AM/GM-Based Optimization: Geometry and Generalizations

DISSERTATION
ZUR ERLANGUNG DES DOKTORGRADES
DER NATURWISSENSCHAFTEN

vorgelegt beim Fachbereich 12
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Helen NAUMANN
aus Frankfurt am Main

Frankfurt, 2021
(D 30)

vom Fachbereich 12 der
Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr.-Ing. Lars Hedrich

Gutachter:

- Prof. Dr. Thorsten Theobald (Goethe-Universität Frankfurt am Main)
- Prof. Dr. Gennadiy Averkov (Brandenburgische Technische Universität Cottbus-Senftenberg)
- Prof. Dr. Timo de Wolff (Technische Universität Braunschweig)

Datum der Disputation: 12.11.2021

Authorship

I, Helen Naumann, hereby certify that the contents of this manuscript are either my own work or represent joint work of my co-authors and myself. All references have been quoted.

Various parts of this thesis are based on joint work with other authors. This thesis focusses on those statements, I had major contributions in. Chapter 3 is based on joint work with Lukas Katthän and Thorsten Theobald and is contained in [KNT21] – although the results were obtained in a different setting. Chapter 4 is based on parts of [Mou+21] and also partially on [Dre+20], the former is joint work with Philippe Moustrou, Cordian Riener, Thorsten Theobald, and Hugues Verdure, the latter is joint work with Mareike Dressler, Janin Heuer, and Timo de Wolff. Chapter 5 is based on joint work with Thorsten Theobald and contained in [NT21b], and Chapter 6 is based on the two works [NT21a] and on selected parts of [MNT20], where the former is joint work with Thorsten Theobald, and the latter is joint work with Riley Murray and Thorsten Theobald. Chapter 7 is my own work, except for the results on symmetries for the X-SAGE-cone, which come from [Mou+21].

Signature: ______________________________

Date: ______________________________

Acknowledgments

During this thesis, I was lucky to have many people to support me, consult me, guide me and encourage me. I am very grateful to have them in my life, and they now deserve to be mentioned in this place.

First and foremost I want to thank my Ph.D. advisor Thorsten Theobald, who encouraged me to stay at university and to work towards gaining a Ph.D., who already introduced me to the fascinating world of SONC and SAGE during my master thesis, who collaborated with me in so many projects by now, who motivated me to give talks and take on responsibility and who, most importantly, helped me face the challenges of research and giving talks so many times. My time as a Ph.D. student would not have been as great if he would not have been my advisor.

Moreover, I want to thank Gennadiy Averkov and Timo de Wolff for their willingness to act as reviewers for this thesis, and Andreas Bernig, Ralph Neininger, Anton Wakolbinger and Tobias Weth who agreed to participate in my doctoral committee.

I had the great opportunity to experience working with various groups on my different projects. Besides Thorsten Theobald, I also want to thank my other coauthors Mareike Dressler, Janin Heuer, Lukas Katthän, Philippe Moustrou, Riley Murray, Cordian Riener, Hugues Verdure, and Timo de Wolff, as well as all the great people I talked to during conferences and workshops. It was a pleasure working with each one of you!

Due to the current situation, the engagement with my working group at Goethe University was interrupted very early. Nevertheless, I want to thank Aenne Benjes, Giulia Codenotti, Stephan Gardoll, Anna Haacke, Max Hahn-Klimroth, Sven Jarohs, Lukas Katthän, Sebastian Manecke, Raman Sanyal, and Mahsa Sayyari for nice lunch chats and even nicer coffee breaks.

A special thanks goes to Giulia Codenotti, Mareike Dressler, Janin Heuer, and Theresa Kumpitsch who have been reading parts (or everything) of this thesis and whose comments have improved this work substantially. Also I would like to thank Melda Görür, Joel Kübler, and Laura Mader, whose councel I could always rely on.

And, of course, I want to thank my family — Pirmin, Sara, Sam, and my parents — for always cheering me up, reading various drafts of this work, or preventing several mental breakdowns.

