Algebraically Solvable Problems: Describing Polynomials as Equivalent to Explicit Solutions

Uwe Schauz

Department of Mathematics
University of Tbingen, Germany
uwe.schauz@gmx.de

Submitted: Nov 14, 2006; Accepted: Dec 28, 2007; Published: Jan 7, 2008
Mathematics Subject Classifications: 41A05, 13P10, 05E99, 11C08, 11D79, 05C15, 15A15

Abstract

The main result of this paper is a coefficient formula that sharpens and generalizes Alon and Tarsi's Combinatorial Nullstellensatz. On its own, it is a result about polynomials, providing some information about the polynomial map $P|_{\mathfrak{X}_1\times\cdots\times\mathfrak{X}_n}$ when only incomplete information about the polynomial $P(X_1,\dots,X_n)$ is given.

In a very general working frame, the grid points $x \in \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n$ which do not vanish under an algebraic solution – a certain describing polynomial $P(X_1,\dots,X_n)$ – correspond to the explicit solutions of a problem. As a consequence of the coefficient formula, we prove that the existence of an algebraic solution is equivalent to the existence of a nontrivial solution to a problem. By a problem, we mean everything that "owns" both, a set $\mathcal{S}$, which may be called the *set of solutions*; and a subset $\mathcal{S}_{\text{triv}} \subseteq \mathcal{S}$, the *set of trivial solutions*.

We give several examples of how to find algebraic solutions, and how to apply our coefficient formula. These examples are mainly from graph theory and combinatorial number theory, but we also prove several versions of Chevalley and Warning's Theorem, including a generalization of Olson's Theorem, as examples and useful corollaries.

We obtain a permanent formula by applying our coefficient formula to the matrix polynomial, which is a generalization of the graph polynomial. This formula is an integrative generalization and sharpening of:

1. Ryser's permanent formula.
2. Alon's Permanent Lemma.
3. Alon and Tarsi's Theorem about orientations and colorings of graphs.

Furthermore, in combination with the Vigneron-Ellingham-Goddyn property of planar n-regular graphs, the formula contains as very special cases:

4. Scheim's formula for the number of edge n-colorings of such graphs.
5. Ellingham and Goddyn's partial answer to the list coloring conjecture.

Introduction

Interpolation polynomials $P = \sum_{\delta \in \mathbb{N}^n} P_\delta X^\delta$ on finite "grids" $\mathfrak{X} := \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n \subseteq \mathbb{F}^n$ $\mathfrak{X}$
are not uniquely determined by the interpolated maps $P|_{\mathfrak{X}} \colon x \mapsto P(x)$. One could re- $P|_{\mathfrak{X}}$
strict the partial degrees to force the uniqueness. If we only restrict the total degree to $\deg(P) \le d_1 + \cdots + d_n$, where $d_j := |\mathfrak{X}_j| - 1$, the interpolation polynomials P are still d_j
not uniquely determined, but they are partially unique. That is to say, there is one (and in general only one) coefficient in $P = \sum_{\delta \in \mathbb{N}^n} P_\delta X^\delta$ that is uniquely determined, namely P_d with $d := (d_1, \ldots, d_n)$. We prove this in Theorem 3.3 by giving a formula for this P_d
coefficient. Our coefficient formula contains Alon and Tarsi's Combinatorial Nullstellensatz [Al2, Th. 1.2], [Al3]:

$$P_d \neq 0 \quad \Longrightarrow \quad P|_{\mathfrak{X}} \not\equiv 0 \ . \tag{1}$$

This insignificant-looking result, along with Theorem 3.3 and its corollaries 3.4, 3.5 and 8.4, are astonishingly flexible in application. In most applications, we want to prove the existence of a point $x \in \mathfrak{X}$ such that $P(x) \neq 0$. Such a point x then may represent a coloring, a graph or a geometric or number-theoretic object with special properties. In the simplest case we will have the following correspondence:

$$\begin{array}{rcl} \mathfrak{X} & \longleftrightarrow & \text{Class of Objects} \\ x & \longleftrightarrow & \text{Object} \\ P(x) \neq 0 & \longleftrightarrow & \text{“Object is interesting (a solution).”} \\ P|_{\mathfrak{X}} \not\equiv 0 & \longleftrightarrow & \text{“There exists an interesting object (a solution).”} \end{array} \tag{2}$$

This explains why we are interested in the connection between P and $P|_{\mathfrak{X}}$: In general, we try to retrieve information about the polynomial map $P|_{\mathfrak{X}}$ using incomplete information about P. One important possibility is if there is (exactly) one trivial solution x_0 to a problem, so that we have the information that $P(x_0) \neq 0$. If, in this situation, we further know that $\deg(P) < d_1 + \ldots + d_n$, then Corollary 3.4 already assures us that there is a second (nontrivial) solution x, i.e., an $x \neq x_0$ in $\mathfrak{X}$ such that $P(x) \neq 0$. The other important possibility is that we do not have any trivial solutions at all, but we know that $P_d \neq 0$ and $deg(P) \le d_1 + \ldots + d_n$. In this case, $P|_{\mathfrak{X}} \not\equiv 0$ follows from (1) above or from our main result, Theorem 3.3. In other cases, we may instead apply Theorem 3.2, which is based on the more general concept from Definition 3.1 of d-leading coefficients.

In Section 4, we demonstrate how most examples from [Al2] follow easily from our coefficient formula and its corollaries. The new, quantitative version 3.3 (i) of the Combinatorial Nullstellensatz is, for example, used in Section 5, where we apply it to the matrix polynomial – a generalization of the graph polynomial – to obtain a permanent formula. This formula is a generalization and sharpening of several known results about permanents and graph colorings (see the five points in the abstract). We briefly describe how these results are derived from our permanent formula.

We show in Theorem 6.5 that it is theoretically always possible, both, to represent the solutions of a given problem $\mathcal{P}$ (see Definition 6.1) through some elements x in some grid $\mathfrak{X}$, and to find a polynomial P, with certain properties (e.g., $P_d \neq 0$ as in (1) above), that describes the problem:

$$P(x) \neq 0 \quad \Longleftrightarrow \quad \text{“}x \text{ represents a solution of } \mathcal{P}.\text{”} \tag{3}$$

We call such a polynomial P an *algebraic solution* of $\mathcal{P}$, as its existence guarantees the existence of a nontrivial solution to the problem $\mathcal{P}$.

Sections 4 and 5 contain several examples of algebraic solutions. Algebraic solutions are particularly easy to find if the problems possess exactly one trivial solution: due to Corollary 3.4, we just have to find a describing polynomial P with degree $\deg(P) < d_1 + \ldots + d_n$ in this case. Loosely speaking, Corollary 3.4 guarantees that every problem which is not too complex, in the sense that it does not require too many multiplications in the construction of P, does not possess exactly one (the trivial) solution.

In Section 7 we give a slight generalization of the (first) Combinatorial Nullstellensatz – a sharpened specialization of Hilbert's Nullstellensatz – and a discussion of Alon's original proving techniques. Note that, in Section 3 we used an approach different from Alon's to verify our main result. However, we will show that Alon and Tarsi's so-called polynomial method can easily be combined with interpolation formulas, such as our inversion formula 2.9, to reach this goal.

Section 8 contains further generalizations and results over the integers $\mathbb{Z}$ and over $\mathbb{Z}/m\mathbb{Z}$. Corollary 8.2 is a surprising relative to the important Corollary 3.4, one which works without any degree restrictions. Theorem 8.4, a version of Corollary 3.5, is a generalization of Olson's Theorem.

Most of our results hold over integral domains, though this condition has been weakened in this paper for the sake of generality (see 2.8 for the definition of integral grids). In the important case of the Boolean grid $\mathfrak{X} = \{0,1\}^n$, our results hold over arbitrary commutative rings $\mathcal{R}$. Our coefficient formulas are based on the interpolation formulas in Section 2, where we generalize known expressions for interpolation polynomials over fields to commutative rings $\mathcal{R}$. We frequently use the constants and definitions from Section 1.

For newcomers to this field, it might be a good idea to start with Section 4 to get a first impression.

We will publish two further articles: One about a sharpening of Warning's classical result about the number of simultaneous zeros of systems of polynomial equations over finite fields [Scha2], the other about the numerical aspects of using algebraic solutions to find explicit solutions, where we present two polynomial-time algorithms that find nonzeros of polynomials [Scha3].

1 Notation and constants

$\mathcal{R}$ is always a commutative ring with $1 \neq 0$. $\quad\mathcal{R}$

$\mathbb{F}_{p^k}$ denotes the field with p^k elements (p prime) and $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$. $\quad\mathbb{F}_{p^k}, \mathbb{Z}_m$

We write $p \,\lfloor\, n$ (or $n \,\rfloor\, p$) for "p *divides* n" and abbreviate $\mathcal{S}\backslash s := \mathcal{S} \setminus \{s\}$. $\quad p \,\lfloor\, n, \mathcal{S}\backslash s$

For $n \in \mathbb{N} := \{0, 1, 2, \ldots\}$ we set: $\quad\mathbb{N}$

$(n] = (0, n] := \{1, 2, \ldots, n\}$, $\quad(n]$

$[n) = [0, n) := \{0, 1, \ldots, n{-}1\}$, $\quad[n)$

$[n] = [0, n] := \{0, 1, \ldots, n\}$. (Note that $0 \in [n]$.) $\quad[n]$

For statements $\mathcal{A}$ the "Kronecker query" $?_{(\mathcal{A})}$ is defined by:

$$?_{(\mathcal{A})} := \begin{cases} 0 & \text{if } \mathcal{A} \text{ is false,} \\ 1 & \text{if } \mathcal{A} \text{ is true.} \end{cases}$$ $?_{(\mathcal{A})}$

For finite tuples (and maps) $d = (d_j)_{j\in J}$ and sets Γ we define:

$\Pi d := \prod_{j\in J} d_j$, $\Pi\Gamma := \prod_{\gamma\in\Gamma} \gamma$ and $\quad\Pi d, \Pi\Gamma$

$\Sigma d := \sum_{j\in J} d_j$, $\Sigma\Gamma := \sum_{\gamma\in\Gamma} \gamma$. $\quad\Sigma d, \Sigma\Gamma$

For maps $y, z \colon \mathfrak{X} \longrightarrow \mathcal{R}$ with finite domain we identify the map $y \colon x \longmapsto y(x)$ with the tuple $(y(x))_{x\in\mathfrak{X}} \in \mathcal{R}^{\mathfrak{X}}$. Consequently, the product with matrices $\Psi = (\psi_{\delta,x}) \in \mathcal{R}^{D\times\mathfrak{X}}$ is given by $\Psi y := \big(\sum_{x\in\mathfrak{X}} \psi_{\delta,x}\, y(x)\big)_{\delta\in D} \in \mathcal{R}^D$. $\quad y, \Psi y$

We write yz for the pointwise product, $(yz)(x) := y(x)z(x)$. If nothing else is said, y^{-1} is also defined pointwise, $y^{-1}(x) := y(x)^{-1}$, if $y(x)$ is invertible for all $x \in \mathfrak{X}$. $\quad yz, y^{-1}$

We define $\mathrm{supp}(y) := \{\, x \in \mathfrak{X} \mid y(x) \neq 0 \,\}$. $\quad\mathrm{supp}(y)$

The tensor product $\bigotimes_{j\in(n]} y_j$ of maps $y_j \colon \mathfrak{X}_j \longrightarrow \mathcal{R}$ is a map $\mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n \longrightarrow \mathcal{R}$, it is defined by $(\bigotimes_{j\in(n]} y_j)(x) := \prod_{j\in(n]} y_j(x_j)$. $\quad\otimes$

Hence, the tensor product $\bigotimes_{j\in(n]} a^j$ of tuples $a^j := (a^j_{x_j})_{x_j\in\mathfrak{X}_j}$, $j \in (n]$, is the tuple $\bigotimes_{j\in(n]} a^j := \big(\prod_{j\in(n]} a^j_{x_j}\big)_{x\in\mathfrak{X}_1\times\cdots\times\mathfrak{X}_n}$.

The tensor product $\bigotimes_{j\in(n]} \Psi^j$ of matrices $\Psi^j = (\psi^j_{\delta_j,x_j})_{\substack{\delta_j\in D_j \\ x_j\in\mathfrak{X}_j}}$, $j \in (n]$, is the matrix $\bigotimes_{j\in(n]} \Psi^j := \big(\prod_{j\in(n]} \psi^j_{\delta_j,x_j}\big)_{\substack{\delta\in D_1\times\cdots\times D_n \\ x\in\mathfrak{X}_1\times\cdots\times\mathfrak{X}_n}}$.

Tensor product and matrix-tuple multiplication go well together:

$$\begin{aligned} \Big(\bigotimes_{j\in(n]} \Psi^j\Big) \bigotimes_{j\in(n]} a^j &= \Big(\prod_{j\in(n]} \psi^j_{\delta_j,x_j}\Big)_{\substack{\delta\in D \\ x\in\mathfrak{X}}} \Big(\prod_{j\in(n]} a^j_{x_j}\Big)_{x\in\mathfrak{X}} = \Big(\sum_{x\in\mathfrak{X}} \prod_{j\in(n]} \psi^j_{\delta_j,x_j} a^j_{x_j}\Big)_{\delta\in D} \\ &= \Big(\prod_{j\in(n]} \sum_{x_j\in\mathfrak{X}_j} \psi^j_{\delta_j,x_j} a^j_{x_j}\Big)_{\delta\in D} = \bigotimes_{j\in(n]} \Big(\sum_{x_j\in\mathfrak{X}_j} \psi^j_{\delta_j,x_j} a^j_{x_j}\Big)_{\delta_j\in D_j} = \bigotimes_{j\in(n]} (\Psi^j a^j) . \end{aligned} \tag{4}$$

In the whole paper we work over Cartesian products $\mathfrak{X} := \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n$ of subsets $\mathfrak{X}_j \subseteq \mathcal{R}$ of size $d_j + 1 := |\mathfrak{X}_j| < \infty$. We define:

Definition 1.1 (d-grids $\mathfrak{X}$). $\mathfrak{X}, [d]$

$d = d(\mathfrak{X})$

For all $j \in (n]$ we define:	In n dimensions we define:
$\mathfrak{X}_j \subseteq \mathcal{R}$ is always a finite set $\neq \varnothing$.	$\mathfrak{X} := \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n \subseteq \mathcal{R}^n$ is a *d-grid* for
$\boxed{d_j = d_j(\mathfrak{X}_j) := \lvert\mathfrak{X}_j\rvert - 1}$ and	$d = d(\mathfrak{X}) := (d_1, \ldots, d_n)$.
$[d_j] := \{0, 1, \ldots, d_j\}$.	$[d] := [d_1] \times \cdots \times [d_n]$ is a d-grid in $\mathbb{Z}^n$.

