\!\{v_q\}\!\!\}\}\!\!\}\}\!\!\}$.

By (5.6), we have $u = \big\{\!\!\big\{v_q\big\{\!\!\big\{\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - v_q\{\!\!\{v_p\}\!\!\} + v_p v_q\big\}\!\!\big\}\big\}\!\!\big\} = \big\{\!\!\big\{v_q\big\{\!\!\big\{\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - v_{[q,p]} + r\big\}\!\!\big\}\big\}\!\!\big\}$, and so,

$$u = \begin{cases} \big\{\!\!\big\{v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - v_q\{\!\!\{v_{[q,p]}\}\!\!\} + v_q\{\!\!\{r\}\!\!\} + v_q\big\}\!\!\big\} = \big\{\!\!\big\{v_{[[q,p],q]} - v_q\{\!\!\{v_{[q,p]}\}\!\!\} + v_q\{\!\!\{r\}\!\!\} + v_q\big\}\!\!\big\} \\ \qquad \text{if } -1 \leq \{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - \{\!\!\{v_{[q,p]}\}\!\!\} + \{\!\!\{r\}\!\!\} < 0 \\ \big\{\!\!\big\{v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - v_q\{\!\!\{v_{[q,p]}\}\!\!\} + v_q\{\!\!\{r\}\!\!\}\big\}\!\!\big\} = \big\{\!\!\big\{v_{[[q,p],q]} - v_q\{\!\!\{v_{[q,p]}\}\!\!\} + v_q\{\!\!\{r\}\!\!\}\big\}\!\!\big\} \\ \qquad \text{if } 0 \leq \{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - \{\!\!\{v_{[q,p]}\}\!\!\} + \{\!\!\{r\}\!\!\} < 1 \\ \big\{\!\!\big\{v_q\{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - v_q\{\!\!\{v_{[q,p]}\}\!\!\} + v_q\{\!\!\{r\}\!\!\} - v_q\big\}\!\!\big\} = \big\{\!\!\big\{v_{[[q,p],q]} - v_q\{\!\!\{v_{[q,p]}\}\!\!\} + v_q\{\!\!\{r\}\!\!\} - v_q\big\}\!\!\big\} \\ \qquad \text{if } 1 \leq \{\!\!\{v_p\}\!\!\}\{\!\!\{v_q\}\!\!\} - \{\!\!\{v_{[q,p]}\}\!\!\} + \{\!\!\{r\}\!\!\} < 2 \end{cases}$$

Next,

$$\big\{\!\!\big\{v_q\{\!\!\{r\}\!\!\}\big\}\!\!\big\} = \big\{\!\!\big\{\{\!\!\{v_q\}\!\!\}\{\!\!\{r\}\!\!\} - r\{\!\!\{v_q\}\!\!\} + rv_q\big\}\!\!\big\} = \big\{\!\!\big\{\{\!\!\{v_q\}\!\!\}\{\!\!\{r\}\!\!\} - v_{[r,q]} + rq\big\}\!\!\big\} = \big\{\!\!\big\{\{\!\!\{v_q\}\!\!\}\{\!\!\{r\}\!\!\} - v_{[r,q]} + t\big\}\!\!\big\}$$

and

$$\{v_q\{v_{[q,p]}\}\} = \big\{\{v_q\}\{v_{[q,p]}\} - v_{[q,p]}\{v_q\} + v_{[q,p]}v_q\big\} = \big\{\{v_q\}\{v_{[q,p]}\} - v_{[[q,p],q]} + v_q\{v_p\}v_q\big\}$$
$$= \big\{\{v_q\}\{v_{[q,p]}\} - v_{[[q,p],q]} + v_{[s,p]}\big\}$$

Combining, we get

$$u = \begin{cases} \Big\{2\{v_{[[q,p],q]}\} - \{v_q\}\{v_{[q,p]}\} - \{v_{[s,p]}\} + \{v_q\}\{r\} - \{v_{[r,q]}\} + \{t\} + \{v_q\}\Big\} \\ \qquad \text{if } -1 \le \{v_p\}\{v_q\} - \{v_{[q,p]}\} + \{r\} < 0 \\ \Big\{2\{v_{[[q,p],q]}\} - \{v_q\}\{v_{[q,p]}\} - \{v_{[s,p]}\} + \{v_q\}\{r\} - \{v_{[r,q]}\} + \{t\}\Big\} \\ \qquad \text{if } 0 \le \{v_p\}\{v_q\} - \{v_{[q,p]}\} + \{r\} < 1 \\ \Big\{2\{v_{[[q,p],q]}\} - \{v_q\}\{v_{[q,p]}\} - \{v_{[s,p]}\} + \{v_q\}\{r\} - \{v_{[r,q]}\} + \{t\} - \{v_q\}\Big\} \\ \qquad \text{if } 1 \le \{v_p\}\{v_q\} - \{v_{[q,p]}\} + \{r\} < 2 \end{cases}$$

(We do not open the outer brackets in order to avoid, in each case, considering 8 different sub-cases, dependently on the value of the integer part of the expression inside the brackets.)

7. Independence of basic gen-polynomials

In this section we will prove Theorem 0.1, namely, that if $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ is a system of polynomials on $\mathbb{Z}^d$ which is $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n] + \mathbb{R}$, then the gen-polynomials $\{v_\alpha(\mathcal{P})\}$, $\alpha \in \mathcal{B}(\mathcal{A})$, are jointly well distributed.

Let $\{u_\alpha\}$, $\alpha \in \mathcal{B} = \mathcal{B}(\mathcal{A})$, be the free coordinate gen-polynomials. By definition, u_α are gen-polynomials in the variables x_β with $\beta \in \mathcal{B}$, but, starting from this point, we will assume that $x_\beta = 0$ for $\beta \notin \mathcal{A}$, and consider u_α as gen-polynomials in the variables x_β with $\beta \in \mathcal{A}$ only. We know from Proposition 2.4 that the gen-polynomials $\{u_\alpha(\mathcal{P})\}$, $\alpha \in \mathcal{B}$, are jointly well distributed; moreover, for any sublattice Λ of $\mathbb{Z}^d$ and any $n_0 \in \mathbb{Z}^d$ the restrictions of $\{u_\alpha(\mathcal{P})\}$, $\alpha \in \mathcal{B}$, on $n_0 + \Lambda$ are also jointly well distributed.

