Aalto University
School of Science
Bachelor's Programme in Science and Technology

Detecting multilinear monomials with algebraic fingerprints

Bachelor's Thesis

March 12, 2023

Onni Miettinen

Aalto University
School of Science
Bachelor's Programme in Science and Technology

ABSTRACT OF
BACHELOR'S THESIS

Author:	Onni Miettinen
Title of thesis:	Detecting multilinear monomials with algebraic fingerprints
Date:	March 11, 2023
Pages:	x
Major:	Computer Science
Code:	SCI3027
Supervisor:	Prof. Eero Hyvönen
Instructor:	Augusto Modanese (Department of Computer Science)
abstract to be	
Keywords:	key, words, the same as in FIN/SWE
Language:	English

Contents

1 Introduction

In recent years, there have been rapid advances in the algorithms for combinatorial problems. This has been greatly sparked by the development in algebraic methods for solving the multilinear monomial detection problem, i.e., finding whether a multivariate polynomial contains a multilinear monomial. Namely, the technique of algebraic fingerprinting first introduced by Koutis in [Kou08] and further developed by Williams in [Wil09] has proven to be very successful.

With algebraic fingerprinting, the k-path problem (see problem def), that previously could be solved in $\mathcal{O}(something)$ time by X in [ref], could be solved in $\mathcal{O}^*(2^{3k/2})$ time in [Kou08]. This result was quickly improved in [Wil09], where a $\mathcal{O}^*(2^k)$ algorithm was given.

Of course, this technique has been further developed, and in [Bjö14] Björklund et al. showed an algorithm that solved the Hamiltonian problem (Hamiltonicity), i.e., finding whether a given graph contains a path that visits every vertex once, in $\mathcal{O}^*(something)$ time. The previous fastest algorithm for Hamiltonicity by Y in [ref] ran in $\mathcal{O}(something)$ time with the use of color coding[??]. [INSERT PREVIOUS MTEHOD] This was a significant improvement on a problem that had seen no progress in nearly fifty years.

The technique of algebraic fingerprinting is present in multilinear monomial detection. Multilinear monomial detection is a fundamental problem, since many important combinatorial problems can be reduced into multilinear monomial detection via a problem specific algebraization. The goal of such algebraization is to form a multivariate polynomial that encodes the combinations, i.e. the solutions and non-solutions, into multivariate monomials where multilinear monomials correspond to solutions to the given problem.

Before discussing multilinear monomial detection and algebraic fingerprints, the thesis covers algebraization, and shows how a combinatorial problem can be reduced into a multilinear monomial detection problem.

[TODO: go through the structure of the thesis]

1.1 Algebraization of combinatorial problems

A combinatorial problem asks whether a given finite set of objects satisfies some given constraints. For example, the k-path problem asks for, given a finite set of vertices and edges, a path of k vertices. The solutions and non-solutions to combinatorial problems can be thought of as combinations of the given objects. The solution space for the k-path problem consists of combinations of k vertices and k-1 edges. A non-solution combination would contain duplicate vertices or edges that contain vertices outside the combination.

Algebraization is reducing a given problem into an algebraic problem, i.e., a question regarding some algebraic property of some algebraic entity. In an algebraization of a combinatorial problem, the algebraic entity can be constructed from algebraic elements defined from the set of objects given as an input. The motivation behind the construction is some algebraic property that, when satisfied, gives a solution to the problem.

[TODO: rewrite this paragraph, explain with generating polynomial and expasion into sum of monomials] Multilinear monomial detection has proven to be a useful algebraization. First, we introduce multiple variables that correspond to elements from the set of objects given as input. Then, we construct a multivariate polynomial such that it encodes all solutions and non-solutions as multivariate monomials, with solutions encoded specifically as multilinear monomials. Thus, the task of finding a satisfying combination to the combinatorial problem has been reduced to finding a multilinear monomial from the multivariate polynomial. It follows, that a decision problem is answered by the existence of a multilinear monomial and a counting problem by the number of multilinear monomials.

Appropriate definitions for the variables are problem specific. In the following section, this thesis shows a reduction into multilinear monomial detection, introduced in [KW15], for the k-3D matching problem.

1.2 Reducing k-3D matching into multilinear monomial detection

The k-3D matching problem is defined as follows:

k-3D MATCHING

Input: Three disjoint sets A, B and C, and a set of triples $T \subset A \times B \times C$.