The following function $N \colon \mathfrak{X} \longrightarrow \mathcal{R}$ will be used throughout the whole paper. The $\psi_{\delta,x}$ are the coefficients of the Lagrange polynomials $L_{\mathfrak{X},x}$, as we will see in Lemma 1.3 . We define:

Definition 1.2 ($N_{\mathfrak{X}}$, $\Psi_{\mathfrak{X}}$, $L_{\mathfrak{X},x}$ and e_x).
Let $\mathfrak{X} := \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n \subseteq \mathcal{R}^n$ be a d-grid, i.e., $d_j = |\mathfrak{X}_j| - 1$ for all $j \in (n]$.

e_x , $L_{\mathfrak{X},x}$

N, Ψ

For $x \in \mathfrak{X}_j$ and $\delta \in [d_j]$ we set:		For $x \in \mathfrak{X}$ and $\delta \in [d]$ we set:	
$e^j_x \colon \mathfrak{X}_j \to \mathcal{R}, \quad e^j_x(\tilde{x}) := ?_{(\tilde{x}=x)}$.		$e_x := \bigotimes_{j\in(n]} e^j_{x_j} = \left(\tilde{x} \mapsto ?_{(\tilde{x}=x)}\right)$.	
$L_{\mathfrak{X}_j\setminus x}(X) := \prod_{\hat{x}\in\mathfrak{X}_j\setminus x}(X - \hat{x})$.		$L_{\mathfrak{X},x}(X_1, \ldots, X_n) := \prod_j L_{\mathfrak{X}_j\setminus x_j}(X_j)$.	
$N_j = N_{\mathfrak{X}_j} : \mathfrak{X}_j \longrightarrow \mathcal{R}$ is defined by:		$N = N_{\mathfrak{X}} \colon \mathfrak{X} \longrightarrow \mathcal{R}$ is defined by:	
$\boxed{N_j(x) := L_{\mathfrak{X}_j\setminus x}(x)}$.		$\boxed{N := \bigotimes_{j\in(n]} N_j = \big(x \mapsto L_{\mathfrak{X},x}(x)\big)}$.	
$\Psi^j := (\psi^j_{\delta,x})_{\substack{\delta\in[d_j]\\ x\in\mathfrak{X}_j}}$ with $\psi^j_{\delta,x} := \sum_{\substack{\Gamma\subseteq\mathfrak{X}_j\setminus x\\ \lvert\Gamma\rvert=d_j-\delta}} \Pi(-\Gamma)$ and in particular $\psi^j_{d_j,x} = 1$.	(5)	$\Psi = (\psi_{\delta,x})_{\substack{\delta\in[d]\\ x\in\mathfrak{X}}} := \bigotimes_{j\in(n]} \Psi^j$, i.e., $\psi_{\delta,x} := \prod_{j\in(n]} \psi^j_{\delta_j,x_j}$ and in particular $\psi_{d,x} = 1$.	(6)

We use multiindex notation for polynomials, i.e., $X^{(\delta_1,\ldots,\delta_n)} := X_1^{\delta_1} \cdots X_n^{\delta_n}$ and we $X^{(\delta_1,\ldots,\delta_n)}$
define $P_\delta = (P)_\delta$ to be the coefficient of X^δ in the standard expansion of $P \in \mathcal{R}[X] :=$ $P_\delta = (P)_\delta$
$\mathcal{R}[X_1, \ldots, X_n]$. That means $P = P(X) = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta$ and $(X^\varepsilon)_\delta = ?_{(\delta=\varepsilon)}$. $\mathcal{R}[X]$

Conversely, for tuples $P = (P_\delta)_{\delta \in \mathcal{D}} \in \mathcal{R}^{\mathcal{D}}$, we set $P(X) := \sum_{\delta \in \mathcal{D}} P_\delta X^\delta$. In this way we identify the set of tuples $\mathcal{R}^{[d]} = \mathcal{R}^{[d_1]\times\cdots\times[d_n]}$ with $\mathcal{R}[X^{\le d}]$, the set of polynomials $P = \sum_{\delta \le d} P_\delta X^\delta$ with restricted partial degrees $\deg_j(P) \le d_j$. It will be clear from the context whether we view P as a tuple (P_δ) in $\mathcal{R}^{[d]}$, a map $[d] \longrightarrow \mathcal{R}$ or a polynomial $P(X)$ in $\mathcal{R}[X^{\le d}]$. $P(X)|_{\mathfrak{X}}$ stands for the map $\mathfrak{X} \longrightarrow \mathcal{R}$, $x \longmapsto P(x)$.

$P(X)$ $\mathcal{R}^{[d]}$ $\mathcal{R}[X^{\le d}]$ $P(X)|_{\mathfrak{X}}$

We have introduced the following four related or identified objects:

Maps:	Tuples:	Polynomials:	Polynomial Maps:	
$\delta \mapsto P_\delta$,	$P = (P_\delta)$	$P(X) = \sum P_\delta X^\delta$	$P(X)\|_{\mathfrak{X}}\colon x \mapsto P(x)$,	
$[d] \to \mathcal{R}$	$\in \mathcal{R}^{[d]}$	$\in \mathcal{R}[X^{\le d}]$	$\mathfrak{X} \to \mathcal{R}$	(7)

With these definitions we get the following important formula:

Lemma 1.3 (Lagrange polynomials).

$$\boxed{(\Psi e_x)(X) \;:=\; \sum_{\delta \in [d]} \psi_{\delta,x} X^\delta \;=\; \prod_{j \in (n]} \prod_{\hat{x}_j \in \mathfrak{X}_j \setminus x_j} (X_j - \hat{x}_j) \;=:\; L_{\mathfrak{X},x}}\;.$$

Proof. We start with the one-dimensional case. Assume $x \in \mathfrak{X}_j$, then

$$\begin{aligned}
(\Psi^j e^j_x)(X_j) &= \Big(\sum_{\delta \in [d_j]} \psi^j_{\delta,x} X_j^\delta\Big) \\
&= \sum_{\delta \in [d_j]} \sum_{\substack{\Gamma \subseteq \mathfrak{X}_j \setminus x \\ |\Gamma| = d_j - \delta}} X_j^\delta \, \Pi(-\Gamma) \\
&= \sum_{\hat{\Gamma} \subseteq \mathfrak{X}_j \setminus x} X_j^{|(\mathfrak{X}_j \setminus x)\setminus \hat{\Gamma}|} \, \Pi(-\hat{\Gamma}) \\
&= \prod_{\hat{x} \in \mathfrak{X}_j \setminus x} (X_j - \hat{x})\;.
\end{aligned} \tag{8}$$

In n dimensions and for $x \in \mathfrak{X}$ we conclude:

$$\begin{aligned}
(\Psi e_x)(X) &= \Big(\big(\textstyle\bigotimes_j \Psi^j\big) \textstyle\bigotimes_j e^j_{x_j}\Big)(X) \\
&\overset{(4)}{=} \Big(\textstyle\bigotimes_j \big(\Psi^j e^j_{x_j}\big)\Big)(X) \\
&= \textstyle\prod_j \big((\Psi^j e^j_{x_j})(X_j)\big) \\
&\overset{(8)}{=} \prod_{j \in (n]} \prod_{\hat{x}_j \in \mathfrak{X}_j \setminus x_j} (X_j - \hat{x}_j)\;.
\end{aligned} \tag{9}$$

□

We further provide the following specializations of the ubiquitous function $N \in \mathcal{R}^{\mathfrak{X}}$, $N(x) = \prod_{j \in (n]} N_j(x_j)$:

Lemma 1.4. *Let* $E_l := \{\, c \in \mathcal{R} \;\vert\; c^l = 1 \,\}$ *denote the set of the* l^{th} *roots of unity in* $\mathcal{R}$ *. For* $x \in \mathfrak{X}_j \subseteq \mathcal{R}$ *hold:*

(i) If $\mathfrak{X}_j = E_{d_j+1}$ *(* $|E_{d_j+1}| = d_j + 1$ *) and if* $\mathcal{R}$ *is an integral domain:*
$$\boxed{N_j(x) \;=\; (d_j+1)\, x^{-1}\,.}$$

(ii) If $\mathfrak{X}_j \uplus \{0\}$ *is a finite subfield of* $\mathcal{R}$ *:*
$$\boxed{N_j(x) \;=\; -x^{-1}\,.}$$

(iii) If $\mathfrak{X}_j = E_{d_j} \uplus \{0\}$ *(* $|E_{d_j}| = d_j$ *) and if* $\mathcal{R}$ *is an integral domain:*
$$\boxed{N_j(x) \;=\; \begin{cases} d_j 1 & \text{for } x \neq 0\,, \\ -1 & \text{for } x = 0\,. \end{cases}}$$

(iv) If $\mathfrak{X}_j$ *is a finite subfield of* $\mathcal{R}$ *:*
$$\boxed{N_j(x) \;=\; -1 \;\;.}$$

(v) If $\mathfrak{X}_j = \{0, 1, \ldots, d_j\} \subseteq \mathbb{Z}$ *:*
$$\boxed{N_j(x) \;=\; (-1)^{d_j+x}\, d_j!\, \tbinom{d_j}{x}^{-1}.}$$

(vi) For $\alpha \in \mathcal{R}$ *we have:*
$$\boxed{N_{\mathfrak{X}_j+\alpha}(x+\alpha) \;=\; N_{\mathfrak{X}_j}(x) \;\;.}$$

Proof. For finite subsets $\mathcal{D} \subseteq \mathcal{R}$ we define
$$L_{\mathcal{D}}(X) \;:=\; \prod_{\hat{x}\in\mathcal{D}} (X - \hat{x}) \;\;. \tag{10}$$
It is well-known that, if E_l contains l elements and lies in an integral domain,
$$L_{E_l}(X) \;=\; \prod_{\hat{x}\in E_l} (X - \hat{x}) \;=\; X^l - 1 \;=\; (X-1)(X^{l-1} + \cdots + X^0) \;\;. \tag{11}$$
Thus
$$L_{E_l\setminus 1}(1) \;=\; \frac{\prod_{\hat{x}\in E_l}(X-\hat{x})}{X-1}\,\Big|_{X=1} = \frac{X^l-1}{X-1}\,\Big|_{X=1} = X^{l-1} + \cdots + X^0\,\Big|_{X=1} = l1 \;. \tag{12}$$
Using this, we get for $x \in E_l$
$$L_{E_l\setminus x}(x) \;=\; L_{x(E_l\setminus 1)}(x) \;=\; \prod_{\hat{x}\in E_l\setminus 1} (x - x\hat{x}) \;=\; x^{l-1} L_{E_l\setminus 1}(1) \;=\; l x^{-1} \;\;. \tag{13}$$
This gives (i) with $l = |\mathfrak{X}_j| = d_j + 1$.

Part (ii) is a special case of part (i), where $\mathfrak{X}_j \;=\; F_{p^k}\setminus 0 \;=\; E_{p^k-1}$ and where consequently $d_j + 1 = |\mathfrak{X}_j| = (p^k - 1) \equiv -1 \pmod{p}$.

To get $N_j(x) \;=\; L_{\{0\}\uplus E_l\setminus x}(x)$ with $x \neq 0$ in part (iii) and part (iv) we multiply Equation (13) with $x - 0$ and use $l = |\mathfrak{X}_j| - 1 = p^k - 1 \equiv -1 \pmod{p}$ for part (iv) and $l = |\mathfrak{X}_j| - 1 = d_j$ for part (iii). For $x = 0$ we obtain in part (iii) and part (iv)
$$N_j(0) \;=\; L_{E_l}(0) \;=\; \prod_{\hat{x}\in E_l} (-\hat{x}) \;=\; -\prod_{\hat{x}\in E_l\setminus\{1,-1\}} (-\hat{x}) \;=\; -1 \;\;, \tag{14}$$

since each subset $\{\hat{x},\hat{x}^{-1}\}\subseteq E_l\setminus\{1,-1\}$ contributes $(-\hat{x})(-\hat{x}^{-1})=1$ to the product – as $\hat{x}\neq\hat{x}^{-1}$, since $\hat{x}^2-1=0$ holds only for $\hat{x}=\pm 1$ – and $E_l\setminus\{1,-1\}$ is partitioned by such subsets. This completes the proofs of parts (iii) and (iv).

We now turn to part (v):

$$N_j(x) \;=\; \Big(\prod_{0\le\hat{x}<x}(x-\hat{x})\Big)\prod_{x<\hat{x}\le d_j}(x-\hat{x}) \;=\; x!\,(d_j-x)!\,(-1)^{d_j-x} \;=\; (-1)^{d_j+x}\,d_j!\binom{d_j}{x}^{-1}. \tag{15}$$

Part (vi) is trivial. □

2 Interpolation polynomials and inversion formulas

This section may be skipped at a first reading; the only things you need from here to understand the rest of the paper are:

– the fact that grids $\mathfrak{X}:=\mathfrak{X}_1\times\cdots\times\mathfrak{X}_n\subseteq\mathcal{R}^n$ over integral domains R are always *integral grids*, in the sense of Definition 2.5, and
– the inversion formula 2.9, which is, in this case, just the well-known interpolation formula for polynomials applied to polynomial maps $P|_{\mathfrak{X}}$.

The rest of this section is concerned with providing some generality that is not really used in the applications of this paper.

We have to investigate the canonical homomorphism $\varphi\colon P\longmapsto P|_{\mathfrak{X}}$ that maps polynomials P to polynomial maps $P|_{\mathfrak{X}}\colon x\mapsto P(x)$ on a fixed d-grid $\mathfrak{X}\subseteq\mathcal{R}^n$. As the monic polynomial $L_j=L_{\mathfrak{X}_j}(X_j):=\prod_{\hat{x}\in\mathfrak{X}_j}(X_j-\hat{x})$ maps all elements of $\mathfrak{X}_j$ to 0, we may replace each given polynomial P by any other polynomial of the form $P+\sum_{j\in(n]}H_jL_j$ without changing its image $P|_{\mathfrak{X}}$. By applying such modifications, we may assume that P has partial degrees $\deg_j(P)\le|\mathfrak{X}_j|-1=d_j$ (see Example 7.1 for an illustration of this method). Hence the image of φ does not change if we regard φ as a map on $\mathcal{R}[X^{\le d}]$ (which we identify with $\mathcal{R}^{[d]}$ by $P\mapsto(P_\delta)_{\delta\in[d]}$). The resulting map

φ

L_j

φ

$$\varphi\colon\mathcal{R}[X^{\le d}]=\mathcal{R}^{[d]}\longrightarrow\mathcal{R}^{\mathfrak{X}},\;\; P\longmapsto P|_{\mathfrak{X}}:=(x\mapsto P(x)) \tag{16}$$

is in the most important cases an isomorphism or at least a monomorphism, as we will see in this section. In general, however, the situation is much more complicated, we give a short example and make a related, more general remark:

Example 2.1. Over $\mathcal{R}=\mathbb{Z}_6:=\mathbb{Z}/6\mathbb{Z}$ we have $X^3|_{\mathbb{Z}_6}=X|_{\mathbb{Z}_6}$ and $3X^2|_{\mathbb{Z}_6}=3X|_{\mathbb{Z}_6}$, so that each polynomial map $\mathfrak{X}:=\mathbb{Z}_6\longrightarrow\mathbb{Z}_6$ can be represented by a polynomial of the form aX^2+bX+c, with $a\in\{0,1,-1\}$. Hence the corresponding $3\cdot 6^2$ distinct maps are the only maps out of the 6^6 maps from $\mathfrak{X}=\mathbb{Z}_6$ to $\mathbb{Z}_6$ that can be represented by polynomials at all. This simple example shows also that the kernel $\ker(\varphi)$ may be very complicated even in just one dimension.

Remark 2.2. There are some general results for the rings $\mathcal{R} = \mathbb{Z}_m$ of integers mod m:

– In [MuSt] a system of polynomials in $\mathbb{Z}_m[X_1, \ldots, X_n]$ is given that represent all polynomial maps $\mathbb{Z}_m^n \longrightarrow \mathbb{Z}_m$ and the number of all such maps is determined.