Our goal is now to compare u_α and v_α.

Lemma 7.1. *For any $\alpha \in \mathcal{B}$, $v_\alpha = m_\alpha u_\alpha + \varphi_\alpha$, where $m_\alpha \in \mathbb{N}$ and φ_α is a piecewise gen-polynomial with* $\mathrm{cmp}(\varphi_\alpha) \le |\alpha| - 2$ *and* $\mathrm{cmc}(\varphi_\alpha) \le |\alpha| - 1$.

Proof. We use induction on $|\alpha|$; if $|\alpha| = 0$, $v_\alpha = u_\alpha = p_\alpha$. Let $\alpha = [\gamma, m\beta]$. By induction, $v_\gamma = m_\gamma u_\gamma + \varphi_\gamma$ and $v_\beta = m_\beta u_\beta + \varphi_\beta$, where $m_\gamma, m_\beta \in \mathbb{N}$ and $\varphi_\gamma, \varphi_\beta$ are piecewise gen-polynomials with $\mathrm{cmp}(\varphi_\gamma) \le |\gamma| - 2$, $\mathrm{cmc}(\varphi_\gamma) \le |\gamma| - 1$, $\mathrm{cmp}(\varphi_\beta) \le |\beta| - 2$, and $\mathrm{cmc}(\varphi_\beta) \le |\beta| - 1$. Thus,

$$v_\alpha = v_\beta\{v_\gamma\}^m = (m_\gamma u_\gamma + \varphi_\gamma)\{m_\beta u_\beta + \varphi_\beta\}^m.$$

$\{m_\beta u_\beta + \varphi_\beta\} = m_\beta\{u_\beta\} + \{\varphi_\beta\} + \eta$, where η is an (integer-valued) piecewise gen-polynomial with $\mathrm{cmp}(\eta) = 0$ and $\mathrm{cmc}(\eta) = \max\big\{\mathrm{cmp}(\{u_\beta\}), \mathrm{cmp}(\{\varphi_\beta\}), \mathrm{cmc}(\varphi_\beta)\big\} = |\beta|$. So,

$$\{m_\beta u_\beta + \varphi_\beta\}^m = m_\beta^m\{u_\beta\}^m + \psi$$

where ψ is a piecewise gen-polynomial with $\mathrm{cmp}(\psi) \le m|\beta| - 1$ and $\mathrm{cmc}(\psi) = |\beta|$. Hence,

$$v_\alpha = (m_\gamma u_\gamma + \varphi_\gamma)\big(m_\beta^m\{u_\beta\}^m + \psi\big) = m_\gamma m_\beta^m u_\gamma\{u_\beta\}^m + \varphi_\alpha = m_\gamma m_\beta^m m! u_\alpha + \varphi_\alpha,$$

where $\varphi_\alpha = m_\gamma u_\gamma\psi + m_\beta^m\varphi_\gamma\{u_\beta\}^m + \varphi_\gamma\psi$ is a piecewise gen-polynomial with $\mathrm{cmp}(\varphi_\alpha) \le |\gamma| + m|\beta| - 2 = |\alpha| - 2$ and $\mathrm{cmc}(\varphi_\alpha) = \max\{|\beta|, |\gamma| - 1\} \le |\alpha| - 1$. ■

We will also need the version of Theorem 6.1 that deals with gen-polynomials u_α instead of v_α:

Theorem 7.2. *Let u be a bounded gen-polynomial over $\mathbb{Z}^d$. Let $\mathcal{R}$ be the algebra generated by the polynomials that occur in u and let $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ be a system of polynomials such that $\mathrm{span}_{\mathbb{Q}}\mathcal{P} + \mathbb{Q}[n] + \mathbb{R} \supseteq \mathcal{R}$. If $M \in \mathbb{N}$ is divisible enough, then there exists a sublattice Λ in $\mathbb{Z}^d$ such that for any translate $\Lambda' = n_0 + \Lambda$ of Λ one has $u|_{\Lambda'} = f\big(\{u_\alpha(M^{-1}\mathcal{P})\} : \alpha \in \mathcal{B}\big)|_{\Lambda'}$ where f is a pp-function and* $\mathrm{cmp}\big(f\big(\{u_\alpha\} : \alpha \in \mathcal{B}\big)\big) \le \mathrm{cmp}(u)$.

Proof. We copy Step 1 of the proof of Theorem 6.1: passing to a sublattice of $\mathbb{Z}^d$, taking M divisible enough, and dealing with the principal sub-gen-monomials of u only, we represent u in the form of $u = g\big((\{\!\{u_\alpha(M^{-1}\mathcal{P})\}\!\} : \alpha \in \mathcal{B}), w\big)$ where g is a pp-function and w is a gen-polynomial with $\mathrm{cmp}(w) < \mathrm{cmp}(u)$. As for Step 2, we cannot copy it since we have no control on the non-principal sub-gen-monomials of the gen-polynomials u_α. We can, however, apply induction on $\mathrm{cmp}(u)$ and represent w in a similar form, but with a different M. This way we obtain a representation of u of the form $u = f\big(\{\!\{u_\alpha(M_j^{-1}\mathcal{P})\}\!\} : \alpha \in \mathcal{B},\ j = 1, \dots, k\big)$, where f is a pp-function, with $M_1, \dots, M_k \in \mathbb{N}$.