Contents

Chapter 1

Motivation and Historical Background

Global or constrained optimization concern the question of finding the minimum value of a given real function $f : \mathbb{R}^n \to \mathbb{R}$ either over $\mathbb{R}^n$ or over some subset $X \subsetneq \mathbb{R}^n$. This problem occurs in many branches of mathematics and various fields of application. The similar question of deciding whether a real function only takes nonnegative values is a fundamental question in real algebraic geometry. Both problems can be treated as equivalent: The infimum of a function f is the largest real scalar λ such that the function $f - \lambda$ is globally nonnegative, i.e.,

$$f^* = \inf\{f(x) \ : \ x \in \mathbb{R}^n\} \ = \ \sup\{\lambda \in \mathbb{R} \ : \ f - \lambda \text{ is nonnegative on } \mathbb{R}^n\}. \qquad (1.1)$$

In this thesis, our functions of interest are exponential sums $f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle x, \alpha \rangle}$ and sometimes polynomials, which can be seen as a special case of exponential sums: When $\mathcal{A} \subseteq \mathbb{N}^n$, the transformation $x_i = \ln y_i$ gives polynomial functions $y \mapsto \sum_{\alpha \in \mathcal{A}} c_\alpha y^\alpha$ on the positive orthant $\mathbb{R}^n_{>0}$. Nonnegative polynomials or exponential sums and optimization over both types of functions are ubiquitous in applications, and sparsity is one of the central structural properties that provide potential for efficient computation. Besides classical application in control theory and robotics (see, e.g., [HG05], [AM19] and the references therein), the more recent applications of nonnegative polynomials and polynomial optimization can be seen in the optimal power flow problem [Jos16], collision avoidance [AM16], shape-constrained regression [Hal18], chemical reaction networks [Mül+15; MHR19], aircraft design optimization [ÖS19; YHD18], and epidemiological process control [NPP17; Pre+14]; see also [EPR20] and the references therein.

Computing the minimal value f^* from (1.1) of a given real function is NP-hard [MK87], even in the polynomial case [Lau09]. Hence, the idea is to search for efficiently computable sufficient conditions for nonnegativity — so called *nonnegativity certificates*. Ideally, a huge amount of elements in the nonnegativity cone satisfy these conditions. A well-known and large subset of nonnegative polynomials are *SOS-polynomials* — sums of squares of other polynomials. They provide a relaxation for finding the minimal value of a given polynomial. Nonnegativity of *these* polynomials can be certified via semidefinite programming [Las00; Par00]; for a broader overview of this topic see Chapter 2. A problem, however, is that a possibly unbounded degree can occur in an SOS-decomposition causing a possibly infinite size of the semidefinite program.

Recently, several researchers have developed sufficient conditions for nonnegativity based on the (weighted) arithmetic-geometric (AM/GM) mean inequality. While sum of squares nonnegativity certificates work with the degree of a polynomial and, in particular, require integer exponents, this is not the case for the techniques based on the

AM/GM inequality. Hence, they also work for exponential sums. Building upon previous results by Reznick [Rez89], these techniques were developed further in the works of Pantea, Koeppl, and Craciun [PKC12], Iliman and de Wolff [IW16a], and Chandrasekaran and Shah [CS16]. In contrast to SOS-based certificates, AM/GM-based nonnegativity certificates preserve sparsity of a given exponential sum or polynomial [MCW21a; Wan18a].

To introduce some notation, an exponential sum (or *signomial*) supported on a finite set $\mathcal{A} \subseteq \mathbb{R}^n$ is a sum $\sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle}$ with real coefficients c_α. Observe that for a set of support points $A \subseteq \mathbb{R}^n, \beta \in \mathbb{R}^n$ and coefficients $\lambda \in \mathbb{R}_+^A$ satisfying $\mathbf{1}^T \lambda = 1$ and $\sum_{\alpha \in A} \lambda_\alpha \alpha = \beta$, the weighted AM/GM inequality yields

$$\sum_{\alpha \in A} \lambda_\alpha e^{\langle \alpha, x \rangle} \geqslant \prod_{\alpha \in A} e^{\lambda_\alpha \langle \alpha, x \rangle} = e^{\langle \beta, x \rangle}.$$

Hence, nonnegativity of exponential sums of the form

$$\sum_{\alpha \in A} \lambda_\alpha e^{\langle \alpha, x \rangle} - e^{\langle \beta, x \rangle}$$

and of similar forms on $\mathbb{R}^n$ can be certified via the AM/GM inequality.

The setting that is relevant in this thesis has been introduced under various names: In 2016, Iliman and de Wolff introduced the concept of *sums of nonnegative circuit polynomials* (SONC) [IW16a]. A circuit is a tuple (A, β) with $A \subseteq \mathbb{R}^n$ affinely independent and $\beta \in \operatorname{relint} \operatorname{conv}(A)$, and a circuit polynomial is a polynomial with exponents in $A \cup \{\beta\}$ such that coefficients corresponding to A are nonnegative. A polynomial supported on a circuit can only be *globally nonnegative* if $A \subseteq (2\mathbb{N})^n$; for a proof, see, e.g., [Fel+20].