– In [Sp] it is shown that the Newton algorithm can be used to determine interpolation polynomials, if they exist. The "divided differences" in this algorithm are, like the interpolation polynomials themselves, not uniquely determined over arbitrary commutative rings, and exist if and only if interpolation polynomials exist.

But back to the main subject. In which situations does $\varphi \colon P \longmapsto P|_{\mathfrak{X}}$ become an isomorphism, or equivalently, when does its representing matrix Φ possess an inverse? Over commutative rings $\mathcal{R}$, square matrices $\Phi \in \mathcal{R}^{m\times m}$ with nonvanishing determinant do not have an inverse, in general. However, there is the matrix $\mathrm{Adj}(\Phi)$ – the adjoint or cofactor matrix – that comes close to being an inverse: $\mathrm{Adj}(\Phi)$

$$\Phi\,\mathrm{Adj}(\Phi) \;=\; \mathrm{Adj}(\Phi)\,\Phi \;=\; \det(\Phi)\,\mathbf{1}\ . \tag{17}$$

In our concrete situation, where $\Phi \in \mathcal{R}^{\mathfrak{X}\times[d]}$ is the matrix of φ (a tensor product of Vandermonde matrices), we work with Ψ (from Definition 1.2) instead of the adjoint matrix $\mathrm{Adj}(\Phi)$. Ψ comes closer than $\mathrm{Adj}(\Phi)$ to being a right inverse of Φ. The following theorem shows that Φ Ψ

$$\Phi\Psi \;=\; \big(N(x)\,?_{(\tilde{x}=x)}\big)_{\tilde{x},x\in\mathfrak{X}}\ , \tag{18}$$

and the entries $N(x)$ of this diagonal matrix divide the entries $\det(\Phi)$ of $\Phi\,\mathrm{Adj}(\Phi)$, so that $\Phi\Psi$ is actually closer than $\Phi\,\mathrm{Adj}(\Phi)$ to the unity matrix (provided we identify the column indices $x \in \mathfrak{X}$ and row indices $\delta \in [d]$ in some way with the numbers $1, 2, \ldots, |\mathfrak{X}| = |[d]|$, in order to make $\det(\Phi)$ and $\mathrm{Adj}(\Phi)$ defined).

However, we used the matrix $\Phi \in \mathcal{R}^{\mathfrak{X}\times[d]}$ of $\varphi \colon P \longmapsto P|_{\mathfrak{X}}$ here just to explain the role of Ψ. In what follows, we do not use it any more; rather, we prefer notations with "φ" or "$|_{\mathfrak{X}}$." For maps/tuples $y \in \mathcal{R}^{\mathfrak{X}}$, we write $(\Psi y)(X) \in \mathcal{R}[X^{\le d}]$, as already defined, for the polynomial whose coefficients form the tuple $\Psi y \in \mathcal{R}^{[d]}$, i.e., $(\Psi y)(X) = \Psi y$ by identification. We have: Φ, φ $(\Psi y)(X)$

Theorem 2.3 (Interpolation). *For maps $y \colon \mathfrak{X} \longrightarrow \mathcal{R}$,*

$$\boxed{(\Psi y)(X)|_{\mathfrak{X}} \;=\; Ny}\ .$$

Proof. As both sides of the equation are linear in y, it suffices to prove the equation for the maps $y = e_{\tilde{x}}$, where $\tilde{x}$ ranges over $\mathfrak{X}$. Now we see that, at each point $x \in \mathfrak{X}$, we actually have

$$(\Psi e_{\tilde{x}})(X)|_{\mathfrak{X}}(x) \;\overset{1.3}{=}\; L_{\mathfrak{X},\tilde{x}}(x) \;=\; N(x)\,?_{(x=\tilde{x})} \;=\; (Ne_{\tilde{x}})(x)\ . \tag{19}$$

□

With this theorem, we are able to characterize the situations in which $\varphi\colon P \longmapsto P|_{\mathfrak{X}}$ is an isomorphism:

Equivalence and Definition 2.4 (Division grids). *We call a d-grid* $\mathfrak{X} \subseteq \mathcal{R}^n$ *a* division grid *(over* $\mathcal{R}$ *) if it has the following equivalent properties:*

(i) *For all* $j \in (n]$ *and all* $x, \tilde{x} \in \mathfrak{X}_j$ *with* $x \neq \tilde{x}$ *the difference* $x - \tilde{x}$ *is invertible.*

(ii) $N = N_{\mathfrak{X}}$ *is pointwise invertible, i.e., for all* $x \in \mathfrak{X}$*,* $N(x)$ *is invertible.*

(iii) ΠN *is invertible.*

(iv) $\varphi\colon \mathcal{R}[X^{\leq d}] = \mathcal{R}^{[d]} \longrightarrow \mathcal{R}^{\mathfrak{X}}$ *is bijective.*

Proof. The equivalence of (i),(ii) and (iii) follows from the Definition 1.2 of N, the definition $\Pi N = \prod_{x\in\mathfrak{X}} N(x)$ and the associativity and commutativity of $\mathcal{R}$.

Assuming (ii), it follows from Theorem 2.3 that $y \longmapsto (\Psi(N^{-1}y))(X)$ is a right inverse of $\varphi\colon P \longmapsto P|_{\mathfrak{X}}$:

$$y \longmapsto (\Psi(N^{-1}y))(X) \overset{\varphi}{\longmapsto} N(N^{-1}y) = y \ . \tag{20}$$

It is even a two-sided inverse, since square matrices Φ over a commutative ring $\mathcal{R}$ are invertible from both sides if they are invertible at all (since $\Phi\,\mathrm{Adj}(\Phi) = \det(\Phi)\mathbb{1}$). This gives (iv).

Now assume (iv) holds; then for all $x \in \mathfrak{X}$,

$$\left(\psi_{\delta,x}\right)_{\delta\in[d]} \;=\; \Psi e_x \overset{2.3}{=} \varphi^{-1}(Ne_x) \;=\; N(x)\,\varphi^{-1}(e_x) \ , \tag{21}$$

and in particular,

$$1 \overset{(6)}{=} \psi_{d,x} \;=\; N(x)\left(\varphi^{-1}(e_x)\right)_d \ . \tag{22}$$

Thus the $N(x)$ are invertible and that is (ii). $\square$

If $\varphi\colon \mathcal{R}[X^{\leq d}] \longrightarrow \mathcal{R}^{\mathfrak{X}}$ is an isomorphism, then $\varphi^{-1}(y)$ is the unique polynomial in $\mathcal{R}[X^{\leq d}]$ that interpolates a given map $y \in \mathcal{R}^{\mathfrak{X}}$, so that, by Theorem 2.3, it has to be the polynomial $\Psi(N^{-1}y) \in \mathcal{R}^{[d]} = \mathcal{R}[X^{\leq d}]$. This yields the following result: φ^{-1}

Theorem 2.5 (Interpolation formula). *Let* $\mathfrak{X}$ *be a division grid (e.g., if* $\mathcal{R}$ *is a field or if* $\mathfrak{X}$ *is the* Boolean grid $\{0,1\}^n$ *). For* $y \in \mathcal{R}^{\mathfrak{X}}$*,*

$$\boxed{\varphi^{-1}(y) \;=\; \Psi(N^{-1}y)}\ .$$

This theorem can be found in [Da, Theorem 2.5.2], but just for fields $\mathcal{R}$ and in a different representation (with $\varphi^{-1}(y)$ as a determinant).

Additionally, if $\mathfrak{X}$ is not a division grid, we may apply the canonical localization homomorphism π $\mathcal{S}, \mathcal{R}_N$

$$\pi\colon \mathcal{R} \longrightarrow \mathcal{R}_N := \mathcal{S}^{-1}\mathcal{R}\ ,\ \ r \longmapsto r^\pi := \tfrac{r}{1} \quad \text{with} \quad \mathcal{S} := \{\, (\Pi N)^m \mid m \in \mathbb{N} \,\}\ , \tag{23}$$

and exert our theorems in this situation. As π and $\mathcal{R}_N$ have the universal property with respect to the invertibility of $(\Pi N)^\pi$ in $\mathcal{R}_N$ (as required in 2.4(iii)), π and $\mathcal{R}_N$ are the best choices. This means specifically that if $(\Pi N)^\pi$ is not invertible in the codomain $\mathcal{R}_N$ of π, then no other homomorphism π' has this property, either. In general, π does not have this property itself: By definition,

$$\frac{r_1}{s_1} = \frac{r_2}{s_2} \quad :\Longleftrightarrow \quad \exists\, s \in \mathcal{S}\colon\ s\, r_1 s_2 \,=\, s\, r_2 s_1\ , \tag{24}$$

and hence $\ker(\pi)$

$$\ker(\pi) \,=\, \{\, r \in \mathcal{R} \mid \exists\, m \in \mathbb{N}\colon (\Pi N)^m\, r = 0 \,\}\ , \tag{25}$$

so that $(\Pi N)^\pi = 0$ is possible. Localization works in the following situation:

Equivalence and Definition 2.6 (Affine grids). *We call a d-grid $\mathfrak{X} \subseteq \mathcal{R}^n$* affine *(over $\mathcal{R}$) if it has the following equivalent properties:*

(i) ΠN *is not nilpotent.*

(ii) $\pi \neq 0$.

(iii) $(\Pi N)^\pi$ *is invertible in* $\mathcal{R}_N$.

(iv) $\pi \neq 0$ *is injective on the* $\mathfrak{X}_j$
and hence induces a bijection $\mathfrak{X} \longrightarrow \mathfrak{X}^\pi := \mathfrak{X}_1^\pi \times \cdots \times \mathfrak{X}_n^\pi$. $\mathfrak{X}^\pi$

Proof. Part (ii) is equivalent to $1^\pi \neq 0$, and this means that $s1 \neq 0$ for all s in the multiplicative system $\mathcal{S} = \{\, (\Pi N)^m \mid m \in \mathbb{N} \,\}$; thus $(i) \Longleftrightarrow (ii)$.

Of cause $(\Pi N)^\pi \frac{1}{\Pi N} = \frac{1}{1}$ is the unity in $\mathcal{R}_N$, provided $\frac{1}{1} = 1^\pi \neq 0$; thus $(ii) \Longrightarrow (iii)$.

If (iii) holds then $(\Pi N)^\pi$ and its factors $(x_j - \tilde{x}_j)^\pi$ do not vanish; thus $(iii) \Longrightarrow (iv)$.

Finally, the implication $(iv) \Longrightarrow (ii)$ is trivial. □

If $\mathfrak{X} \subseteq \mathcal{R}^n$ is affine, then $\mathfrak{X}^\pi := \mathfrak{X}_1^\pi \times \cdots \times \mathfrak{X}_n^\pi \subseteq \mathcal{R}_N^n$ is a division d-grid over $\mathfrak{X}^\pi$
$\mathcal{R}_N$ by 2.6 (iv), 2.6 (iii) and 2.4 (iii). Now, Theorem 2.5 applied to $y := P^\pi|_{\mathfrak{X}^\pi}$ with P^π
$P^\pi = \sum_{\delta \in [d]} P_\delta^\pi X^\delta$ yields

$$P^\pi \,=\, \Psi_{\mathfrak{X}^\pi}\big((N_{\mathfrak{X}^\pi})^{-1} (P^\pi|_{\mathfrak{X}^\pi}) \big)\ , \tag{26}$$

along with the associated constants $N_{\mathfrak{X}^\pi} \in \mathcal{R}_N^{\mathfrak{X}^\pi}$ and $\Psi_{\mathfrak{X}^\pi} \in \mathcal{R}_N^{[d] \times \mathfrak{X}^\pi}$ of $\mathfrak{X}^\pi$.

With componentwise application of $\pi\colon r\mapsto \frac{r}{1}$ to $P|_{\mathfrak{X}}$, $N\in\mathcal{R}^{\mathfrak{X}}$ and to $\Psi\in\mathcal{R}^{[d]\times\mathfrak{X}}$ so that $(P|_{\mathfrak{X}})^\pi$, $N^\pi\in\mathcal{R}_N^{\mathfrak{X}}$ and $\Psi^\pi\in\mathcal{R}_N^{[d]\times\mathfrak{X}}$, we obtain: N^π, Ψ^π

Theorem 2.7 (Inversion formula). *Let $\mathfrak{X}$ be affine (e.g., if $\mathcal{R}$ does not possess nilpotent elements). For $P\in\mathcal{R}[X^{\le d}]=\mathcal{R}^{[d]}$,*

$$\boxed{P^\pi \;=\; \Psi^\pi\big((N^\pi)^{-1}(P|_{\mathfrak{X}})^\pi\big)}\ .$$

If π is injective on its whole domain $\mathcal{R}$ then $\mathcal{R}$ is a subring of $\mathcal{R}_N$ and we may omit π in formula 2.7. In fact, we will see that this is precisely when φ is injective, as seen in the following characterization:

Equivalence and Definition 2.8 (Integral grids). *We call a d-grid $\mathfrak{X}\subseteq\mathcal{R}^n$* integral *(over $\mathcal{R}$) if it has the following, equivalent properties:*

(i) For all $j\in(n]$ and all $x,\tilde{x}\in\mathfrak{X}_j$ with $x\neq\tilde{x}$, $x-\tilde{x}$ is not a zero divisor.

(ii) For all $x\in\mathfrak{X}$, $N(x)$ is not a zero divisor.

(iii) ΠN is not a zero divisor.

(iv) π is injective ($\mathcal{R}\subseteq\mathcal{R}_N$).

(v) $\varphi\colon\mathcal{R}[X^{\le d}]=\mathcal{R}^{[d]}\longrightarrow\mathcal{R}^{\mathfrak{X}}$ is injective.

Proof. The equivalence of (i),(ii) and (iii) follows from the Definition 1.2 of N, the definition $\Pi N=\prod_{x\in\mathfrak{X}}N(x)$ and the associativity and commutativity of $\mathcal{R}$.

As already mentioned $\ker(\pi)=\{\,r\in\mathcal{R}\mid\exists\, m\in\mathbb{N}\colon (\Pi N)^m r=0\,\}$, so $(iii)\Longrightarrow(iv)$.

If (iv) holds, then ΠN is invertible in $\mathcal{R}_N$. By Equivalence 2.4, it follows that $\varphi\colon\mathcal{R}_N[X^{\le d}]\longrightarrow\mathcal{R}_N^{\mathfrak{X}}$ is bijective, so that $(iv)\Longrightarrow(v)$.