We may now utilize the natural origination of the gen-polynomials u_α:

Lemma 7.3. *For any $M \in \mathbb{N}$ and any $\alpha \in \mathcal{B}$ there exists a pp-function h such that $\{\!\{u_\alpha(x_1, \dots, x_r)\}\!\} = h\big(\{\!\{u_\beta(M^{-1}x_1, \dots, M^{-1}x_r)\}\!\} : |\beta| \leq |\alpha|\big)$, and $\mathrm{cmp}\big(h(\{\!\{u_\beta\}\!\} : |\beta| \leq |\alpha|)\big) = \mathrm{cmp}(u)$.*

Proof. Let $X = F/\Delta$ be a free nilmanifold large enough so that u_α is a coordinate gen-polynomial coming from X. Let $e_1, \dots, e_r \in \Delta$ be the free generators of Δ (and of F), and let Δ_M be the discrete group generated by $e_1^M, \dots, e_r^M$. Then Δ_M is a free group cocompact in F, so $X_M = F/\Delta_M$ is also a free nilmanifold, for which X is a factor; let $\pi \colon F \longrightarrow X$, $\pi_M \colon F \longrightarrow X_M$, and $\zeta \colon X_M \longrightarrow X$ be the natural projections.

$$\begin{array}{ccc} & F & \\ {\scriptstyle \pi_M}\swarrow & & \searrow{\scriptstyle \pi} \\ X_M & \xrightarrow{\ \zeta\ } & X \end{array}$$

Let $x_1, \dots, x_r, \dots$ be the coordinates on F with respect to the basis $e_1, \dots, e_r, \dots$, and let $y_1, \dots, y_r, \dots$ be the coordinates on F with respect to the basis $e_1^M, \dots, e_r^M, \dots$; then $y_i = M^{-1}x_i$, $i = 1, \dots, r$. In these coordinates, the α-component of the projection π is the gen-polynomial $\{\!\{u_\alpha(x_1, \dots, x_r)\}\!\}$, the β-component of the projection π_M is $\{\!\{u_\beta(y_1, \dots, y_r)\}\!\}$, $\beta \in \mathcal{B}$. The projection ζ is given in coordinates on X_M and on X by pp-functions; let h be the α-component of this mapping. Then $u_\alpha(x_1, \dots, x_r) = h\big(u_\beta(y_1, \dots, y_r) : |\beta| \leq |\alpha|\big) = h\big(u_\beta(M^{-1}x_1, \dots, M^{-1}x_r) : |\beta| \leq |\alpha|\big)$. Since the weighted degree of $h\big(\{\!\{y_\beta\}\!\} : |\beta| \leq |\alpha|\big)$ is equal to $|\alpha|$ assuming y_β has degree $|\beta|$, $\beta \in \mathcal{B}$, we also have $\mathrm{cmp}\big(h\big(\{\!\{u_\beta\}\!\} : |\beta| \leq |\alpha|\big)\big) = \mathrm{cmp}(u)$. ∎

Now, for M being a common multiple of $M_1, \dots, M_k$, Theorem 7.2 follows. ∎

Example. We will demonstrate how Lemma 7.3 works on the case where F is the free 3-step nilpotent group with only two (continuous) generators a and b. The basic commutators in F are

$$a,\ b,\ [b,a],\ [b,2a],\ \text{and } [[b,a],b]. \tag{7.1}$$

Let $M \in \mathbb{N}$. Let Δ_M be the group generated by a^M and b^M. The basic commutators in Δ_M are

$$\begin{aligned} a^M, \quad b^M, \quad [b^M, a^M] = [b,a]^{M^2}[b,2a]^{\binom{M}{2}M}[[b,a],b]^{\binom{M}{2}M}, \quad [b^M, 2a^M] = [b,2a]^{M^3}, \\ \text{and } [[b^M,a^M],b^M] = [[b,a],b]^{M^3}. \end{aligned} \tag{7.2}$$

A point in F with coordinates $(y_1, \dots, y_5)$ with respect to the basis (7.2) is

$$\begin{aligned} &(a^M)^{y_1}(b^M)^{y_2}\big([b,a]^{M^2}[b,2a]^{\binom{M}{2}M}[[b,a],b]^{\binom{M}{2}M}\big)^{y_3}([b,2a]^{M^3})^{y_4}([[b,a],b]^{M^3})^{y_5} \\ &\qquad = a^{My_1}b^{My_2}[b,a]^{M^2y_3}[b,2a]^{\binom{M}{2}My_3 + M^3y_4}[[b,a],b]^{\binom{M}{2}My_3 + M^3y_5}, \end{aligned}$$

and hence has coordinates

$$x_1 = My_1, \quad x_2 = My_2, \quad x_3 = M^2y_3, \quad x_4 = \tbinom{M}{2}My_3 + M^3y_4, \quad \text{and } x_5 = \tbinom{M}{2}My_3 + M^3y_5$$

with respect to the the basis (7.1). Next, one can check that the projection to X of a point with coordinates $x_1, \dots, x_5$ in F has coordinates

$$\begin{aligned} &\{\!\{x_1\}\!\}, \quad \{\!\{x_2\}\!\}, \quad \{\!\{(\{\!\{x_1\}\!\} - x_1)x_2 + x_3\}\!\}, \quad \Big\{\!\Big\{\tbinom{\{\!\{x_1\}\!\} - x_1}{2}x_2 + (\{\!\{x_1\}\!\} - x_1)x_3 + x_4\Big\}\!\Big\}, \\ &\qquad \text{and } \Big\{\!\Big\{(\{\!\{x_1\}\!\} - x_1)x_2(\{\!\{x_2\}\!\} - x_2) + (\{\!\{x_1\}\!\} - x_1)\tbinom{x_2}{2} + (\{\!\{x_2\}\!\} - x_2)x_3 + x_5\Big\}\!\Big\} \end{aligned}$$