Question: Is there a subset $M \subseteq T$, such that $|M| = k$ and $\forall m \in M$: None of the elements in m appear in $M \backslash \{m\}$?

We begin by defining new variables corresponding to the elements in A, B and C, labeled as a_i, b_j and c_k, respectively, where $i \in [|A|]$, $j \in [|B|]$ and $k \in [|C|]$.

For every triple $t \in T$, we define a multilinear monomial x that is a product of the elements in t. We introduce a set X that satisfies the following:

$$\forall x \in X\colon\ x = abc :\ (a, b, c) \in T.$$

Next, we define multivariate polynomials P_1 and P_k as follows:

$$P_1 = \sum_X ,\ P_k = P_1^k.$$

Following this construction, we observe that P_k, when expanded into a sum of multivariate monomials, contains a multilinear term if and only if the original k-3D matching instance can be answered in the positive. Furthermore, every multilinear monomial in the expanded P_k corresponds to a solution to the problem, and the solutions can be directly found from the variables in the multilinear monomial. Thus, a successful reduction into multilinear monomial detection has been given for the k-3D mapping.

An example instance of k-3D matching with this exact algebraization can be found in [10.1145/2742544]. [show the example?]

2 Preliminaries

It is necessary to recall basic algebraic concepts before further discussing multilinear monomial detections and algebraic fingerprinting. In this section, definitions for a group, ring and field are given, and some useful concepts regarding them. The second subsection goes through general notation and terminology used throughout the thesis.

2.1 Groups, rings and fields

A group $\mathbf{G}$ is a tuple $(G, +)$, where G is a set of elements, $+ : G \times G \longrightarrow G$ is a binary operation closed under the elements in G, $+$ is associative, every element $g \in G$ has an inverse $g^{-1} \in G$, and G contains an identity element e such that $g + e = g$, $g + g^{-1} = e$ and $e = e^{-1}$. Moreover, $\mathbf{G}$ is called *Abelian* if $+$ is also commutative.

A ring $\mathbf{R}$ is a tuple $(R, \cdot)$, where $R = (G, +)$ is an Abelian group, $\cdot : G \times G \longrightarrow G$ is a binary operation closed under G. We call the binary operations $+$ and $\cdot$ as addition

and multiplication, respectively. Note, that from here on we use R as the set of elements defined for $\mathbf{R}$. In general, a bold typeface $\mathbf{X}$ represents a group, ring or field and X its set of elements. R must contain a multiplicative identity $\mathbf{1} \in R$ such that $\forall a \in R\colon\ a \cdot \mathbf{1} = a$. We notate the additive identity e required for the group R as $\mathbf{0}$ from here on. Observe, that for any $\mathbf{R} \neq \{\mathbf{0}\}$, $\mathbf{1} \neq \mathbf{0}$. Left and right distributive laws hold for rings, i.e.,

$$\forall a, b, c \in R\colon\ a \cdot (b + c) = (a \cdot b) + (a \cdot c) \wedge (b + c) \cdot a = (b \cdot a) + (c \cdot a).$$

$u \in R$ is called *unit* if it holds that $\exists v \in R\colon\ u \cdot v = v \cdot u = \mathbf{1}$, i.e., it has a multiplicative inverse $v \in R$.

A field $\mathbf{F} = (F, +, \cdot)$ is defined with the following conditions:

- $(F, +)$ is an Abelian group
- $(F \backslash \{\mathbf{0}\}, \cdot)$ is an Abelian group
- Left and right distributive laws hold for $\mathbf{F}$

Equivalently, a ring is a field if every non-zero element is unit, $\mathbf{1} \neq \mathbf{0}$, and multiplication is commutative. The *characteristic* of a field $\mathbf{F}$ is defined as follows:

$$char(\mathbf{F}) = \begin{cases} min\{n \in \mathbb{N} : n \cdot \mathbf{1} = \mathbf{0}\} & \\ 0 & \text{if such } n \text{ does not exist} \end{cases} \tag{1}$$

Note, that a field $\mathbf{F}$ with characeristic 2 satisfies the following:

$$\forall u \in F\colon\ u + u = u \cdot (\mathbf{1} + \mathbf{1}) = u \cdot \mathbf{0} = \mathbf{0}$$

TODO: polynomial rings, group algebra

2.2 Notation and other terminology

TODO: create a table or like, list terms: multilinearity, multivariety, sum of monomials form & generating form (arithmetic circuit) of polynomial, degree of multivariate monomial, $\mathcal{O}$, Θ, FPT, $\mathcal{O}^*$, determinism & non-determinism

3 General multilinear monomial detection

The detection of multilinear monomials is a fundamental problem, since many important problems can be reduced to it [TODO: quick examples (just refs?)]. Therefore, any

progress in general multilinear monomial detection directly implies faster algorithms for all problems, that are reduced to and solved with general multilinear monomial detection. The general, parameterized multilinear monomial problem is defined as follows:

k-MULTILINEAR MONOMIAL DETECTION

Input: A commutative arithmetic circuit A over a set of variables X.