Now suppose that (ii) does not hold, so that there are a point $x\in\mathfrak{X}$ and a constant $M\in\mathcal{R}\backslash 0$ with

$$MN(x)\;=\;0\ . \tag{27}$$

Then

$$P\;:=\;\Psi(Me_x)\;\neq\;0\ , \tag{28}$$

as

$$P_d\;=\;M(\Psi(e_x))_d\;=\;M\psi_{d,x}\overset{(6)}{=}M\;\neq\;0\ . \tag{29}$$

However,

$$\varphi(P)\overset{2.3}{=}NMe_x\;=\;MN(x)e_x\;\equiv\;0\ , \tag{30}$$

so that (v) does not hold, either. Thus $(v)\Longrightarrow(ii)$. ☐

Any integral grid $\mathfrak{X}$ over $\mathcal{R}$ is, in fact, a division grid over $\mathcal{R}_N \supseteq \mathcal{R}$, since ΠN becomes invertible in $\mathcal{R}_N$. Formula 2.5 applied to $y := P|_{\mathfrak{X}}$ yields the following specialization of Theorem 2.7:

Theorem 2.9 (Inversion formula). *Let $\mathfrak{X}$ be integral (e.g., if $\mathcal{R}$ is an integral domain). For $P \in \mathcal{R}[X^{\le d}] = \mathcal{R}^{[d]}$,*

$$\boxed{P \;=\; \Psi(N^{-1}P|_{\mathfrak{X}})}\;.$$

From the case $P = 1$, we see that $N^{-1}P|_{\mathfrak{X}}$ inside this formula does not lie in $\mathcal{R}^{\mathfrak{X}}$ in general (of course $N^{-1}P|_{\mathfrak{X}} \in \mathcal{R}_N^{\mathfrak{X}}$). This also shows that, in general, the maps of the form Ny, with $y \in \mathcal{R}^{\mathfrak{X}}$, in Theorem 2.3 are not the only maps that can be represented by polynomials over $\mathcal{R}$, i.e., $\{\, Ny \mid y \in \mathcal{R}^{\mathfrak{X}} \,\} \subsetneq \mathrm{Im}(\varphi)$. However, the maps of the form Ny are exactly the linear combinations of Lagrange's polynomial maps $Ne_x = L_{\mathfrak{X},x}|_{\mathfrak{X}}$ over the grid $\mathfrak{X}$; and if we view, a bit more generally, Lagrange polynomials $L_{\tilde{\mathfrak{X}},x}$ over subgrids $\tilde{\mathfrak{X}} = \tilde{\mathfrak{X}}_1 \times \cdots \times \tilde{\mathfrak{X}}_n \subseteq \mathfrak{X}$, then the maps of the form $L_{\tilde{\mathfrak{X}},x}|_{\mathfrak{X}}$ span $\mathrm{Im}(\varphi)$, as one can easily show.

On the other hand, in general, $\mathrm{Im}(\varphi) \subsetneq \mathcal{R}^{\mathfrak{X}}$, so that not every map $y \in \mathcal{R}^{\mathfrak{X}}$ can be interpolated over $\mathcal{R}$. If $\mathfrak{X}$ is integral, then interpolation polynomials exist over the bigger ring $\mathcal{R}_N$. The univariate polynomials $\binom{X}{k} := \frac{X(X-1)\cdots(X-k+1)}{k!}$, for example, describe integer-valued maps (on the whole domain $\mathbb{Z}$), but do not lie in $\mathbb{Z}[X]$. More information about such "overall" integer-valued polynomials over quotient fields can be found, for example, in [CCF] and [CCS], and in the literature cited there.

The reader might find it interesting that the principle of inclusion and exclusion follows from Theorem 2.9 as a special case:

Proposition 2.10 (Principle of inclusion and exclusion).
Let $\mathfrak{X} := \{0,1\}^n = [d]$ and $x \in \mathfrak{X}$; then $x^\delta = ?_{(\delta \le x)}$ for all $\delta \in [d]$. Thus, for arbitrary $P = (P_\delta) \in \mathcal{R}^{[d]} = \mathcal{R}[X^{\le d}]$,

$$P(x) \;=\; \sum\nolimits_{\delta \le x} P_\delta \;. \tag{31}$$

Formula 2.9 is the Möbius inversion to Equation (31):

$$\begin{aligned} P_\delta \;&\overset{2.9}{=}\; \sum_{x \in [d]} \psi_{\delta,x} N^{-1}(x) P(x) \\ &\overset{1.2}{=}\; \sum_{x \in [d]} \Big[\prod_{j \in (n]} ?_{(x_j \le \delta_j)}\,(-1)^{1-\delta_j}\Big]\Big[\prod_{j \in (n]} (-1)^{1-x_j}\Big] P(x) \\ &=\; \sum\nolimits_{x \le \delta} (-1)^{\Sigma(\delta - x)} P(x) \;. \end{aligned} \tag{32}$$

3 Coefficient formulas – the main results

The applications in this paper do not start with a map $y \in \mathcal{R}^{\mathfrak{X}}$ that has to be interpolated by a polynomial P. Rather, we start with a polynomial P, or with some information about a polynomial $P \in \mathcal{R}[X]$, which describes the very map $y := P|_{\mathfrak{X}}$ that we would like to understand. Normally, we will not have complete information about P, so that we do not usually know all coefficients P_δ of P. However, there may be a coefficient P_δ in $P = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta$ that, on its own, allows conclusions about the map $P|_{\mathfrak{X}}$. We define (see also figure 1 below):

Definition 3.1. Let $P = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta \in \mathcal{R}[X]$ be a polynomial. We call a multiindex $\varepsilon \le d \in \mathbb{N}^n$ *d-leading* in P if for each monomial X^δ in P, i.e., each δ with $P_\delta \neq 0$, holds either

- (case 1) $\delta = \varepsilon$; *or*
- (case 2) there is a $j \in (n]$ such that $\delta_j \neq \varepsilon_j$ but $\delta_j \le d_j$.

Note that the multiindex d is d-leading in polynomials P with $\deg(P) \le \Sigma d$. In this situation, case 2 reduces to "there is a $j \in (n]$ such that $\delta_j < d_j$," and, as $\Sigma\delta \le \Sigma d$ for all X^δ in P, we can conclude:

$$\text{“not case 2”} \implies \delta \ge d \implies \delta = d \implies \text{“case 1”} . \tag{33}$$

Thus d really is d-leading in P (see also figure 2 on page 28). Of course, if all partial degrees are restricted by $\deg_j(P) \le d_j$ then all multiindices $\delta \le d$ are d-leading. Figure 1 (below) shows a nontrivial example $P \in \mathcal{R}[X_1, X_2]$. The monomials X^δ of P ($P_\delta \neq 0$), and the $2^n - 1 = 3$ "forbidden areas" of each of the two d-leading multiindices, are marked.

In what follows, we examine how the preconditions of the inversion formula 2.9 may be weakened. It turns out that formula 2.9 holds without further degree restrictions for the d-leading coefficients P_ε of P. The following theorem is a generalization and a sharpening of Alon and Tarsi's (second) Combinatorial Nullstellensatz [Al2, Theorem 1.2]:

Theorem 3.2 (Coefficient formula). *Let $\mathfrak{X}$ be an integral d-grid. For each polynomial $P = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta \in \mathcal{R}[X]$ with d-leading multiindex $\varepsilon \le d \in \mathbb{N}^n$,*

(i) $\boxed{P_\varepsilon = (\Psi(N^{-1}P|_{\mathfrak{X}}))_\varepsilon}$ $\quad$ *(* $= \sum_{x\in\mathfrak{X}} \psi_{\varepsilon,x} N(x)^{-1} P(x)$ *), and*

(ii) $\boxed{P_\varepsilon \neq 0 \implies P|_{\mathfrak{X}} \not\equiv 0}$.

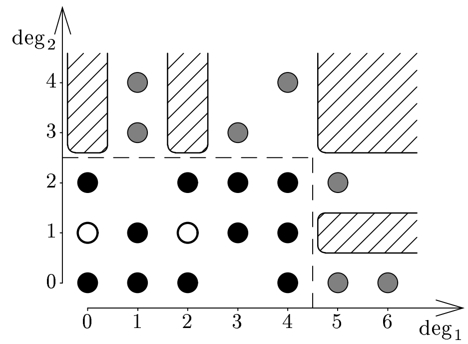

Figure 1: Monomials of a polynomial P with $(4,2)$-leading multiindices $(0,1)$ and $(2,1)$.

Proof. In our first proof we use the tensor product property (4) and the linearity of the map $P \mapsto (\Psi(N^{-1}P|_{\mathfrak{X}}))_\varepsilon$ to reduce the problem to the one-dimensional case. The one-dimensional case is covered by the inversion formula 2.9. Another proof, following Alon and Tarsi's polynomial method, is described in Section 7.

Since both sides of the Equation (i) are linear in the argument P it suffices to prove $(X^\delta)_\varepsilon = (\Psi(N^{-1}X^\delta|_{\mathfrak{X}}))_\varepsilon$ in the two cases of Definition 3.1. In each case,

$$\begin{aligned}\Big(\Psi(N^{-1}X^\delta|_{\mathfrak{X}})\Big)_\varepsilon &= \Big(\Psi\Big(\big(\textstyle\bigotimes_j N_j^{-1}\big)\textstyle\bigotimes_j (X_j^{\delta_j}|_{\mathfrak{X}_j})\Big)\Big)_\varepsilon \\ &= \Big(\big(\textstyle\bigotimes_j \Psi^j\big)\textstyle\bigotimes_j (N_j^{-1}X_j^{\delta_j}|_{\mathfrak{X}_j})\Big)_\varepsilon \\ &\overset{(4)}{=} \Big(\textstyle\bigotimes_j \big(\Psi^j(N_j^{-1}X_j^{\delta_j}|_{\mathfrak{X}_j})\big)\Big)_\varepsilon \\ &= \prod_{j\in(n]} \Big(\Psi^j(N_j^{-1}X_j^{\delta_j}|_{\mathfrak{X}_j})\Big)_{\varepsilon_j} .\end{aligned} \tag{34}$$

Using the one-dimensional case of the inversion formula 2.9 we also derive

$$\Big(\Psi^j(N_j^{-1}X_j^{\delta_j}|_{\mathfrak{X}_j})\Big)_{\varepsilon_j} = (X_j^{\delta_j})_{\varepsilon_j} = ?_{(\delta_j=\varepsilon_j)} \qquad \text{for all } j\in(n] \text{ with } \delta_j \le d_j . \tag{35}$$

Thus in case 1 ($\forall\, j\in(n]:\ \ \delta_j=\varepsilon_j\le d_j$) ,

$$\Big(\Psi(N^{-1}X^\delta|_{\mathfrak{X}})\Big)_\varepsilon = 1 = (X^\delta)_\varepsilon . \tag{36}$$

And in case 2 ($\exists\, j\in(n]:\ \ \varepsilon_j\ne\delta_j\le d_j$) ,

$$\Big(\Psi(N^{-1}X^\delta|_{\mathfrak{X}})\Big)_\varepsilon = 0 = (X^\delta)_\varepsilon . \tag{37}$$

□

Note that the one-dimensional case of Theorem 3.2 (ii) is nothing more then the well-known fact that polynomials $P(X_1)\ne 0$ of degree at most d_1 have at most d_1 roots.

With the remark after Definition 3.1, and the knowledge that $\psi_{d,x} \overset{(6)}{=} 1$ for all $x\in\mathfrak{X}$, we get our main result as an immediate consequence of Theorem 3.2:

Theorem 3.3 (Coefficient formula). *Let $\mathfrak{X}$ be an integral d-grid. For each polynomial $P = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta \in \mathcal{R}[X]$ of total degree* $\deg(P) \le \Sigma d$ *,*

(i) $\boxed{P_d \;=\; \Sigma(N^{-1}P|_{\mathfrak{X}})}$ $\quad$ *(* $= \sum_{x\in\mathfrak{X}} N(x)^{-1}P(x)$ *), and*

(ii) $\boxed{P_d \ne 0 \implies P|_{\mathfrak{X}} \not\equiv 0}$.

This main theorem looks simpler then the more general Theorem 3.2, and you do not have to know the concept of d-leading multiindices to understand it. Furthermore, the applications in this paper do not really make use of the generality in Theorem 3.2 . However, we tried to provide as much generality as possible, and it is of course interesting to understand the role of the degree restriction in Theorem 3.3 .

The most important part of this results, the implication in Theorem 3.3 (*ii*), which is known as Combinatorial Nullstellensatz was already proven in [Al2, Theorem 1.2], for integral domains. Note that $P_d = 0$ whenever $\deg(P) < \Sigma d$, so that the implication seems to become useless in this situation. However, one may modify P, or use smaller sets $\mathfrak{X}_j$ (and hence smaller d_j), and apply the implication then. So, if $P_\delta \ne 0$ for a $\delta \le d$ with $\Sigma\delta = \deg(P)$ then it still follows that $P|_{\mathfrak{X}} \not\equiv 0$. De facto, such δ are d-leading.

If, on the other hand, $\deg(P) = \Sigma d$, then P_d is, in general, the only coefficient that allows conclusions on $P|_{\mathfrak{X}}$ as in Theorem 3.3 (*ii*). This follows from the modification methods of Section 7 . More precisely, if we do not have further information about the d-grid $\mathfrak{X}$, then the d-leading coefficients are the only coefficients that allow such conclusions. For special grids $\mathfrak{X}$, however, there may be some other coefficients P_δ with this property, e.g., P_0 in the case $0 = (0,\ldots,0) \in \mathfrak{X}$.

Note further that for special grids $\mathfrak{X}$, the degree restriction in Theorem 3.3 may be weakened slightly. If, for example, $\mathfrak{X} = \mathbb{F}_q^n$, then the restriction $\deg(P) \le \Sigma d + q - 2$ suffices; see the footnote on page 28 for an explanation.

The following corollary is a consequence of the simple fact that vanishing sums – the case $P_d = 0$ in Theorem 3.3 (*i*) – do not have exactly one nonvanishing summand. It is very useful if a problem possesses exactly one trivial solution: if we are able to describe the problem by a polynomial of low degree, we just have to check the degree, and Corollary 3.4 guarantees a second (in this case, nontrivial) solution. There are many elegant applications of this; for some examples see Section 4 . We will work out a general working frame in Section 6 . We have:

Corollary 3.4. *Let $\mathfrak{X}$ be an integral d-grid. For polynomials P of degree* $\deg(P) < \Sigma d$ *(or, more generally, for polynomials with vanishing d-leading coefficient $P_d = 0$),*

$$\boxed{\big|\{\, x \in \mathfrak{X} \;\vert\; P(x) \ne 0 \,\}\big| \ne 1} \;.$$

If the grid $\mathfrak{X}$ has a special structure – for example, if $\mathfrak{X} \subseteq \mathbb{R}_{>0}{}^n$ – this corollary may also hold for polynomials P with vanishing d-leading coefficient $P_\varepsilon = 0$ for some $\varepsilon \neq d$. The simple idea for the proof of this, which uses Theorem 3.2 instead of Theorem 3.3, leads to the modified conclusion that

$$\left|\{\, x \in \mathfrak{X} \;\middle\vert\; \psi_{\varepsilon,x} P(x) \neq 0 \,\}\right| \neq 1 \,. \tag{38}$$

Note further that the one-dimensional case of Corollary 3.4 is just a reformulation of the well-known fact that polynomials $P(X_1)$ of degree less than d_1 do not have $d_1 = |\mathfrak{X}_1| - 1$ roots, except if $P = 0$.

The example $P = 2X_1 + 2 \in \mathbb{Z}_4[X_1]$, $\mathfrak{X} = \{0, 1, -1\}$ shows that Corollary 3.4 does not hold over arbitrary grids. However, if $\mathfrak{X} = \mathbb{Z}_m{}^n =: \mathcal{R}^n$ with m not prime, the grid $\mathfrak{X}$ is not integral; yet assertion 3.4 holds anyway. Astonishingly, in this case the degree condition can be dropped, too. We will see this in Corollary 8.2.