in X. Combining, we find that, in the coordinates on X_M and on X,

$$\begin{aligned}\zeta(y_1,y_2,y_3,y_4,y_5) = \Big(&\{\!\{My_1\}\!\},\ \{\!\{My_2\}\!\},\ \{\!\{(\{\!\{My_1\}\!\}-My_1)My_2+M^2y_3\}\!\},\\ &\Big\{\!\!\Big\{\tbinom{\{\!\{My_1\}\!\}-My_1}{2}My_2+(\{\!\{My_1\}\!\}-My_1)M^2y_3+\tbinom{M}{2}My_3+M^3y_4\Big\}\!\!\Big\},\\ &\Big\{\!\!\Big\{(\{\!\{My_1\}\!\}-My_1)My_2(\{\!\{My_2\}\!\}-My_2)+(\{\!\{My_1\}\!\}-My_1)\tbinom{My_2}{2}+(\{\!\{My_2\}\!\}-My_2)M^2y_3+\tbinom{M}{2}My_3+M^3y_5\Big\}\!\!\Big\}\Big)\end{aligned}$$

(which is a pp-function since $y_1,\dots,y_5$ are assumed to belong to $[0,1]$). Finally, replacing in this formula $y_1,\dots,y_5$ respectively by $u_a(M^{-1}x_1,M^{-1}x_2),\dots,u_{[b,a],b]}(M^{-1}x_1,M^{-1}x_2)$, we obtain that

$$\begin{aligned}\{\!\{u_a(x_1,x_2)\}\!\} &= \{\!\{M\{\!\{u_a(M^{-1}x_1,M^{-1}x_2)\}\!\}\}\!\}\\ \{\!\{u_b(x_1,x_2)\}\!\} &= \{\!\{M\{\!\{u_b(M^{-1}x_1,M^{-1}x_2)\}\!\}\}\!\}\\ \{\!\{u_{[b,a]}(x_1,x_2)\}\!\} &= \{\!\{(\{\!\{M\{\!\{u_a(M^{-1}x_1,M^{-1}x_2)\}\!\}\}\!\}-M\{\!\{u_a(M^{-1}x_1,M^{-1}x_2)\}\!\})M\{\!\{u_b(M^{-1}x_1,M^{-1}x_2)\}\!\}\\ &\qquad +M^2\{\!\{u_{[b,a]}(M^{-1}x_1,M^{-1}x_2)\}\!\}\}\!\}\\ \{\!\{u_{[b,2a]}(x_1,x_2)\}\!\} &= \dots\\ \{\!\{u_{[[b,a],b]}(x_1,x_2)\}\!\} &= \dots\end{aligned}$$

Proof of Theorem 0.1. By Lemma 7.1, for any $\alpha\in\mathcal{B}$ we have

$$\{\!\{v_\alpha\}\!\} = m_\alpha\{\!\{u_\alpha\}\!\}+\{\!\{\varphi_\alpha\}\!\} \bmod 1$$

with $\mathrm{cmp}(\{\!\{\varphi_\alpha\}\!\}),\mathrm{cmc}(\{\!\{\varphi_\alpha\}\!\})\le|\alpha|-1$. By Theorem 7.2, there exists a sublattice Λ of $\mathbb{Z}^d$ such that for every translate $\Lambda'=n_0+\Lambda$ of Λ one has $\{\!\{\varphi_\alpha\}\!\}_{|\Lambda'}=f\big(\{\!\{u_\beta\}\!\}:|\beta|<|\alpha|\big)_{|\Lambda'}$, where f is a pp-function. Since $\{\!\{u_\alpha\}\!\}_{|\Lambda'}$, $\alpha\in\mathcal{B}$, are jointly well distributed for every translate Λ' of Λ, $\{\!\{v_\alpha\}\!\}_{|\Lambda'}$ are also jointly well distributed for every translate Λ' of Λ, and thus $\{\!\{v_\alpha\}\!\}$ are jointly well distributed. ∎

We now have the uniqueness part of Theorem 0.2:

Corollary 7.4. *In the formulation of Theorem 0.2 or Theorem 6.1, if $M\in\mathbb{N}$ and a family $\mathcal{P}$ are fixed and $\mathcal{P}$ is $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$, the pp-function f is defined uniquely up to a pp-function vanishing a.e.*

Proof. Indeed, let a pp-function f on $[0,1]^l$, a sublattice Λ' of Λ, and distinct indices $\alpha_1,\dots,\alpha_l\in\mathcal{B}$ be such that $f\big(\{\!\{v_{\alpha_1}(M^{-1}\mathcal{P})\}\!\},\dots,\{\!\{v_{\alpha_l}(M^{-1}\mathcal{P})\}\!\}\big)_{|\Lambda'}=0$. Let $[0,1]^l=\bigcup_{i=1}^k Q_i$, where Q_i are defined by polynomial inequalities and $f_{|Q_i}$ is a polynomial for every i. Then, since the gen-polynomials $\{\!\{v_{\alpha_j}(M^{-1}\mathcal{P})\}\!\}_{|\Lambda'}$, $j=1,\dots,l$, are jointly well distributed in $[0,1]^l$, it must be that $f_{|Q_i}=0$ for every i such that Q_i has positive measure. Thus, $f=0$ a.e. on $[0,1]^l$. ∎

8. Finding the limiting distribution of values of bounded gen-polynomials

Given one or several bounded gen-polynomials $u_1,\dots,u_k$, to find the limiting distribution of values of the vector gen-polynomial $\mathbf{u}=(u_1,\dots,u_k)$ we choose a system $\mathcal{P}=\{p_\alpha:\alpha\in\mathcal{A}\}$ of ordinary polynomials that spans, modulo $\mathbb{Q}[n]+\mathbb{R}$, the algebra generated by the polynomials occurring in $u_1,\dots,u_k$, and write $\mathbf{u}$ in the form $\mathbf{u}=F\big(v_{\alpha_1}(M^{-1}\mathcal{P}),\dots,v_{\alpha_l}(M^{-1}\mathcal{P})\big)$, where $\alpha_1,\dots,\alpha_l\in\mathcal{B}(\mathcal{A})$, $M\in\mathbb{N}$, and F is a vector-valued pp-function; since the gen-polynomials $v_\alpha(M^{-1}\mathcal{P})$, $\alpha\in\mathcal{B}(\mathcal{A})$, are jointly well distributed in $[0,1]^l$, the limiting distribution of $\mathbf{u}$ will be given by F. As a corollary, we obtain the following theorem, proved in [BL]:

Theorem 8.1. *Let $u_1,\dots,u_k\colon\mathbb{Z}^d\longrightarrow\mathbb{R}$ be bounded gen-polynomials. There exist $l\in\mathbb{N}$ and a pp-function $F\colon[0,1]^l\longrightarrow\mathbb{R}^k$ such that the values of the function $\mathbf{u}=(u_1,\dots,u_k)$ are well distributed in the set $F([0,1]^l)$ with respect to the measure $F(\mu)$, where μ is the Lebesgue measure on $[0,1]^l$.*

Example. Let p,q be polynomials $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$ and let the $\mathbb{R}^2$-valued gen-polynomial u have the canonical form

$$\mathbf{u}=\begin{cases}\Bigl(\{v_p\}\{v_q\}^2\{v_{[q,p]}\}+2\{v_p\}\{v_q\}\{v^3_{[[q,p],q]}\},\{3v_p\}^2\{v_q\}-4\{v_{[[q,p],q]}\}\Bigr) & \text{if } 0\le\{v_p\}+\{v_q\}\{v_{[q,p]}\}<1\\ \Bigl(5\{v_p\}^2\{v_q\}-6\{v_p\}\{v_q\},7\{v_p\}^2\{v_q\}+8\{v_{[[q,p],q]}\}+9\Bigr) & \text{if } 1\le\{v_p\}+\{v_q\}\{v_{[q,p]}\}<2\end{cases}$$

where $v_p=p$, $v_q=q$, $v_{[q,p]}=q\{p\}$, and $v_{[[q,p],q]}=q\{p\}\{q\}$. Define

$$F(x_1,x_2,x_3,x_4)=\begin{cases}\bigl(x_1x_2^2x_3+2x_1x_2x_4^3,3x_1^2x_2-4x_4\bigr) & \text{if } 0\le x_1+x_2x_3<1\\ \bigl(5x_1^2x_2-6x_1x_2,7x_1^2x_2+8x_4+9\bigr) & \text{if } 1\le x_1+x_2x_3<2,\end{cases}$$

$x_1,x_2,x_3,x_4\in[0,1]$. Then the values of $\mathbf{u}$ are well distributed on the surface $F([0,1]^4)$.

The following proposition can be used to simplify the process of finding the distribution of values of bounded gen-polynomials:

Proposition 8.2. *Let $\mathcal{A}$ be a well-ordered set, and let gen-polynomials w_α, $\alpha\in\mathcal{B}=\mathcal{B}(\mathcal{A})$, be such that for every α the only principal sub-gen-monomial of w_α equals $\lambda_\alpha v_\alpha$ with $\lambda_\alpha\in\mathbb{Q}\setminus\{0\}$. Let $\mathcal{P}=\bigl\{p_\alpha:\alpha\in\mathcal{A}\bigr\}$ be a system of polynomials on $\mathbb{Z}^d$, $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$. Then the gen-polynomials $\{w_\alpha(\mathcal{P})\}$, $\alpha\in\mathcal{B}$, are jointly well distributed.*

Proof. Let $\alpha_1,\ldots,\alpha_l\in\mathcal{B}$. Let M be divisible enough; represent the gen-polynomials $\{w_{\alpha_1}(\mathcal{P})\},\ldots,\{w_{\alpha_l}(\mathcal{P})\}$ as pp-functions of the gen-polynomials $\{v_\alpha(M^{-1}\mathcal{P})\}$, $\alpha\in\mathcal{B}$; then, since the only principal gen-monomial of w_{α_i} is $\lambda_{\alpha_i}v_{\alpha_i}$, it follows from (the proof of) Theorem 6.1 that $\{w_{\alpha_i}(\mathcal{P})\}=\bigl\{\lambda_iM^{|\alpha_i|}\{v_{\alpha_i}(M^{-1}\mathcal{P})\}+f_{\alpha_i}\bigl(\{v_\alpha(\mathcal{P})\}:\alpha\in\mathcal{B},\ |\alpha|<|\alpha_i|\bigr)\bigr\}$, $i=1,\ldots,l$. Since $\{v_\alpha(M^{-1}\mathcal{P})\}$, $\alpha\in\mathcal{B}$, are jointly well distributed and $\lambda_iM^{|\alpha_i|}\in\mathbb{Z}\setminus\{0\}$, $i=1,\ldots,l$, if M is divisible enough, we obtain that the gen-polynomials $\{w_{\alpha_i}(\mathcal{P})\}$, $i=1,\ldots,l$, are also jointly well distributed. ∎

Example. The assumption (and so, the conclusion) of Proposition 8.2 holds for the gen-polynomials w_α defined inductively by: $w_\alpha=x_\alpha$ for $\alpha\in\mathcal{A}$, $w_\alpha=\lambda_\alpha w_\gamma[w_\beta]^m$, with $0\ne\lambda_\alpha\in\mathbb{Q}$, for $\alpha=[\gamma,m\beta]\in\mathcal{B}$.