Question: Does the sum of monomials form of the polynomial $P(X)$ represented by A contain a multilinear monomial of degree k?

3.1 Non-algebraic methods

Of course, fast evaluation of the expanded polynomial is also an important aspect, since it contributes directly to the general multilinear monomial detection problem. However, a naive expansion and evaluation is non-optimal. When the problem domain has n variables in an N-degree polynomial, the number of possible monomials is $\binom{n+N}{n} = 2^{\Theta(N)}$.

This motivates the detection of multilinear monomials without fully expanding the polynomial into a sum of monomials. On this section, the thesis gives a quick overview on non-algebraic methods for multilinear monomial detection that are faster than naive expansion and evaluation. These methods, however, are outperformed by the purely algebraic technique of algebraic fingerprinting which is covered later.

[TODO: motivate quickly why we still go through them anyway]

[dynamic programming is used in random color coding, random assignment is used in algebraic fingerprinting, dynamic programming's set rule of any var squared = 0 can be achieved with matrices]

3.1.1 Dynamic programming for smart expansion of the polynomial

In multilinear monomial detection, only the multilinear terms are important. This implies that any non-linear term can be discarded as soon as they are formed, since the degree will not decrease during the expansion of the polynomial. Therefore, a dynamic programming algorithm can be designed for the expansion with an additional rule: any squared variable can be instantly discarded, or in algebraic terms, set to zero.

With this additional rule, the following is deduced: if there are no solutions to the original problem, the polynomial will identically evaluate to zero. Note, however, that the polynomial evaluating to zero does not imply that no solutions exist. Indeed, we will come across a problem where multilinear monomials cancel each other, and thus evaluate

to zero (see chapther x).

3.1.2 Non-deterministic color coding for faster evaluation

[TODO: go over random assignment with colors, which results a polynomial of smaller domain (less variables), thus making the evaluation faster]

[TODO: end with hints toward matrix assignment (a non-zero matrix squared can equal zero)]

3.2 Non-deterministic color coding with matrices

[TODO: go through idea of matrix assignment and specifications for a suitable algebra, then lead to fingerprinting]

3.3 Algebraic fingerprinting to prevent unwanted cancellation

[TODO: add subsections for problem specific implementations with different utilizations of fingerprints]

3.4 Limits of general multilinear monomial detection

[TODO: explain the limit in general multilinear detection with this algebra (impossible to find better algebra than what is used for the current fastest k-mld algorithm)]

[TODO: lead to problem specific implementations (we can use fingerprinting cleverly, when we understand the underlying problem well)]

4 Problem specific implementations of algebraic fingerprints

4.1 Fingerprinting for cancellation of non-solutions

[TODO: go over Björklund et al. for k-path, managed to insert fingerprints such that non-solutions cancel]

References

[Kou08] Ioannis Koutis. "Faster Algebraic Algorithms for Path and Packing Problems". In: *Automata, Languages and Programming*. Ed. by Luca Aceto et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 575–586. ISBN: 978-3-540-70575-8.

[Wil09] Ryan Williams. "Finding paths of length k in O*(2k) time". In: *Information Processing Letters* 109.6 (2009), pp. 315–318. ISSN: 0020-0190. DOI: https://doi.org/10.1016/j.ipl.2008.11.004. URL: https://www.sciencedirect.com/science/article/pii/S0020019008003396.

[Bjö14] Andreas Björklund. "Determinant Sums for Undirected Hamiltonicity". In: *SIAM Journal on Computing* 43.1 (2014), pp. 280–299. DOI: 10.1137/110839229. eprint: https://doi.org/10.1137/110839229. URL: https://doi.org/10.1137/110839229.

[KW15] Ioannis Koutis and Ryan Williams. "Algebraic Fingerprints for Faster Algorithms". In: *Commun. ACM* 59.1 (Dec. 2015), pp. 98–105. ISSN: 0001-0782. DOI: 10.1145/2742544. URL: https://doi.org/10.1145/2742544.