We also present another proof of Corollary 3.4 that uses only the weaker part (*ii*) of Theorem 3.2, to demonstrate that the well-known Combinatorial Nullstellensatz, our Theorem 3.3 (*ii*), would suffice for the proof of the main part of the corollary:

Proof. Suppose P has exactly one nonzero $x_0 \in \mathfrak{X}$. Then

$$Q \;:=\; P - P(x_0) N^{-1}(x_0) L_{\mathfrak{X},x_0} \;\in\; \mathcal{R}[X] \tag{39}$$

vanishes on the whole grid $\mathfrak{X}$, but possesses the nonvanishing and d-leading coefficient

$$Q_d \;=\; -P(x_0) N^{-1}(x_0) \;\neq\; 0 \;, \tag{40}$$

in contradiction to Theorem 3.2 (*ii*). □

A further useful corollary, and a version of Chevalley and Warning's classical result – Theorem 4.3 in this paper – is the following result (see also [Scha2] for a sharpening of Warning's Theorem, and Theorem 8.4 for a similar result over $\mathbb{Z}_{p^k}$):

Corollary 3.5. *Let $\mathfrak{X} \subseteq \mathbb{F}_{p^k}^n$ be a d-grid and $P_1, \ldots, P_m \in \mathbb{F}_{p^k}[X_1, \ldots, X_n]$. If $(p^k - 1) \sum_{i \in (m]} \deg(P_i) < \Sigma d$, then*

$$\boxed{\left|\left\{\, x \in \mathfrak{X} \;\middle\vert\; P_1(x) = \cdots = P_m(x) = 0 \,\right\}\right| \;\neq\; 1} \,.$$

Proof. Define

$$P \;:=\; \prod_{i \in (m]} (1 - P_i^{\,p^k - 1}) \;; \tag{41}$$

then for points $x = (x_1, \ldots, x_n)$,

$$P(x) \neq 0 \iff \forall\, i \in (m] \colon\; P_i(x) = 0 \;, \tag{42}$$

and hence

$$\left|\left\{ x \in \mathfrak{X} \mid P_1(x) = \cdots = P_m(x) = 0 \right\}\right| \;=\; \left|\left\{ x \in \mathfrak{X} \mid P(x) \neq 0 \right\}\right| \overset{3.4}{\neq} 1 \;, \tag{43}$$

since

$$\deg(P) \;\leq\; \sum_{i \in (m]} (p^k - 1) \deg(P_i) \;<\; \Sigma d \;. \tag{44}$$

□

4 First applications and the application principles

In this section we present some short and elegant examples of how our theorems may be applied. They are all well-known, but we wanted to have some examples to demonstrate the flexibility of these methods. This flexibility will also be emphasized through the general working frame described in Section 6, for which the applications of this section may serve as examples. Alon used them already in [Al2] to demonstrate the usage of implication 3.3 (ii); whereas we prove them by application of Theorem 3.3 (i), and corollaries 3.4 and 3.5, an approach which is – in most cases – more straightforward and more elegant. The main advantage of the coefficient formula 3.3 (i) can be seen in the proof of Theorem 4.3, where the implication 3.3 (ii) does not suffice to give a proof of the full theorem. Section 5 will contain another application that puts the new quantitative aspect of coefficient formula 3.3 into the spotlight.

Our first example was originally proven in [AFK]:

Theorem 4.1. *Every loopless* 4*-regular multigraph plus one edge* $G = (V, E \uplus \{e_0\})$ *contains a nontrivial* 3*-regular subgraph.*

See [AFK2] and [MoZi] for further similar results. The additional edge e_0 in our version is necessary as the example of a triangle with doubled edges shows.

We give a comprehensive proof in order to outline the principles:

Proof. Of course, the empty graph $(\varnothing, \varnothing)$ is a (trivial) 3-regular subgraph. So there is one "solution," and we just have to show that there is not exactly one "solution." This is where Corollary 3.5 comes in. Systems of polynomials of low degree do not have exactly one common zero. Thus, if the 3-regular subgraphs correspond to the common zeros of such a system of polynomials we know that there has to be a second (nontrivial) "solution."

The subgraphs without isolated vertices can be identified with the subsets S of the set of all edges $\bar{E} := E \uplus \{e_0\}$. Now, an edge $e \in \bar{E}$ may or may not lie in a subgraph $S \subseteq \bar{E}$. We represent these two possibilities by the numbers 1 and 0 in $\mathfrak{X}_e := \{0, 1\}$ (the first step in the algebraization), we define

$$\chi(S) \;:=\; \left(?_{(e \in S)}\right)_{e \in \bar{E}} \;\in\; \mathfrak{X} \;:=\; \{0,1\}^{\bar{E}} \;\subseteq\; \mathbb{F}_3^{\bar{E}} \;. \tag{45}$$

With this representation, the subgraphs S correspond to the points $x = (x_e)$ of the Boolean grid $\mathfrak{X} := \{0,1\}^{\bar{E}} \subseteq \mathbb{F}_3^{\bar{E}}$; and it is easy to see that the polynomials

$$P_v := \sum_{e \ni v} X_e \in \mathbb{F}_3[\, X_e \mid e \in \bar{E}\,] \quad \text{for all } v \in V \tag{46}$$

do the job, i.e., they have sufficient low degrees and the common zeros $x \in \mathfrak{X}$ correspond to the 3-regular subgraphs. To see this, we have to check for each vertex $v \in V$ the number $\big|\{\, e \ni v \mid x_e = 1\,\}\big| \leq 5$ of edges e connected to v that are "selected" by a common zero $x \in \mathfrak{X} = \{0,1\}^{\bar{E}}$:

$$P_v(x) = 0 \quad \Longleftrightarrow \quad \sum_{e \ni v} x_e = 0 \quad \Longleftrightarrow \quad \big|\{\, e \ni v \mid x_e = 1\,\}\big| \in \{0,3\} \ . \tag{47}$$

Furthermore, we have to check the degree condition of Corollary 3.5, and that is where we need the additional edge e_0 :

$$(3^1 - 1) \sum_{v \in V} \deg(P_v) \;=\; 2|V| \;=\; |E| \;<\; |\bar{E}| \;=\; \Sigma d(\mathfrak{X}) \ . \tag{48}$$

By Corollary 3.5, the trivial graph $\varnothing \subseteq \bar{E}$ ($x = 0$) cannot be the only 3-regular subgraph. □

The following simple, geometric result was proven by Alon and Füredi in [AlFü], and answers a question by Komjáth. Our proof uses Corollary 3.4:

Theorem 4.2. *Let $H_1, H_2, \ldots, H_m$ be affine hyperplanes in $\mathbb{F}^n$ ($\mathbb{F}$ a field) that cover all vertices of the unit cube $\mathfrak{X} := \{0,1\}^n$ except one, then $m \geq n$.*

Proof. Let $\sum_{j \in (n]} a_{i,j} X_j = b_i$ be an equation defining H_i , and set

$$P := \prod_{i \in (m]} \sum_{j \in (n]} (a_{i,j} X_j - b_i) \;\in\; \mathbb{F}[X_1, \ldots, X_n] \ ; \tag{49}$$

then for points $x = (x_1, \ldots, x_n)$;

$$P(x) \neq 0 \quad \Longleftrightarrow \quad \Big(\forall\, i \in (m]\colon \sum_{j \in (n]} a_{i,j} x_j \neq b_i \Big) \quad \Longleftrightarrow \quad x \notin \bigcup_{j \in (m]} H_j \ . \tag{50}$$

If we now suppose $m < n$, then it follows that

$$\deg(P) \;\leq\; m \;<\; n \;=\; \Sigma d(\mathfrak{X}) \ , \tag{51}$$

and hence,

$$\Big|\mathfrak{X} \setminus \textstyle\bigcup_{j \in (m]} H_j\Big| \;=\; \big|\{\, x \in \mathfrak{X} \mid P(x) \neq 0\,\}\big| \;\overset{3.4}{\neq}\; 1 \ . \tag{52}$$

This means that there is not one unique uncovered point x in $\mathfrak{X} = \{0,1\}^n$ – $m < n$ hyperplanes are not enough to achieve that. □

Our next example is a classical result of Chevalley and Warning that goes back to a conjecture of Dickson and Artin. There are a lot of different sharpenings to it; see [MSCK], [Scha2], Corollary 3.5 and Theorem 8.4 . In the proof of the classical version, presented below, we do not use the Boolean grid $\{0,1\}^n$, as in the last two examples. We also have to use Theorem 3.3 (i) instead of its corollaries. What remains the same as in the proof of the closely related Corollary 3.5 is that we have to translate a system of equations into a single inequality:

Theorem 4.3. *Let p be a prime and $P_1, P_2, \ldots, P_m \in \mathbb{F}_{p^k}[X_1, \ldots, X_n]$.*
If $\sum_{i\in(m]} \deg(P_i) < n$, then

$$p \,\Big\lfloor\, \left|\{\, x \in \mathbb{F}_{p^k}^n \;\mid\; P_1(x) = \cdots = P_m(x) = 0 \,\}\right| \ ,$$

and hence the P_i do not have one unique common zero x .

Proof. Define

$$P \;:=\; \prod_{i\in(m]} (1 - P_i^{p^k-1}) \ ; \tag{53}$$

then

$$P(x) \;=\; \begin{cases} 1 & \text{if } P_1(x) = \cdots = P_m(x) = 0\,, \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } \ x \in F_{p^k}^{\;n}\,, \tag{54}$$

thus, with $\mathfrak{X} := \mathbb{F}_{p^k}^n$,

$$\left|\left\{x \in \mathbb{F}_{p^k}^n \;\mid\; P_1(x) = \cdots = P_m(x) = 0\right\}\right| \cdot 1 \;=\; \sum_{x\in\mathfrak{X}} P(x) \;\overset{\substack{3.3\\1.4}}{=}\; (-1)^n\,(P)_{d(\mathfrak{X})} \;\overset{(56)}{=}\; 0 \ , \tag{55}$$

where the last two equalities hold as

$$\deg(P) \;\le\; (p^k-1)\sum_{i\in(m]} \deg(P_i) \;<\; (p^k-1)\,n \;=\; \Sigma d(\mathfrak{X}) \ . \tag{56}$$

□

The Cauchy-Davenport Theorem is another classical result. It was first proven by Cauchy in 1813, and has many applications in additive number theory. The proof of this result is as simple as the last ones, but here we use the coefficient formula 3.3 (i) in the other direction – we know the polynomial map $P|_{\mathfrak{X}}$, and use it to determine the coefficient P_d :

Theorem 4.4. *If p is a prime, and A and B are two nonempty subsets of $\mathbb{Z}_p := \mathbb{Z}/p\mathbb{Z}$, then*

$$|A+B| \ge \min\{\, p\,,\, |A| + |B| - 1 \,\} \ .$$

Proof. We assume $|A+B| \leq |A|+|B|-2$, and must prove $|A+B| \geq p$.

Define

$$P \;:=\; \prod_{c\in A+B} (X_1+X_2-c) \;\in\; \mathbb{Z}_p[X_1,X_2] \;\;, \tag{57}$$

set

$$\mathfrak{X}_1 := A \;\;, \tag{58}$$

and choose a subset

$$\varnothing \neq \mathfrak{X}_2 \subseteq B \tag{59}$$

of size

$$|\mathfrak{X}_2| \;=\; |A+B|-|A|+2 \quad (\;\leq\; |B|\;) \;\;. \tag{60}$$

Now

$$P|_{\mathfrak{X}_1\times\mathfrak{X}_2} \;\equiv\; 0 \;\;, \tag{61}$$

and

$$\deg(P) \;=\; |A+B| \;=\; |\mathfrak{X}_1|+|\mathfrak{X}_2|-2 \;=\; d_1(\mathfrak{X}_1)+d_2(\mathfrak{X}_2) \;\;, \tag{62}$$

so that

$$\tbinom{|A+B|}{d_1}\cdot 1 \;=\; P_{(d_1,|A+B|-d_1)} \;=\; P_d \overset{3.3}{=} \sum_{x\in\mathfrak{X}_1\times\mathfrak{X}_2} 0 \;=\; 0 \;\in\; \mathbb{Z}_p \;\;. \tag{63}$$

Hence

$$p \;\lfloor\; \tbinom{|A+B|}{d_1} \;\;, \tag{64}$$

and it follows that

$$|A+B| \;\geq\; p \;\;. \tag{65}$$

□

There are some further number-theoretic applications, for example, Erdős, Ginzburg and Ziv's Theorem, which also can be found in [Al2].

5 The matrix polynomial – another application

In this section we apply our results to the matrix polynomial $\Pi(AX)$, a generalization of the graph polynomial (see also [AlTa2] or [Ya]).

We always assume $A=(a_{i,j}) \in \mathcal{R}^{m\times n}$, and the product of this matrix with the tuple $X := (X_1,\ldots,X_n) \in \mathcal{R}[X]^n$ is $AX := (\sum_{j\in(n]} a_{ij}X_j)_{i\in(m]}$. Now, $\Pi(AX)$ is defined in accordance with the definition of Π in Section 1, as follows: A, X AX

Definition 5.1 (Matrix polynomial). The *matrix polynomial* of $A=(a_{i,j}) \in \mathcal{R}^{m\times n}$ is given by $\Pi(AX)$

$$\Pi(AX) \;:=\; \prod_{i\in(m]} \sum_{j\in(n]} a_{ij}X_j \;\in\; \mathcal{R}[X] \;\;.$$

It turns out that the coefficients of the matrix polynomial are some kind of permanents:

Definition 5.2 (δ-permanent). For $\delta \in \mathbb{N}^n$ we define the *δ-permanent* of $A = (a_{i,j}) \in \mathcal{R}^{m\times n}$ through $\mathrm{per}_\delta(A)$

$$\mathrm{per}_\delta(A) \;:=\; \sum_{\substack{\sigma:\,(m]\to(n]\\ |\sigma^{-1}|=\delta}} \pi_A(\sigma) \;\;,$$

where $\pi_A(\sigma)$

$$\pi_A(\sigma) \;:=\; \prod_{i\in(m]} a_{i,\sigma(i)} \quad \text{and} \quad |\sigma^{-1}| \;:=\; \big(|\sigma^{-1}(j)|\big)_{j\in(n]} \;.$$ $|\sigma^{-1}|$

Obviously, $\mathrm{per}_\delta(A) = 0$ if $\Sigma\delta \neq m$. If $m = n$ then $\mathrm{per} := \mathrm{per}_{(1,1,\ldots,1)}$ is the usual permanent; and, if $\Sigma\delta = m$, it is easy to see that per

$$\big(\textstyle\prod_{j\in(n]} \delta_j!\big)\; \mathrm{per}_\delta(A) \;=\; \mathrm{per}(A\langle|\delta\rangle), \tag{66}$$

where $A\langle|\delta\rangle$ is a matrix that contains the j^{th} column of A exactly δ_j times. But note that $\mathrm{per}_\delta(A)$ is, in general, not determined by $\mathrm{per}(A\langle|\delta\rangle)$. If, for example, $(\prod_{j\in(n]} \delta_j!)\,1 = 0$ in $\mathcal{R}$, the δ-permanent $\mathrm{per}_\delta(A)$ may take arbitrary values, while $\mathrm{per}(A\langle|\delta\rangle) = 0$. $A\langle|\delta\rangle$

As an immediate consequence of the definitions, we have

Lemma 5.3.

$$\boxed{\Pi(AX) \;=\; \textstyle\sum_{\delta\in\mathbb{N}^n} \mathrm{per}_\delta(A)\, X^\delta}\;.$$

The next theorem now easily follows from our main result, Theorem 3.3. It is an integrative generalization of Alon's Permanent Lemma [Al2, Section 8], and of Ryser's permanent formula [BrRy, p.200], which follow as the special cases:

- $m = n$, $d = (1,1,\ldots,1)$ *of the following 5.4 (ii) over fields,*
- $m = n$, $d = (1,1,\ldots,1)$, $\mathfrak{X} = \{0,1\}^n$, $b = (0,0,\ldots,0)$ *of 5.4 (i) over fields.*

We already proved a slightly weaker version for $\mathfrak{X} \subseteq \mathbb{N}^n \subseteq \mathcal{R}^n$ in [Scha, 1.14 & 1.15]. This proof was based on Ryser's formula, and is a little more technical. [Scha, 1.10] is the special case $\mathfrak{X} = [d] \subseteq \mathbb{N}^n \subseteq \mathcal{R}^n$, but you will have to use 1.4 (v) to see this. For some additional tricks over fields of characteristic $p > 0$, see [DeV]. We have:

Theorem 5.4 (Permanent formula). *Suppose $A = (a_{ij}) \in \mathcal{R}^{m\times n}$ and $b = (b_i) \in \mathcal{R}^m$ are given, and let $\mathfrak{X} \subseteq \mathcal{R}^n$ be an integral d-grid. If $m \leq \Sigma d$, then* A

(i) $\boxed{\mathrm{per}_d(A) \;=\; \textstyle\sum_{x\in\mathfrak{X}} N(x)^{-1}\, \Pi(Ax - b)}$, *and*

(ii) $\boxed{\mathrm{per}_d(A) \neq 0 \quad\Longrightarrow\quad \exists\, x \in \mathfrak{X}\colon\; (Ax)_1 \neq b_1\,,\, \ldots,\, (Ax)_m \neq b_m}$.