9. Nested gen-polynomials, and gen-polynomials of the product type.

Nested gen-polynomials are those in which no products of closed gen-polynomials occur: $p_1\{p_2\{p_3\}+p_4\{p_5\{p_6\}\}\}+p_7$ is nested, $p_1\{p_2\{p_3\}+p_4\{p_5\}\{p_6\}\}$ is not. Nested gen-monomials, in the variables $x_1,\ldots,x_r$, have the form $x_{i_1}\{x_{i_2}\{\ldots\{x_{i_{l-1}}\{x_{i_l}\}\}\ldots\}\}$, and the corresponding commutators in the free group $\mathcal{F}$ with generators $e_1,\ldots,e_r$ are $[e_{i_1},[e_{i_2},[\ldots,[e_{i_{l-1}},e_{i_l}]]\ldots]]$. It is well known that the set $\mathcal{N}$ of commutators of this form "spans" $\mathcal{F}$, in the sense that for any j the set $\mathcal{N}\cap\mathcal{F}_j$ spans $\mathcal{F}_j$ modulo $\mathcal{F}_{j+1}$. In accordance with this fact, nested gen-monomials also "generate" the space of all bounded gen-polynomials: using identities (5.1)–(5.7), one can show that any bounded gen-polynomial is expressible as a pp-function of bounded nested gen-monomials. (This gives another proof of Theorem 2.3, since, as it is also easy to show, any nested gen-polynomial can be read off from of a nilmanifold.) However, it is hard to describe a basis consisting of "nested" commutator expressions, and thus to choose a basis consisting of nested gen-monomials; we do not try to do this.

Gen-polynomials *of the product type* are gen-polynomials of depth 1, that is, linear combinations of gen-polynomials of the form $p_0\{p_1\}\{p_2\}\ldots\{p_k\}$, $k\ge0$. The basic gen-polynomials of the product type can be described very easily: these are the gen-monomials of the form $x_{i_0}\{x_{i_1}\}^{m_1}\{x_{i_2}\}^{m_2}\ldots\{x_{i_k}\}^{m_k}$ with $i_1<\ldots<i_k$, $k\ge0$, $m_1,\ldots,m_k\in\mathbb{N}$, and $i_0>i_1$. (These gen-monomials are the principal sub-gen-monomials of the coordinate gen-polynomials corresponding to the group $\mathcal{F}/[\mathcal{F}_2,\mathcal{F}_2]$.) It follows from Theorem 6.1 that if $u=\{u'\}$ where u' is a gen-polynomial of the product type, then the basic gen-polynomials v_α occurring in the canonical representation $u=f(\{v_\alpha\}:\alpha\in\mathcal{B})$ of u have depth ≤1, and so, are of the product type only.

10. Unbounded gen-polynomials

Any (unbounded) gen-polynomial $u\colon\mathbb{Z}^d \longrightarrow \mathbb{R}$ can be rewritten in the form $u(n) = \sum_{i=1}^r u_i n^{\nu_i} + u_0$, where ν_i are different nontrivial multiindices, and "the coefficients" u_i have the form $\sum_j \lambda_j \{\!\{w_{j,1}\}\!\} \dots \{\!\{w_{j,r_j}\}\!\}$, $\lambda_j \in \mathbb{R}$, and so, are bounded gen-polynomials.

Example.

$$(n^2+2n+3)\{\!\{\sqrt{2}n\}\!\} + (\sqrt{3}n-1)\{\!\{\sqrt{5}n^2\{\!\{\sqrt{7}n\}\!\}\}\!\}\{\!\{\sqrt{11}n^3\}\!\}$$
$$= \{\!\{\sqrt{2}n\}\!\}n^2 + \Big(2\{\!\{\sqrt{2}n\}\!\} + \sqrt{3}\{\!\{\sqrt{5}n^2\{\!\{\sqrt{7}n\}\!\}\}\!\}\{\!\{\sqrt{11}n^3\}\!\}\Big)n + \Big(3\{\!\{\sqrt{2}n\}\!\} - \{\!\{\sqrt{5}n^2\{\!\{\sqrt{7}n\}\!\}\}\!\}\{\!\{\sqrt{11}n^3\}\!\}\Big)$$

Next, using Theorem 0.2, we can represent the gen-polynomials u_i, $i = 0,\dots,r$, in the form $u_i = f_i\big(v_{\alpha_1}(M^{-1}\mathcal{P}),\dots,v_{\alpha_l}(M^{-1}\mathcal{P})\big)$, where $\mathcal{P} = \{p_\alpha : \alpha \in \mathcal{A}\}$ is a system of polynomials $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n] + \mathbb{R}$, $M \in \mathbb{N}$, $\alpha_1,\dots,\alpha_l \in \mathcal{B}(\mathcal{A})$, and $f_0,\dots,f_r$ are pp-functions on $[0,1]^l$. So, we get a representation of u in the form

$$u = \sum_{i=1}^r f_i\big(\{\!\{v_{\alpha_1}(M^{-1}\mathcal{P})\}\!\},\dots,\{\!\{v_{\alpha_l}(M^{-1}\mathcal{P})\}\!\}\big)n^{\nu_i} + f_0\big(\{\!\{v_{\alpha_1}(M^{-1}\mathcal{P})\}\!\},\dots,\{\!\{v_{\alpha_l}(M^{-1}\mathcal{P})\}\!\}\big) \tag{10.1}$$

For a set $S \subseteq \mathbb{Z}^d$, *the density* $D(S)$ of S is $\lim_{N\to\infty} \frac{|S\cap[-N,N]^d|}{(2N)^d}$, if it exists.