Later in the same year, Chandrasekaran and Shah introduced the concept of *sums of arithmetic-geometric exponentials* (SAGE) [CS16]. An arithmetic-geometric exponential (AGE exponential, sometimes also called AM/GM exponential) is an exponential sum with at most one negative coefficient. This definition exactly covers the previously mentioned subclass of exponential sums whose nonnegativity can be certified by the AM/GM inequality (explaining the name of this class of exponentials).

In various articles, the SONC approach was used for a relaxation of the global optimization problem (1.1), i.e., the nonnegativity condition in the problem formulation was replaced by a SONC condition

$$f^{\text{sonc}} = \sup\{\lambda \in \mathbb{R} \ : \ f - \lambda \text{ is SONC}\} \ \leq \ \inf\{f(x) \ : \ x \in \mathbb{R}^n\}. \tag{1.2}$$

Iliman, de Wolff, and Dressler et al. examined a polynomial nonnegativity certificate called the *circuit number*, resulting in a geometric program for certain subclasses of SONC polynomials [IW16a; IW16b; DIW17; DKW21]. Karaca et al. studied a combined SONC and Sums-of-Squares approach to polynomial optimization on the nonnegative orthant [Kar+17]. Chandrasekaran, Shah, Murray and Wierman focused on relative entropy programming approaches for both exponential sums and polynomials [CS16; CS17; MCW21a], which can be computed efficiently and can be used for general classes of exponential sums. Seidler and de Wolff proposed an algorithm to compute circuit decompositions [Sd18]. Together with Magron [MSW19], they compared the existing approaches on SONC-, SAGE- and SOS-based relaxations of global optimization problems with each other. Using similar approaches to those discussed in Chapter 3 of this thesis, Papp developed an algorithm for computing the optimal circuit decomposition of a given polynomial [Pap19]. Many of these optimization approaches are implemented in [Sd19]. In 2019, Forsgård and de Wolff examined the algebraic boundary of the SONC-cone using a new setting (for exponential sums, [FW19]) using the theory of regular subdivisions, A-discriminants, and tropical geometry.

Besides the two formerly discussed optimization approaches, there also exists a relation to semidefinite programming. Similar to [Kar+17], Averkov combined SONC approaches with semidefinite approaches to polynomial optimization. He was the first to prove that the cone of SONC polynomials can be represented as the projection of a spectrahedron [Ave19]. Wang and Magron provided an explicit representation of the primal SONC-cone as a second order-cone program [WM20a].

Already in 2016, Iliman and de Wolff showed that global nonnegativity of a circuit polynomial can be reduced to nonnegativity of an exponential sum with the same support and potentially slightly modified coefficients [IW16a]. In 2018, Murray, Chandrasekaran, and Wierman proved that this holds for general polynomials with at most one negative term (i.e., at most one term $c_\beta x^\beta$ with $c_\beta < 0$ or $\beta \notin (2\mathbb{N})^n$) [MCW21a]. Wang observed earlier in the same year that the polynomial form of an AGE exponential is a SONC polynomial [Wan18a], and Murray, Chandrasekaran, and Wierman gave an independent proof of the explicit statement that the extreme rays of the cone of AGE exponentials are supported on circuits (again, see [MCW21a]). Both thus showed the equivalence of SONC and SAGE respective to the setting of polynomials or exponential sums.

Hence, for results that are relevant for both SONC and SAGE, it is natural to examine only the simpler setting of exponential sums and discuss it under a common name. Building upon the approach of exponential sums introduced by Chandrasekaran and Shah, we use the name *AGE exponentials* for exponential sums with at most one negative term, and *SAGE exponential* for sums of these AGE exponentials. Note that the term *exponential* is here explicitely included in the name in contrast to [CS16] — to stress the setting. Moreover, contributing to the terminology introduced by Iliman and de Wolff, we call AGE exponentials supported on a circuit where the possibly negative term corresponds to the inner exponent of the circuit *circuit exponentials*. These functions are of particular interest because — as mentioned above — all extremal rays of the SAGE-cone are supported on circuits. In this thesis, we put a particular focus on circuits by examining the duality theory and extremality of the SAGE-cone as well as various optimization approaches. We also consider all these topics for constrained optimization, i.e., optimization over a subset $X \subsetneq \mathbb{R}^n$.