Proof. Part (i) follows from Theorem 3.3, as $\deg\big(\Pi(AX-b)\big) = m \le \Sigma d$, and since $\big(\Pi(AX-b)\big)_d = \big(\Pi(AX)\big)_d \overset{5.3}{=} \mathrm{per}_d(A)$. Part (ii) is a simple consequence of part (i). $\square$

We call an element $x \in \mathcal{R}^m$ with $(Ax)_1 \neq 0$, ..., $(Ax)_m \neq 0$ a (correct) *coloring* of A, and a map $\sigma\colon (m] \longrightarrow (n]$ with $\pi_A(\sigma) \neq 0$ and $|\sigma^{-1}| = \delta$ is a δ-*orientation* of A. With this terminology, Theorem 5.4 describes a connection between the orientations and the colorings of A, and it is not too difficult to see that this is a sharpening and a generalization of Alon and Tarsi's Theorem about colorings and orientations of graphs in [AlTa]. That is because, in virtue of the embedding $\vec{G} \longmapsto A(\vec{G})$ described in (67) below, oriented graphs form a subset of the set of matrices, if $-1 \neq 1$ in $\mathcal{R}$. The resulting sharpening 5.5 of the Alon-Tarsi Theorem contains Scheim's formula for the number of edge r-colorings of a planar r-regular graph as a permanent and Ellingham and Goddyn's partial solution of the list coloring conjecture. We briefly elaborate on this; for even more detail, see [Scha], where we described this for grids $\mathfrak{X} \subseteq \mathbb{N}^n \subseteq \mathcal{R}^n$, and where we pointed out that many other graph-theoretic theorems may be formulated for matrices, too. σ

Let $\vec{G} = (V, E, \rightarrow, \leftarrow)$ be a *oriented multigraph* with *vertex set* V, *edge set* E and *defining orientations* $\rightarrow\colon E \longrightarrow V$, $e \longmapsto e^{\rightarrow}$ and $\leftarrow\colon e \longmapsto e^{\leftarrow}$. $\vec{G}$ shall be *loopless*, so that $e^{\rightarrow} \neq e^{\leftarrow}$ for all $e \in E$. We write $v \in e$ instead of $v \in \{e^{\rightarrow}, e^{\leftarrow}\}$ and define the *incidence matrix* $A(\vec{G})$ of $\vec{G}$ by $\vec{G}, e^{\rightarrow}$ $\rightarrow, \leftarrow$ $v \in e$ $A(\vec{G})$

$$A(\vec{G}) := (a_{e,v}) \in \mathcal{R}^{E\times V}, \text{ where } a_{e,v} := ?_{(e^{\rightarrow}=v)} - ?_{(e^{\leftarrow}=v)} \in \{-1,0,1\} \,. \tag{67}$$

With this definition, the *orientations* $\sigma\colon E \ni e \longmapsto e^{\sigma} \in e$ and the *colorings* $x\colon V \longrightarrow \mathcal{R}$ of $\vec{G}$ are exactly the orientations and the colorings of $A(\vec{G})$ as defined above. The orientations σ of $A(\vec{G})$ have the special property $\pi_{A(\vec{G})}(\sigma) = \pm 1$. According to this, we say that an orientation σ of $\vec{G}$ is *even/odd* if $e^{\sigma} \neq e^{\rightarrow}$ (i.e., $e^{\sigma} = e^{\leftarrow}$) holds for even/odd many edges $e \in E$. We write DE_δ / DO_δ for the set of even/odd orientations σ of $\vec{G}$ with $|\sigma^{-1}| = \delta \in \mathbb{N}^V$. With this notation we have: DE_δ DO_δ

Corollary 5.5. *Let $\vec{G} = (V, E, \rightarrow, \leftarrow)$ be a loopless, directed multigraph and $\mathfrak{X} \subseteq \mathcal{R}^V$ be an integral d-grid; where $d = (d_v) \in \mathbb{N}^V$, and $d_v = |\mathfrak{X}_v| - 1$ for all $v \in V$.*
If $|E| \le \Sigma d$, then

(i) $|DE_d| - |DO_d| = \mathrm{per}_d(A(\vec{G})) = \sum_{x\in\mathfrak{X}} N(x)^{-1} \prod_{e\in E}\big(x_{e^{\rightarrow}} - x_{e^{\leftarrow}}\big)$,

(ii) $|DE_d| \neq |DO_d| \implies \exists\, x \in \mathfrak{X}\colon \forall\, e \in E\colon\ x_{e^{\rightarrow}} \neq x_{e^{\leftarrow}}$ *(x is a coloring).*

Furthermore, it is not so hard to see that, if EE / EO is the set of even/odd Eulerian subgraphs of $\vec{G}$, and $\delta := |{\rightarrow}^{-1}|$, we have $|DE_\delta| = |EE|$ and $|DO_\delta| = |EO|$; see also [Scha, 2.6]. EE, EO

Note that even though Corollary 5.5 looks a little simpler than [Scha, 1.14 & 2.4], there is some complexity hidden in the symbol $N(x)$. If the "lists" $\mathfrak{X}_v$ ($\mathfrak{X} = \prod_{v\in V} \mathfrak{X}_v$) are all equal, this becomes less complex. Further, if the graph $\vec{G}$ is the line graph of a r-regular graph, so that its vertex colorings are the edge colorings of the r-regular graph, then the whole right side becomes very simple. The summands are then – up to a constant factor – equal to ± 1; or to 0, if $x = (x_v)_{v \in V}$ is not a correct coloring. The corresponding specialization of equation 5.5 (i) was already obtained in [ElGo] and [Sch].

If in addition $\vec{G}$ is planar this formula becomes even simpler, so that the whole right side is – up to a constant factor – the number of edge r-colorings of the r-regular graph. Scheim [Sch] proved this specialization in his approach to the four color problem for 3-regular graphs using a result of Vigneron [Vig]. However, with Ellingham and Goddyn's generalization [ElGo, Theorem 3.1] of Vigneron's result, this specialization also follows in the r-regular case.

As the left side of our equation does not depend on the choice of the d-grid $\mathfrak{X}$, the right side does not depend on it, either. In our special case, where the right side is the number of r-colorings of the line graph of a planar r-regular graph, this means that if there are colorings to equal lists $\mathfrak{X}_v$ of size r (e.g., $\mathfrak{X} = [r]^V$), then there are also colorings to arbitrary lists $\mathfrak{X}_v$ of size $|\mathfrak{X}_v| = r$ – which is just Ellingham and Goddyn's confirmation of the list coloring conjecture for planar r-regular edge r-colorable multigraphs [ElGo].

6 Algebraically solvable existence problems: Describing polynomials as equivalent to explicit solutions

In this section we describe a general working frame to Theorem 3.3 (ii) and Corollary 3.4, as it may be used in existence proofs, such as those of 3.5, 4.2 or 5.4 (ii). We call the polynomials defined in the equations (41) and (49) or the matrix polynomial $\Pi(AX)$ in our last example, algebraic solutions, and show that such algebraic solutions may be seen as equivalent to explicit solutions. We show that the existence of algebraic solutions, and of nontrivial explicit solutions are equivalent. To make this more exact, we have to introduce some definitions. Our definition of problems should not merely reflect common usage. In fact, the generality gained through an exaggerated extension of the term "problem" through abstraction is desirable.

Definition 6.1 (Problem). A *problem* $\mathcal{P}$ is a pair $(\mathcal{S}, \mathcal{S}_{\text{triv}})$ consisting of a set $\mathcal{S}$, which we call its set of *solutions*; and a subset $\mathcal{S}_{\text{triv}} \subseteq \mathcal{S}$, which we call its set of *trivial solutions.* $\mathcal{P}$ $(\mathcal{S}, \mathcal{S}_{\text{triv}})$

In example 4.1, the set of solutions $\mathcal{S}$ consists of the 3-regular subgraphs and $\mathcal{S}_{\text{triv}} = \{(\varnothing, \varnothing)\}$. These are exact definitions, but it does not mean that we know if there are nontrivial solutions, i.e., if $\mathcal{S} \neq \mathcal{S}_{\text{triv}}$. The set $\mathcal{S}$ is well defined, but we do not know what it looks like; indeed, that is the actual problem.

To apply our theory about polynomials in such general situations, we have to bring in grids $\mathfrak{X}$ in some way. For that, we define impressions:

Definition 6.2 (Impression). A triple $(\mathcal{R},\mathfrak{X},\chi)$ is an *impression* of $\mathcal{P}$ if $\mathcal{R}$ is a commutative ring with $1\neq 0$, if $\mathfrak{X}=\mathfrak{X}_1\times\cdots\times\mathfrak{X}_n\subseteq\mathcal{R}^n$ is a finite integral grid (for some $n\in\mathbb{N}$) and if $\chi\colon\mathcal{S}\longrightarrow\mathfrak{X}$ is a map. $(\mathcal{R},\mathfrak{X},\chi)$

As the set $\mathcal{S}$ of solutions is usually unknown, one may ask how the map $\chi\colon\mathcal{S}\longrightarrow\mathfrak{X}$ can be defined. The answer is that we usually, define χ on a bigger domain at first, as in Equation (45) in example 4.1. Then the unknown set of solutions $\mathcal{S}$ (more precisely, its image $\chi(\mathcal{S})$) is indirectly described:

Definition 6.3 (Describing polynomial). A polynomial $P\in\mathcal{R}[X_1,\dots,X_n]$ is a *describing polynomial* of $\mathcal{P}$ over $(\mathcal{R},\mathfrak{X},\chi)$ if

$$\chi(\mathcal{S}) \;=\; \operatorname{supp}(P|_{\mathfrak{X}}) \;\;.$$

The diagram (2) in the introduction shows a schematic illustration of our concept in the case $\mathcal{S}_{\text{triv}}=\varnothing$. The next question is how it might be possible to reveal the existence of nontrivial solutions using some knowledge about a describing polynomial P, and how to find such an appropriate P. In view of our results from Section 3, we give the following definition:

Definition 6.4 (Algebraic solutions). A describing polynomial P is an *algebraic solution* (over $(\mathcal{R},\mathfrak{X},\chi)$) of a problem of the form $\mathcal{P}=(\mathcal{S},\varnothing)$ if it fulfills

$$deg(P)\;\le\;\Sigma d(\mathfrak{X}) \qquad\text{and}\qquad P_{d(\mathfrak{X})}\;\neq\;0\;\;.$$

It is an *algebraic solution* of a problem $\mathcal{P}=(\mathcal{S},\mathcal{S}_{\text{triv}})$ with $\mathcal{S}_{\text{triv}}\neq\varnothing$ if it fulfills

$$\deg(P)\;<\;\Sigma d(\mathfrak{X}) \qquad\text{and}\qquad \sum_{x\in\chi(\mathcal{S}_{\text{triv}})} N(x)^{-1}P(x)\;\neq\;0 \quad (\text{e.g., if } |\chi(\mathcal{S}_{\text{triv}})|=1\,).$$

The bad news is that now, we do not have a general recipe for finding algebraic solutions that indicate the solvability of problems. However, we have seen that there are several combinatorial problems that are algebraically solvable in an obvious way. The construction of algebraic solutions in these examples follows more or less the same simple pattern, and that constructive approach is the big advantage. Algebraic solutions are easy to construct if the problem is not too complex in the sense that the construction does not require too many multiplications. In many cases algebraic solutions can be formulated for whole classes of problems, e.g., for all extended 4-regular graphs in example 4.1, where the final algebraic solution was hidden in Corollary 3.5. In these cases a maybe infinite number of algebraic solutions fit into one single general form, which can be presented on a finite blackboard or sheet of paper. The concept of algebraic solutions provides a method of resolution to what we call the “finite blackboard problem”, a fundamental problem whenever we view general situations with infinitely many concrete instances.

If an algebraic solution is found, we can apply Theorem 3.3, Corollary 3.4 or the following theorem, which also shows that algebraic solutions always exist, provided there are nontrivial solutions in the first place.

Theorem 6.5. *Let $\mathcal{P} = (\mathcal{S}, \mathcal{S}_{\text{triv}})$ be a problem. The following properties are equivalent:*

(i) There exists a nontrivial solution of $\mathcal{P}$; i.e., $\mathcal{S} \neq \mathcal{S}_{\text{triv}}$.

(ii) There exists an algebraic solution of $\mathcal{P}$ over an impression $(\mathcal{R}, \mathfrak{X}, \chi)$.

(iii) There exist algebraic solutions of $\mathcal{P}$ over each impression $(\mathcal{R}, \mathfrak{X}, \chi)$ that fulfills either

- $|\mathcal{R}| > 2$ *and* $\mathcal{S} \neq \mathcal{S}_{\text{triv}} \Rightarrow \chi(\mathcal{S}) \neq \chi(\mathcal{S}_{\text{triv}})$ *(e.g., if χ is injective or if $\mathcal{S}_{\text{triv}} = \varnothing$);* or
- $|\mathcal{R}| = 2$ *and* $|\chi(\mathcal{S})| + 1 \equiv |\chi(\mathcal{S}_{\text{triv}})| \equiv ?_{(\mathcal{S}_{\text{triv}} \neq \varnothing)} \pmod 2$.

Proof. First, assume (ii), and let P be an algebraic solution. We want to show that (i) holds. For $\mathcal{S}_{\text{triv}} = \varnothing$, this follows from Theorem 3.3 (ii). For $\mathcal{S}_{\text{triv}} \neq \varnothing$, we have

$$0 \;=\; P_d \overset{3.3}{=} \Sigma(N^{-1}P|_{\mathfrak{X}}) \;=\; \sum_{x \in \chi(\mathcal{S}_{\text{triv}})} N(x)^{-1}P(x) \;+\; \sum_{x \in \chi(\mathcal{S}) \setminus \chi(\mathcal{S}_{\text{triv}})} N(x)^{-1}P(x) \;, \tag{68}$$

where the first sum over $\chi(\mathcal{S}_{\text{triv}})$ does not vanish. Hence, the second sum over the set $\chi(\mathcal{S}) \setminus \chi(\mathcal{S}_{\text{triv}})$ does not vanish, either. Thus $\chi(\mathcal{S}) \setminus \chi(\mathcal{S}_{\text{triv}}) \neq \varnothing$, and $\mathcal{S} \neq \mathcal{S}_{\text{triv}}$ follows.