Lemma 10.1. *Let $f(x,n)$ be a polynomial on $[0,1]^l \times \mathbb{Z}^d$ of degree ≥ 1 with respect to n, and let a sequence $(x_n)_{n\in\mathbb{Z}^d}$ be well distributed in $[0,1]^l$. Then the sequence $f(x_n,n)$, $n \in \mathbb{Z}^d$, tends to infinity in density, that is, for any $N > 0$, the set $S = \{n \in \mathbb{Z}^d : |f(x_n,n)| < N\}$ has zero density.*

Proof. Extend f to a polynomial on $[0,1]^l \times \mathbb{R}^d$. Write f in coordinates: $f(x,z) = \sum_{\alpha\in A} h_\alpha(z)x^\alpha$, where A is a set of multiindices and for each $\alpha \in A$, h_α is a polynomial on $\mathbb{R}^d$. Let $r = \max\{\deg h_\alpha,\ \alpha \in A\}$. For each $\alpha \in A$, let $\hat{h}_\alpha$ be the homogeneous part of h_α of degree r. Let Σ be the sphere $\{\xi \in \mathbb{R}^d : |\xi| = 1\}$ and let $\Xi = \{\xi \in \Sigma : \hat{h}_\alpha(\xi) \neq 0$ for some $\alpha \in A\}$. For every $\xi \in \Sigma$ and $\alpha \in A$, $\lim_{t\to\infty} t^{-r}h_\alpha(t\xi) = \hat{h}_\alpha(\xi)$, thus the polynomials $f_t(x,\xi) = t^{-r}f(x,t\xi)$ converge as $t \to \infty$ to the polynomial $f_\xi(x) = \sum_{\alpha\in A}\hat{h}_\alpha(\xi)x^\alpha$ uniformly on $[0,1]^l \times \Sigma$.

Fix $\varepsilon > 0$. For $\xi \in \Xi$, let $R_\xi = \{x \in [0,1]^l : f_\xi(x) = 0\}$ and let $\delta_\xi > 0$ be such that the set $R_{\xi,\delta_\xi} = \{x \in [0,1]^l : |f_\xi(x)| < \delta_\xi\}$ has measure $< \varepsilon$. Let $U_\xi \subset \Xi$ be an open neighborhood of ξ such that $|f_\zeta(x) - f_\xi(x)| < \delta_\xi/3$ for all $\zeta \in U_\xi$ and $x \in [0,1]^l$, then $|f_\zeta(x)| \geq 2\delta_\xi/3$ for all $\zeta \in U_\xi$ and $x \in [0,1]^l \setminus R_{\xi,\delta_\xi}$. Let $t_\xi > 0$ be such that $t_\xi^r > 3N/\delta_\xi$ and $\big|t^{-r}f(x,t\zeta)(x) - f_\zeta(x)\big| < \delta_\xi/3$ for all $t > t_\xi$, $\zeta \in U_\xi$, and $x \in [0,1]^l$. Then $|t^{-r}f(x,t\zeta)(x)| \geq \delta_\xi/3$, and so $|f(x,t\zeta)(x)| \geq t^r\delta_\xi/3 > N$, for all $t > t_\xi$, $\zeta \in U_\xi$, and $x \in [0,1]^l \setminus R_{\xi,\delta_\xi}$. Thus, for any $t > t_\xi$ and $\zeta \in U_\xi$, $\{x \in [0,1]^l : |f(x,t\zeta)| < N\} \subseteq R_{\xi,\delta_\xi}$.

Since the sequence (x_n) is well distributed in $[0,1]^l$, for every $\xi \in \Xi$ there exists $K_\xi \in \mathbb{N}$ such that for any $K > K_\xi$ and any $m \in \mathbb{R}^d$, $\frac{1}{K^d}\big|\{n \in m + [1,K]^d : x_n \in R_{\xi,\delta_\xi}\}\big| < 2\varepsilon$. If $m \in \mathbb{R}^d$ and $K \in \mathbb{N}$ are such that $|m| > t_\xi + \sqrt{d}K$ and $m + [1,K]^d \subset \mathbb{R}_+U_\xi$, then for any $n \in m + [1,K]^d$ we have $\{x \in [0,1]^l : |f(x,n)| < N\} \subseteq R_{\xi,\delta_\xi}$. Thus, for such m and K, with $K > K_\xi$, $\frac{1}{K^d}\big|\{n \in m + [1,K]^d : |f(x_n,n)| < N\}\big| < 2\varepsilon$, and hence, $\frac{1}{K^d}\big|S \cap \big(m + [1,K]^d\big)\big| < 2\varepsilon$.

The set $E = \Sigma \setminus \Xi$ is a proper algebraic subvariety of Σ, therefore there exists a compact set $F \subset \Xi$ such that $D(\mathbb{R}_+F \cap \mathbb{Z}^d) > 1 - \varepsilon$. Let $\xi_1,\dots,\xi_k$ be such that $\bigcup_{j=1}^k U_{\xi_j} \supseteq F$ and let $T = \max_{1\leq j\leq k} t_{\xi_j}$ and $K = \max_{1\leq j\leq k} K_{\xi_j}$. Let $t > T + \sqrt{d}K$ be such that for any cube $C = m + [1,K]^d \subset \mathbb{R}_+F$ with $|m| > t$ we have $C \subset \mathbb{R}_+U_{\xi_j}$ for some j. Then for any such cube C, $\frac{1}{|C|}|S \cap C| < 2\varepsilon$. Thus, $D(S) < 3\varepsilon$. Hence, $D(S) = 0$. ■

As a corollary, we get the following proposition:

Proposition 10.2. *For the representation of u in the form* (10.1) *(assuming that the system $\mathcal{P}$ is $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n] + \mathbb{R}$), let Q be a subset of $[0,1]^l$ defined by a system of polynomial inequalities and such that $f_i|_Q$ are polynomials, $i = 0,\dots,r$, and let $P = \{n \in \mathbb{Z}^d : (\{\!\{v_{\alpha_1}(n)\}\!\},\dots,\{\!\{v_{\alpha_l}(n)\}\!\}) \in Q\}$. If $f_i|_Q$ is nonzero for at least one $i \in \{1,\dots,d\}$, then the sequence $u(n)$, $n \in P$, tends to infinity in density.*

Corollary 10.3. *If u is bounded, then, in formula* (10.1), $f_i = 0$ *a.e. for* $i = 1, \dots, r$.