For SOS polynomials, there are various Positivstellensatz results such as *Putinar's Positivstellensatz* [Put93]. This statement shows that under certain preconditions, a polynomial which is nonnegative over a set of given constraints can be decomposed using SOS and this constraint set. For the SONC and SAGE case, such a Putinar-like Positivstellensatz unfortunately does not exist [DKW21], hence, when it comes to examining the constrained case, the SONC and SAGE approaches suffer from a severe disadvantage compared to the SOS approach. Dressler et. al provided a different Positivstellensatz for SONC polynomials, which was later identified as a special case of the Krivine-Positivstellensatz.

In 2019, Murray, Chandrasekaran, and Wierman proposed another approach to constrained optimization for both exponential sums and polynomials [MCW21b] that results in a certificate for nonnegativity using relative entropy programming. Using the dual Lagrangian they arrived — similarly to the unconstrained case — at a relative entropy program. The underlying set of constraints is represented by a so-called support function. With this relative entropy program, one can find a hierarchy to approximate the optimal value. Wang et al showed that in fact this hierarchy is complete [Wan+20] (see also [DP15]). The implemented program can be found in [Mur20], using the solver MOSEK [DA21].

Further related work considers the exploitation of sparsity and symmetries to derive specific SDP relaxations for polynomial optimization [KKW05; MCD17; Rie+13; WML21b; WML21a; WLT18].

In this thesis, we examine the SAGE-cone, its geometry, and generalizations of it. The thesis consists of three main parts:

1. In the first part, we focus purely on the cone of sums of globally nonnegative exponential sums with at most one negative term, the SAGE-cone. We examine the duality theory, extreme rays of the cone, and provide two efficient optimization approaches over the SAGE-cone and its dual.

2. In the second part, we introduce and study the so-called $\mathcal{S}$-cone, which provides a uniform framework for SAGE exponentials and SONC polynomials. In particular, we focus on second-order representations of the $\mathcal{S}$-cone and its dual using extremality results from the first part.

3. In the third and last part of this thesis, we turn towards examining the *conditional* SAGE-cone. We develop a notion of *sublinear circuits* leading to new duality results and a partial characterization of extremality. In the case of polyhedral constraint sets, this examination is simplified and allows us to classify sublinear circuits and extremality for some cases completely. For constraint sets with certain conditions such as sets with symmetries, conic, or polyhedral sets, various optimization and representation results from the unconstrained setting can be applied to the constrained case.

1.1 Extremality and Duality Theory of the SAGE-Cone

While many aspects of the classes of SAGE exponentials and SONC polynomials are the subject of open questions and research efforts, they clearly exhibit some fundamental structural phenomena that work well with sparse settings. Building upon the earlier work of the author in [DNT21] on the dual cone of SONC polynomials, we start by examining the dual cone of SAGE exponentials. In particular, we derive a projection-free representation of the dual cone of SAGE exponentials and use this representation to completely characterize extreme rays of the *primal* SAGE-cone. These results can be used thereafter to examine efficient optimization approaches based on the SAGE-cone.

We introduce the concept of *reduced circuits*. Reduced circuits are circuits that do not have additional points of the overall support contained in its convex hull. Using this concept, we provide a comprehensive characterization of the dual of the SAGE-cone, see Theorem 3.1.5. In particular, we provide projection-free characterizations in terms of AGE exponentials supported on the particular class of reduced circuits. The characterizations of the dual cone go far beyond the characterizations of the dual cone of SAGE exponentials from [CS16] and the dual cone of SONC polynomials from [DNT21], where the dual cones are described in terms of projections. Our proofs provide a uniform tool set for handling the various types of cones.

Based on the characterizations of the dual of the SAGE-cone, we show that every SAGE exponential can be written as a sum of nonnegative circuit exponentials supported on reduced circuits, see Theorem 3.2.1, and provide an exact characterization of the extreme rays of the SAGE-cone, see Theorem 3.2.4. This characterization substantially sharpens the necessary conditions in [MCW21a] (also see [Wan18a]).