To prove $(i) \Longrightarrow (iii)$ for $\mathcal{S}_{\text{triv}} = \varnothing$, assume (i), and define a map $y \colon \mathfrak{X} \longrightarrow \mathcal{R}$ such that $\operatorname{supp}(y) = \chi(\mathcal{S})$ and $\Sigma y \neq 0$. (In the case $|\mathcal{R}| = 2$, we need $|\chi(\mathcal{S})| \equiv 1 \pmod 2$ to make this possible.) The interpolation polynomial $P := (\Psi y)(X)$ to the map Ny described in Theorem 2.3 now has degree $deg(P) \leq \Sigma d$, and fulfills

$$\operatorname{supp}(P|_{\mathfrak{X}}) \overset{2.3}{=} \operatorname{supp}(y) \;=\; \chi(\mathcal{S}) \tag{69}$$

and

$$P_d \overset{3.3}{=} \Sigma(N^{-1}P|_{\mathfrak{X}}) \overset{2.3}{=} \Sigma y \;\neq\; 0 \;. \tag{70}$$

To prove $(i) \Longrightarrow (iii)$ for $\mathcal{S}_{\text{triv}} \neq \varnothing$, assume (i), and define a map $y \colon \mathfrak{X} \longrightarrow \mathcal{R}$ such that $\operatorname{supp}(y) = \chi(\mathcal{S})$, $\sum_{x \in \chi(\mathcal{S}_{\text{triv}})} y(x) \neq 0$ and $\Sigma y = 0$. (In the case $|\mathcal{R}| = 2$, we need $|\chi(\mathcal{S})| + 1 \equiv |\chi(\mathcal{S}_{\text{triv}})| \equiv 1 \pmod 2$ to make this possible.) Now, the polynomial $P := (\Psi y)(X)$ has partial degrees $\deg_j(P) \leq d_j$, and total degree $deg(P) < \Sigma d$, as

$$P_d \overset{3.3}{=} \Sigma(N^{-1}P|_{\mathfrak{X}}) \overset{2.3}{=} \Sigma y \;=\; 0 \;. \tag{71}$$

It satisfies

$$\operatorname{supp}(P|_{\mathfrak{X}}) \overset{2.3}{=} \operatorname{supp}(y) \;=\; \chi(\mathcal{S}) \tag{72}$$

and

$$\sum_{x \in \chi(\mathcal{S}_{\text{triv}})} N(x)^{-1}P(x) \overset{2.3}{=} \sum_{x \in \chi(\mathcal{S}_{\text{triv}})} y \;\neq\; 0 \;. \tag{73}$$

Finally, to show $(iii) \Longrightarrow (ii)$, we only have to prove that there exists an impression $(\mathcal{R}, \mathfrak{X}, \chi)$ as described in (iii). This is clear, as we may define χ by setting

$$\chi(s) := \begin{cases} x_{\text{triv}} & \text{for } s \in \mathcal{S}_{\text{triv}}\,, \\ x_{\text{good}} & \text{for } s \in \mathcal{S} \setminus \mathcal{S}_{\text{triv}}\,, \end{cases} \tag{74}$$

where x_{triv} and x_{good} are two distinct, arbitrary elements in some suitable grid $\mathfrak{X}$. □

The arguments in this proof also show that the restrictions to the impression $(\mathcal{R}, \mathfrak{X}, \chi)$ in part (iii) are really necessary. If, for example, we had $|\mathcal{R}| = 2$, $\mathcal{S}_{\text{triv}} = \varnothing$ and $|\chi(\mathcal{S})| \equiv 0 \pmod 2$, then $P_d = 0$, and the problem would not be algebraically solvable with respect to the impression $(\mathcal{R}, \mathfrak{X}, \chi)$.

7 The combinatorial nullstellensatz or how to modify polynomials

In this section, we describe a sharpening of a specialization of Hilbert's Nullstellensatz (see e.g. [DuFo]), the so-called (first) Combinatorial Nullstellensatz. This theorem, and the modification methods behind it, can be used in another proof of the coefficient formulas in Section 3.

We start with an example that illustrates the underlying modification method of this section. It also shows that the coefficient P_d in Theorem 3.3 is, in general, the only coefficient that is uniquely determined by $P|_{\mathfrak{X}}$:

Example 7.1. Let $P \in \mathbb{C}[X_1, X_2]$ (i.e., $\mathcal{R} := \mathbb{C}$ and $n := 2$), and define for $j = 1, 2$:

$$L_j \;:=\; \frac{X_j^5 - 1}{X_j - 1} \;=\; X_j^4 + X_j^3 + X_j^2 + X_j^1 + X_j^0 \qquad \textit{and} \tag{75}$$

$$\mathfrak{X}_j \;:=\; \{x \in \mathbb{C} \mid L_j(x) = 0\} \;=\; \{x_1, x_2, x_3, x_4\}\,, \quad \text{where } x_k := e^{\frac{k}{5}2\pi\sqrt{-1}}\,. \tag{76}$$

Then $d = d(\mathfrak{X}) = (3,3)$. Now, for $\varepsilon \in \mathbb{N}^2$, the polynomial $X^\varepsilon L_1$ (and $X^\varepsilon L_2$) vanishes on $\mathfrak{X}$. Therefore, the modified polynomial

$$P' \;:=\; P + c\,X^\varepsilon L_1\;, \qquad \text{where} \quad c \in \mathcal{R}\backslash 0\;, \tag{77}$$

fulfills

$$P'|_{\mathfrak{X}} \;=\; P|_{\mathfrak{X}}\;; \tag{78}$$

but the coefficients $P_{\varepsilon+(0,0)}$, $P_{\varepsilon+(1,0)}$, $P_{\varepsilon+(2,0)}$, $P_{\varepsilon+(3,0)}$ and $P_{\varepsilon+(4,0)}$ have changed:

$$P'_{\varepsilon+(i,0)} \;=\; P_{\varepsilon+(i,0)} + c \;\neq\; P_{\varepsilon+(i,0)} \quad \text{for} \quad i = 0,1,2,3,4\;. \tag{79}$$

In this way we may modify P without changing the map $P|_{\mathfrak{X}}$.

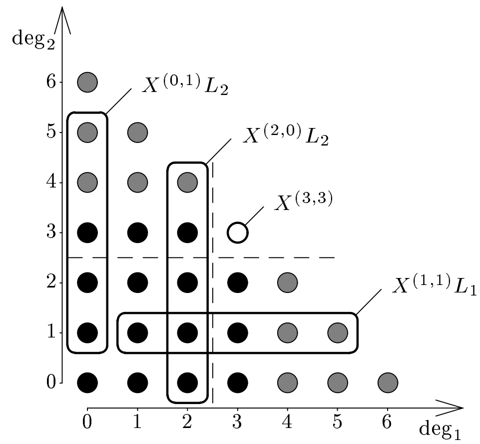

Figure 2: Monomials of degree $\leq 3+3$.

Now, suppose $\deg(P) \leq \Sigma d = 3+3$. Figure 2 illustrates that all coefficients P_δ with $\delta \leq \Sigma d$ – except P_d – can be modified without losing the condition $\deg P \leq \Sigma d$, so that they are not uniquely determined by $P|_{\mathfrak{X}}$. If we try to modify $P_d = P_{(3,3)}$ – for example, by adding $c\,X^{(0,3)}L_1$ (or $c\,X^{(3,0)}L_2$) – we realize that

$$\deg(X^{(0,3)}L_1) \;=\; \deg(X^{(0,7)}) \;=\; 7 \;>\; 3+3 \;=\; \Sigma d \;\; , \tag{80}$$

and $\deg(P') > \Sigma d$ would follow. The coefficient P_d cannot be modified in this way.

This example can also be used to illustrate a second proof of Theorem 3.3 (and Theorem 3.2):

By successive modifications, as above, with L_j

$$L_j = L_{\mathfrak{X}_j}(X_j) := \prod_{\hat{x} \in \mathfrak{X}_j} (X_j - \hat{x}) \tag{81}$$

in the general case, it is possible to modify P into a *trimmed* polynomial $P/\mathfrak{X}$ with the properties $P/\mathfrak{X}$

$$P/\mathfrak{X}|_{\mathfrak{X}} \;=\; P|_{\mathfrak{X}} \qquad \text{and} \qquad \deg_j(P/\mathfrak{X}) \leq d_j \quad \text{for } \; j = 1, \ldots, n\,. \tag{82}$$

By 2.8(v), $P/\mathfrak{X}$ is uniquely determined if $\mathfrak{X}$ is an integral d-grid (e.g., $P/\{x\} = P(x)$). If $\deg(P) \leq \Sigma d$, then it is obviously possible to leave the coefficient P_d unchanged during the modification[1]. Therefore we get

[1] At this point the degree restriction $\deg(P) \leq \Sigma d$ may be weakened slightly if the grid $\mathfrak{X}$ – and hence the L_j – have a special structure. If, e.g., $L_j = X_j^{k+1} - 1$ (or $L_j = X_j^{k+1} - X_j$) for all $j \in (n]$, then $\deg(P) \leq \Sigma d + k$ $[= (n+1)k]$ (respectively $\deg(P) \leq \Sigma d + k - 1$) suffices.

$$P_d \;=\; (P/\mathfrak{X})_d \overset{2.9}{=} (\Psi(N^{-1}P/\mathfrak{X}|_{\mathfrak{X}}))_d \;=\; (\Psi(N^{-1}P|_{\mathfrak{X}}))_d \overset{(6)}{=} \Sigma(N^{-1}P|_{\mathfrak{X}}) \;; \tag{83}$$

and Theorem 3.3 follows immediately.

Theorem 3.2 can also be proven the same way by using the following obvious generalization (Lemma 7.2) of the first equation in (83). Furthermore, we want to mention at this point that the proof above (and the following lemma) may work for some other coefficients P_δ as well if the $L_j = L_{\mathfrak{X}_j}(X_j)$ have a special structure, e.g., $L_j = X^{d_j} - 1$. Of course it works for P_0 if $0 = (0,\ldots,0) \in \mathfrak{X}$, since all L_j lack a constant term in this case. Without further information about the grid $\mathfrak{X}$, we "carry through" only the d-leading coefficients:

Lemma 7.2. *Let $\mathfrak{X}$ be a d-grid. For each polynomial $P = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta \in \mathcal{R}[X]$ with d-leading multiindex $\varepsilon \le d \in \mathbb{N}^n$ (e.g., if $\Sigma\varepsilon = \deg(P)$),*

$$\boxed{(P/\mathfrak{X})_\varepsilon = P_\varepsilon} \;.$$

If we take a closer look at the modification methods above, we see that the difference $P - P/\mathfrak{X}$ can be written as

$$P - P/\mathfrak{X} \;=\; \sum_{j\in(n]} H_j L_j \;\;, \tag{84}$$

with some $H_j \in \mathcal{R}[X]$ of degree $\deg(H_j) \le \deg(P) - \deg(L_j)$.

If $P|_{\mathfrak{X}} \equiv 0$, then $P/\mathfrak{X} = 0$ by the uniqueness of the trimmed polynomial, and (84) yields $P = \sum_{j\in(n]} H_j L_j$. This was proven for integral rings in [Al2, Theorem 1.1]. More formally, we have:

Theorem 7.3 (Combinatorial Nullstellensatz). *Let $\mathfrak{X} = \mathfrak{X}_1 \times \cdots \times \mathfrak{X}_n \subseteq \mathcal{R}^n$ be an integral grid with associated polynomials $L_j := \prod_{\hat{x}\in\mathfrak{X}_j}(X_j - \hat{x})$.*
For any polynomial $P = \sum_{\delta\in\mathbb{N}^n} P_\delta X^\delta \in \mathcal{R}[X]$, the following are equivalent:

(i) $P|_{\mathfrak{X}} \;\equiv\; 0$.

(ii) $P/\mathfrak{X} \;=\; 0$.

(iii) $P \;=\; \sum_{j\in(n]} H_j L_j \;\in\; \mathcal{R}[X]$ *for some polynomials H_j over a ring extension of $\mathcal{R}$.*

(iv) $P \;=\; \sum_{j\in(n]} H_j L_j$ *for some $H_j \in \mathcal{R}[X]$ of degree* $\deg(H_j) \le \deg(P) - |\mathfrak{X}_j|$.

Proof. We already have seen that the implications $(i) \Longrightarrow (ii) \Longrightarrow (iv)$ hold; and the implications $(iv) \Longrightarrow (iii) \Longrightarrow (i)$ are trivial. ∎

The implication "$(i) \implies (iv)$" states that polynomials P with $P|_{\mathfrak{X}} \equiv 0$ may be written as $\sum_{j\in(n]} H_j L_j$. In other words, P lies in the ideal spanned by the polynomials L_j. As we do not know *a priori* that this ideal is a radical ideal, Hilbert's Nullstellensatz would only provide $P^k = \sum_{j\in(n]} H_j L_j$ for some $k \geq 1$, and without degree restrictions for the H_j (provided $\mathcal{R}$ is an algebraically closed field). Alon suggested calling the stronger (with respect to the special polynomials L_j) result "Combinatorial Nullstellensatz." He used it to prove the implication (ii) in the coefficient formula 3.3 [Al2, Theorem 1.2] and recycled the phrase "Combinatorial Nullstellensatz" for the implication 3.3 (ii).

8 Results over $\mathbb{Z}$, $\mathbb{Z}_m$ and other generalizations

There are several ways to generalize the coefficient formulas 3.3 and 3.2. This section will address some of those.

If a grid $\mathfrak{X}$ is just affine but we want to use Theorem 3.3, we may apply the homomorphism $\pi\colon r \mapsto \frac{r}{1}$ from $\mathcal{R}$ to the localization $\mathcal{R}_N$, exactly as in Theorem 2.7. In particular, this leads to the implications: π

$$P_d = 1 \implies P_d^\pi \neq 0 \implies P|_{\mathfrak{X}} \not\equiv 0 \ . \tag{85}$$

It may also be that there is an integral grid $\hat{\mathfrak{X}}$ over a ring $\hat{\mathcal{R}}$, and a homomorphism $\hat{\mathcal{R}} \longrightarrow \mathcal{R}$ that induces a map from $\hat{\mathfrak{X}}$ into $\mathfrak{X}$. Our results may then be applied to a preimage $\hat{P} \in \hat{\mathcal{R}}[X]$ of $P \in \mathcal{R}[X]$. This leads to results about P on not necessarily integral or affine grids $\mathfrak{X}$. If, for example, $\mathcal{R} = \mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$ and $\hat{\mathcal{R}} = \mathbb{Z}$, we may read the following formula 8.1 modulo m (note that it contains only integer coefficients). $\mathbb{Z}_m$

Theorem 8.1. *Assume* $P \in \mathbb{Z}[X]$ *and* $\mathfrak{X} = [d] := [d_1] \times \cdots \times [d_n]$.
If $\deg(P) \leq \Sigma d$, *then*

$$\boxed{(-1)^{\Sigma d}\left[\prod_{j\in(n]}(d_j!)\right] P_d \;=\; \sum_{x\in\mathfrak{X}} \left[\prod_{j\in(n]}(-1)^{x_j}\binom{d_j}{x_j}\right] P(x)} \ .$$

Proof. This follows from Theorem 3.3 and Lemma 1.4 (v). □

With this theorem we get the following special version of Corollary 3.4, which works perfectly well without a degree condition. (See [MuSt] and [Sp] for more information about polynomial maps $\mathbb{Z}_m^n \to \mathbb{Z}_m$.)