It now follows from Corollary 10.3 and from (the second part of) Theorem 0.2 that

Corollary 10.4. *If $\mathcal{P}$ and M are chosen and $\mathcal{P}$ is $\mathbb{Q}$-linearly independent modulo $\mathbb{Q}[n]+\mathbb{R}$, then, in formula* (10.1)*, the pp-functions f_i are defined uniquely up to a pp-function vanishing a.e.*

We also get information about the distribution of values of an unbounded gen-polynomial:

Corollary 10.5. *Let u be a (vector-valued) generalized polynomial $\mathbb{Z}^d \longrightarrow \mathbb{R}^k$; write u in the form (10.1). Then for any set $B \subset \mathbb{R}^k$ with $0 < \mu(B) < \infty$ and $\mu(\partial B) = 0$ the density in $\mathbb{Z}^d$ of the set $u^{-1}(B)$ is equal to the measure of the set $\{x \in [0,1]^l : f_1(x) = \ldots = f_k(x) = 0,\ f_0(x) \in B\}$.*

Example. Let $u(n) = [\sqrt{2}n][\sqrt{3}n] - [\sqrt{6}n^2]$, $n \in \mathbb{Z}$. Then

$$\begin{aligned} u(n) = [\sqrt{2}n][\sqrt{3}n] - [\sqrt{6}n^2] &= -\sqrt{2}n\{\sqrt{3}n\} - \sqrt{3}n\{\sqrt{2}n\} + \{\sqrt{2}n\}\{\sqrt{3}n\} + \{\sqrt{6}n^2\} \\ &= -\big(\sqrt{2}\{\sqrt{3}n\} + \sqrt{3}\{\sqrt{2}n\}\big)n + \big(\{\sqrt{2}n\}\{\sqrt{3}n\} + \{\sqrt{6}n^2\}\big), \end{aligned}$$

and this is the canonical representation of u. "The coefficient" before n in this representation is $\sqrt{2}\{\sqrt{3}n\} + \sqrt{3}\{\sqrt{2}n\} = f_1(\{\sqrt{2}n\}, \{\sqrt{3}n\})$, where $f(x_1, x_2) = \sqrt{2}x_1 + \sqrt{3}x_2$. Since f does not vanish on a subset of positive measure in $[0,1]^2$, $u(n)$ tends to infinity in density as $n \longrightarrow \infty$.

11. Gen-polynomials of real arguments

When, instead of $\mathbb{Z}^d \longrightarrow \mathbb{R}$, we deal with gen-polynomials $\mathbb{R}^d \longrightarrow \mathbb{R}$:

1. We need a version of Theorem 2.2 that deals with "d-dimensional polynomial flows", that is, continuous polynomial actions of $\mathbb{R}^d$, instead of polynomial actions of $\mathbb{Z}^d$; such a theorem is established in [BLM].

2. The values of any nonconstant polynomial $\mathbb{R}^d \longrightarrow \mathbb{R}$ are always well distributed modulo 1, – in contrast with polynomials $\mathbb{Z}^d \longrightarrow \mathbb{R}$, which must have a non-rational coefficient to be well distributed. It follows that polynomials $\mathbb{R}^d \longrightarrow \mathbb{R}$ are jointly well distributed iff they are $\mathbb{Q}$-rationally independent modulo constants. Thus, we do not have to exclude polynomials with rational coefficients from our consideration, and, in the proof of Theorem 6.1, do not have to pass to "sublattices". Therefore, dealing with gen-polynomials of continuous variables, we no longer meet all the problems appearing when we pass from variables x_i to the variables $M^{-1}x_i$; this essentially simplifies the proofs.

As a result, we get Theorems 0.3 and 0.4.

Acknowledgment. I am grateful to Vitaly Bergelson for introducing me to the area of generalized polynomials, and for his constant interest to their life and behavior. I also thank an anonymous referee for careful reading of the preprint, correcting numerous mistakes, and making useful suggestions.

Bibliography

[AGH] L. Auslander, L. Green and F. Hahn, *Flows on homogeneous spaces*, Annals of Math. Studies, vol. **53**, Princeton Univ. Press, 1963.

[BHå] V. Bergelson and I. J. Håland, Sets of recurrence and generalized polynomials, *Convergence in Ergodic Theory and Probability*, Walter de Gruyter, 1996, 91–110.

[BLM] V. Bergelson, A. Leibman, and C.G. Moreira, From discrete time to continuous time ergodic theorems, in preparation.

[GT] B. Green and T. Tao, The quantitative behaviour of polynomial orbits on nilmanifolds, preprint. arXiv:0709.3562

[GO'B] R. Graham and K. O'Bryant, Can you hear the shape of a Beatty sequence?, preprint. arXiv:0809.0004v1

[BL] V. Bergelson and A. Leibman, Distribution of values of bounded generalized polynomials, *Acta Math.* **198** (2007), 155-230.

[Hå1] I. J. Håland, *Uniform distribution of generalized polynomials*, Ph.D. Thesis, The Ohio State University, 1992.

[Hå2] I. J. Håland, Uniform distribution of generalized polynomials, *J. Number Theory* **45** (1993), 327-366.

[Hå3] I. J. Håland, Uniform distribution of generalized polynomials of the product type, *Acta Arith.* **67** (1994), 13-27.

[HåK] I. J. Håland and D. E. Knuth, Polynomials involving the floor function, *Math. Scand.* **76** (1995), 194-200.

[L1] A. Leibman, Polynomial mappings of groups, *Israel J. of Math.* **129** (2002), 29-60.
(Erratum at http://www.math.ohio-state.edu/∼leibman/preprints/)

[L2] A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, *Erg. Th. and Dyn. Syst.* **25** (2005), 201-213.

[L3] A. Leibman, Pointwise convergence of ergodic averages for polynomial actions of $\mathbb{Z}^d$ by translations on a nilmanifold, *Erg. Th. and Dyn. Syst.* **25** (2005), 215-225.

[MKS] W. Magnus, A. Karras, and D. Solitar, *Presentations of Groups in Terms of Generators and Relations*, Dover Publications, 2004.

[Ma] A. Malcev, On a class of homogeneous spaces, *Amer. Math. Soc. Transl.* **9** (1962), 276-307.