Corollary 8.2. *Let* $P \in \mathbb{Z}_m[X]$ *, and set* $\mathfrak{X} := \mathbb{Z}_m^{\,n}$ *, which we identify with* $[m)^n \subseteq \mathbb{Z}^n$*. If* m *is not prime, and* $(m,n) \neq (4,1)$ *, then:*

(i) $\boxed{\big|\{\, x \in \mathfrak{X} \mid P(x) \neq 0 \,\}\big| \,\neq\, 1}$.

(ii) $\boxed{P_0 \neq 0 \implies \exists\, x \in \mathfrak{X}\backslash 0 : \big[\prod_{j\in(n]} \binom{m-1}{x_j}\big]\, P(x) \neq 0 \implies P|_{\mathfrak{X}\backslash 0} \not\equiv 0}$.

(iii) $\boxed{0 \,=\, \sum_{x\in\mathfrak{X}} \big[\prod_{j\in(n]} (-1)^{x_j} \binom{m-1}{x_j}\big]\, P(x)}$.

Proof. Suppose there is an $\hat{x} \in \mathfrak{X} = \mathbb{Z}_m^{\,n}$ with $P(\hat{x}) \neq 0$. By applying the substitutions $X_j = X_j + \hat{x}_j$, we may assume $0 \neq P(0) = P_0$; and part (i) follows from the compounded implication (ii).

Part (ii) follows from part (iii), as the summand $[\prod_{j\in(n]} (-1)^0 \binom{m-1}{0}] P(0) = P_0$ cannot be the only nonvanishing summand in the vanishing sum.

To prove part (iii), we may assume that P has partial degrees $\deg_j(P) \leq d_j = d_j(\mathfrak{X})$. This is so, as the monic polynomial $L_j := \prod_{x\in\mathfrak{X}_j} (X_j - x)$ maps $\mathfrak{X}_j$ to 0, so that we may replace P with any polynomial of the form $P + \sum_{j\in(n]} H_j L_j$ without changing its image $P|_{\mathfrak{X}}$ (see the Example 7.1 and (82) for an illustration of this method). Now let $\hat{P} \in \mathbb{Z}[X]$ be such that

$$P \,=\, \hat{P} + m\mathbb{Z}[X] \;\in\; \mathbb{Z}[X]/m\mathbb{Z}[X] \,=\, \mathbb{Z}_m[X] \qquad \text{and} \qquad deg_j(\hat{P}) \,\leq\, d_j \;. \tag{86}$$

We only have to show that $m \,\lfloor\, (m-1)!^{\,n}$, so that the left side of Equation 8.1, applied to $\hat{P}$, vanishes modulo m , in the relevant case $d_1, \ldots, d_n = m-1$:

If $m \neq 4$ and $m = m_1 m_2$, with $m_1 < m_2 < m$, then $m \,\lfloor\, (m-1)!$.
If $m \neq 4$ and $m = p^2$ with $p > 2$, then $p < 2p < m$ and so $m \,\lfloor\, p\,(2p) \,\lfloor\, (m-1)!$.
If $m = 4$ and $n \geq 2$, then $m = 2^2 \,\lfloor\, 3!^{\,2} \,\lfloor\, (m-1)!^{\,n}$.

□

The examples $X^3 + X + 2$ and $X^3 - 2X^2 - X + 2 \in \mathbb{Z}_4[X]$ show that the very special case $(m,n) = (4,1)$ in 8.2 is really an exception. As one can show, these two examples are the only exceptions to assertion (i) that fulfill the additional normalization conditions $\deg(P) \leq 3$, $P_3 \neq -1$ and that the nonvanishing point is the zero ($P(x) \neq 0 \Leftrightarrow x = 0$).

We also present another version of Corollary 3.5 . For this, we will need the following specialization of [AFK2, Lemma A.2]:

Lemma 8.3. *Let* $p \in \mathbb{N}$ *be prime,* $k > 0$ *and* $c = c(p^k) := \sum_{i\in[k)}(p^i - 1)$ *. For* $y \in \mathbb{Z}$ *,*

(i) $p^c \,\lfloor\, \prod_{0<\hat{y}<p^k} (y - \hat{y})$ *, and*

(ii) $p^{c+1} \not\lfloor \prod_{0<\hat{y}<p^k} (y - \hat{y}) \iff p^k \lfloor y$ *.*

For completeness, we present the relatively short proof:

Proof. For each $j \in (k]$ there are exactly p^{k-j} numbers among the p^k consecutive integers y , $y-1$, $\dots$, $y-(p^k-1)$ that are dividable by p^j. Thus:

If $p^k \lfloor y$, then exactly $p^{k-j}-1$ of the factors $y - \hat{y}$ in the product $\prod_{0<\hat{y}<p^k} (y - \hat{y})$ are dividable by p^j.

If $p^k \not\lfloor y$, then at least $p^{k-j}-1$ of these factors are dividable by p^j ; and in the case $j = k$, strictly more than $p^{k-j}-1 = 0$ are multiples of $p^j = p^k$.

It follows:

If $p^k \lfloor y$, then $p^c \lfloor \prod_{0<\hat{y}<p^k}(y-\hat{y})$, but $p^{c+1} \not\lfloor \prod_{0<\hat{y}<p^k}(y-\hat{y})$.

If $p^k \not\lfloor y$, then $p^{c+1} \lfloor \prod_{0<\hat{y}<p^k}(y-\hat{y})$.

□

The following version of Corollary 3.5 (see also Theorem 4.3 and [Scha2]) reduces to Olson's Theorem [AFK2, Theorem 2.1], if $\deg(P_1) = \dots = \deg(P_m) = 1$ and if we set $\mathfrak{X} := \{0,1\}^n$. Olson's Theorem can be used, for example, to prove generalizations of Theorem 4.1 about regular subgraphs, such as those in [AFK2]. Here we view, more generally, arbitrary polynomials and arbitrary *p-integral* grids – i.e., grids $\mathfrak{X} \subseteq \mathbb{Z}^n$ with the property:

$$\text{For all } j \in (n] \text{ and all } x, \tilde{x} \in \mathfrak{X}_j \text{ with } x \neq \tilde{x}, \quad p \not\lfloor x - \tilde{x} \ . \tag{87}$$

We have:

Theorem 8.4. *Let* $p \in \mathbb{N}$ *be a prime and* $\mathfrak{X} \subseteq \mathbb{Z}^n$ *a p-integral d-grid.*
For polynomials $P_1, \dots, P_m \in \mathbb{Z}[X_1, \dots, X_n]$ *, and numbers* $k_1, \dots, k_m > 0$ *small enough so that* $\sum_{i\in(m]}(p^{k_i}-1)\deg(P_i) < \Sigma d$ *,*

$$\boxed{\left|\left\{\, x \in \mathfrak{X} \;\middle\vert\; \forall\, i \in (m]:\ p^{k_i} \lfloor P_i(x) \,\right\}\right| \neq 1} \ .$$

Proof. Set

$$c \;:=\; \sum_{i\in(m]} c_i \quad \text{where} \quad c_i \;=\; c(p^{k_i}) \;:=\; \sum_{j\in[k_i)} (p^j - 1) \;\; , \tag{88}$$

define

$$P \;:=\; \prod_{i\in(m]} \; \prod_{0<\hat{y}<p^{k_i}} (P_i - \hat{y}) \;\in\; \mathbb{Z}[X] \tag{89}$$

and let

$$\bar{P} \;:=\; P + p^{c+1}\mathbb{Z}[X] \;\in\; \mathbb{Z}[X]/p^{c+1}\mathbb{Z}[X] \;=\; \mathbb{Z}_{p^{c+1}}[X] \;\; . \tag{90}$$

For points $x = (x_1, \ldots, x_n) \in \mathbb{Z}^n$, set

$$\bar{x} \;:=\; (x_1 + p^{c+1}\mathbb{Z}\,, \ldots, x_n + p^{c+1}\mathbb{Z}) \;\in\; (\mathbb{Z}_{p^{c+1}})^n \;\; ; \tag{91}$$

then

$$\bar{\mathfrak{X}} \;:=\; \{\,\bar{x} \;\mid\; x \in \mathfrak{X}\,\} \;\subseteq\; (\mathbb{Z}_{p^{c+1}})^n \tag{92}$$

is an integral d-grid, and $x \longmapsto \bar{x}$ induces a bijection from $\mathfrak{X}$ to $\bar{\mathfrak{X}}$.
Now it follows that

$$\begin{aligned} \bar{P}(\bar{x}) \neq 0 \quad &\Longleftrightarrow \quad p^{c+1} \nmid P(x) \\ &\overset{8.3\,(i)}{\Longleftrightarrow} \quad \forall\, i\colon\; p^{c_i+1} \nmid \prod_{0<\hat{y}<p^{k_i}} (P_i(x) - \hat{y}) \\ &\overset{8.3\,(ii)}{\Longleftrightarrow} \quad \forall\, i\colon\; p^{k_i} \mid P_i(x) \;\; , \end{aligned} \tag{93}$$

and since

$$\deg(\bar{P}) \;\leq\; \deg(P) \;\leq\; \sum_{i\in(m]} (p^{k_i} - 1)\deg(P_i) \;<\; \Sigma d \;\; , \tag{94}$$

we obtain

$$\left|\left\{\, x \in \mathfrak{X} \;\mid\; \forall\, i \in (m]\colon\; p^{k_i} \mid P_i(x) \,\right\}\right| \;=\; \left|\left\{\, \bar{x} \in \bar{\mathfrak{X}} \;\mid\; \bar{P}(\bar{x}) \neq 0 \,\right\}\right| \;\overset{3.4}{\neq}\; 1 \;\; . \tag{95}$$

□

Our result can be generalized further, in the obvious way, by using [AFK2, Lemma A.2], instead of our Lemma 8.3. However, the result would look a bit more technical. In [AFK2] Alon, Friedland and Kalai also made the following conjecture:

Conjecture 8.5. *Set* $\mathfrak{X} := \{0,1\}^n$ *and let* $P_1, \ldots, P_m \in \mathbb{Z}[X_1, \ldots, X_n]$ *be homogenous polynomials of degree* 1. *If* $k \in \mathbb{N}$ *is small enough so that* $(k-1)\,m \,<\, n$, *then*

$$\boxed{\left|\left\{\, x \in \mathfrak{X} \;\mid\; \forall\, i \in (m]\colon\; k \mid P_i(x) \,\right\}\right| \;\neq\; 1} \;\; .$$

Acknowledgement:
We are grateful for interesting and useful comments by the referee. He also helped, together with Alexandra Kallia and Michael Harrison, to improve my English. Many thanks to all.

References

[Al] N. Alon: Restricted colorings of graphs.
In "Surveys in combinatorics, 1993", London Math. Soc. Lecture Notes Ser. 187, Cambridge Univ. Press, Cambridge 1993, 1-33.

[Al2] N. Alon: Combinatorial Nullstellensatz.
Combin. Probab. Comput. 8, No. 1-2 (1999), 7-29.

[Al3] N. Alon: Discrete Mathematics: Methods and Challenges.
Proc. of the International Congress of Mathematicians (ICM), Beijing 2002, China, Higher Education Press (2003), 119-135.

[AFK] N. Alon, S. Friedland, G. Kalai: Every 4-regular graph plus an edge contains a 3-regular subgraph. *J. Combin. Theory Ser. B 37 (1984), 92-93.*

[AFK2] N. Alon, S. Friedland, G. Kalai: Regular subgraphs of almost regular graphs.
J. Combin. Theory Ser. B 37 (1984), 79-91.

[AlFü] N. Alon, Z. Füredi: Covering the cube by affine hyperplans.
European J. Combinatorics 14 (1993), 79-83.

[AlTa] N. Alon, M. Tarsi: Colorings and orientations of graphs.
Combinatorica 12 (1992), 125-134.

[AlTa2] N. Alon, M. Tarsi: A nowhere-zero point in linear mappings.
Combinatorica 9 (1989), 393-395.

[ApHa] K. I. Appel, W. Haken, J. Koch: Every planar map is four colorable.
Illinois J. Math. 21 (1977), 429-567.

[BrRy] R. A. Brualdi, H. J. Ryser: Combinatorial matrix theory.
Cambridge University Press, Cambridge 1991.

[CCF] P. J. Cahen, J. L. Chabert, S. Frisch: Interpolation domains.
J. Algebra 225 (2000), 794–803.

[CCS] J. L. Chabert, S. T. Chapman, W. W. Smith:
The Skolem Property in rings of integer-valued polynomials.
Proceedings of the American Mathematical Society, Vol. 126 No. 11 (1998), 3151-3159.

[Da] P. J. Davis: Interpolation and approximation.
Dover Books on Advanced Mathematics. New York 1975.

[DeV] M. DeVos: Matrix choosability. *J. Combin. Theory Ser. A 90 (2000), 197-209.*

[Di] R. Diestel: Graph theory. *Springer, Berlin 2000.*

[DuFo] D. S. Dummit, R. M. Foote: Abstract Algebra.
John Wiley and Sons, Inc. 2004.

[ElGo] M. N. Ellingham, L. Goddyn: List edge colourings of some 1-factorable multigraphs. *Combinatorica 16 (1996), 343-352.*

[JeTo] T. R. Jensen,B. Toft: Graph coloring problems. *Wiley, New York 1995.*

[Minc] H. Minc: Permanents. *Addison-Wesley, London 1978.*

[MSCK] O. Moreno, K. W. Shum, F. N. Castro, P. V. Kumar:
Tight bounds for Chevalley-Warning-Ax-Katz type estimates, with improved applications. *Proc. London Math. Soc. (3) 88 (2004), 545-564.*

[MoZi] O. Moreno, V. A. Zinoviev: Tree-regular subgraphs of four-regular graphs.
European J. Combinatorics 19, No. 3, (1998), 369-373.

[MuSt] G. Mullen, H. Stevens: Polynomial functions (mod m).
Acta Math. Hung. 44 (1984), 237-241.

[Scha] U. Schauz: Colorings and orientations of matrices and graphs.
The Electronic Journal of Combinatorics 13 (2006), #R61.

[Scha2] U. Schauz: On the Dispersions of the Polynomial Maps over Finite Fields, and a Sharpening of Warning's Theorem.
Submitted to the Electronic Journal of Combinatorics.

[Scha3] U. Schauz: Mr. Paint and Mrs. Rubber: How To Find Nonzeros and Colorings of Polynomials. *In preparation.*

[Sch] D. E. Scheim: The number of edge 3-colorings of a planar cubic graph as a permanent. *Discrete Math. 8 (1974), 377-382.*

[Sp] R. Spira: Polynomial interpolation over commutative rings.
Amer. Math. Monthly 75 (1968), 638–640.

[Vig] L. Vigneron: Remarques sur les réseaux cubiques de classe 3 associés au problème des quatre couleurs. *C. R. Acad. Sci. Paris T. 223 (1946), 770-772.*

[Ya] Yu Yang: The permanent rank of a matrix.
J. Combin. Theory Ser. A 85(2) (1999), 237-242.