Symmetry Reduction in AM/GM-based Optimization

From an algebraic point of view, a problem is *symmetric* when it is invariant under some group action. Symmetries are ubiquitous in the context of polynomials or exponential sums and optimization since they manifest both in the problem formulation and the solution set. This often makes it possible to reduce the complexity of the corresponding algorithmic questions. Regarding the set of solutions, in 1840 it was observed by Terquem that a symmetric polynomial does not always have a fully symmetric minimizer (see also Waterhouse's survey [Wat83]). However, in many instances, the set of minimizers contains highly symmetric points, see, e.g., [FRS18; MRV21; Rie12; Tim03]. With respect to the problem formulation, symmetry reduction has provided essential advances in many situations, see, e.g., [BV08; KS10; DV15], especially in the context of sums of squares, see, e.g., [Bac+12; BR21; DR20; GP04; HHS21; Ray+18; Rie+13].

We examine to which extent symmetries can be exploited in AM/GM-based optimization assuming that the problem affords symmetries. With this, we provide a first systematic study of the AM/GM-based approaches in G-invariant situations under the action of a group G.

We prove a symmetry-adapted decomposition theorem and develop a symmetry-adapted relative entropy formulation of SAGE exponentials in a general G-invariant setting. This adaption reduces the size of the resulting relative entropy programs or geometric programs, see Theorem 4.2.1, Theorem 4.2.3, and Corollary 4.2.6. As revealed by these statements, the gain depends on the orbit structure of the group action.

We evaluate the structural results in the thesis in terms of computations. In situations with strong symmetry structure, the number of variables and the number of equations and inequalities becomes substantially smaller. Accordingly, the interior-point solvers underlying the computation of SAGE bounds then show strong reductions in computation time. In various cases, the symmetry-adapted computation succeeds when the conventional SAGE computation fails.

We mostly focus on unconstrained optimization but the techniques can also generally be extended to the constrained case, see, e.g., Corollary 7.1.2.

Global Optimization via the Dual SAGE-Cone and Linear Programming

Using the dual cone of sums of AGE exponentials, we provide a relaxation of the global optimization problem to minimize an exponential sum and, as a special case, a multivariate real polynomial. The key idea of this optimization approach is to relax the problem (1.1) via optimizing over a cone with coefficients induced by the dual SONC-cone. Our approach is motivated by the recent works [DNT21], [MCW21a], and Chapter 3, and builds on two key observations, which are the main theoretical contributions:

1. The dual cone of AGE exponentials is contained in the primal one; see Proposition 4.3.1. Also, a variation of the dual SAGE-cone is contained in the primal SAGE-cone.

2. Optimizing over this modified dual cone can be carried out by solving a linear program; see Proposition 4.3.4.

We emphasize that neither the primal nor the dual SAGE-cone is polyhedral; in this context see also the results in [FW19]. The approach works as follows: First, we investigate a lifted version of the dual cone involving additional linear auxiliary variables (Theorem 4.1.6 (3)). Second, we show that the coefficients of a given exponential sum can be interpreted as being induced by variables in the dual cone; see (4.21). Third, we observe that fixing these coefficient variables yields a linear optimization problem only involving auxiliary variables; see Proposition 4.3.4.

Based on our two key observations stated above, we present two linear programs solving a relaxation of (1.1).

1.2 A Primal-Dual View on Second-Order Representability

As explained earlier, the cones of sums of arithmetic-geometric exponentials and sums of nonnegative circuit polynomials provide nonnegativity certificates based on the arithmetic-geometric inequality and are particularly useful in the context of sparse polynomials and exponential sums.

In Chapter 5, we introduce and study a cone which consists of a class of generalized polynomial functions and which provides a common framework for recent nonnegativity certificates of polynomials and exponntial sums in sparse settings. Specifically, this $\mathcal{S}$-cone generalizes and unifies both the cone of SONC polynomials and the cone of SAGE exponentials. In particular, several results in the context of these cones — such as the characterizations of the dual cone as well as of the extreme rays — can be transfered to the $\mathcal{S}$-cone.

Since nonnegativity of a polynomial function $f(x_1, \ldots, x_n)$ on $\mathbb{R}^n_+$ is equivalent to nonnegativity of $f(|x_1|, \ldots, |x_n|)$ on $\mathbb{R}^n$, we consider the more general functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ of the form

$$f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta, \tag{1.3}$$

with sets of exponents $\mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$, which also include exponential sums. Based on a subset of these functions, we define the *$\mathcal{S}$-cone* $C_\mathcal{S}(\mathcal{A}, \mathcal{B})$ which provides the common generalization of the cones mentioned above, see Definition 5.1.3. Elements in this $\mathcal{S}$-cone are called *AG functions* or *sums of AG functions.* These AG functions are functions of the form (1.3) with strong support conditions. They can be seen as a (non-polynomial) generalization of polynomials coming from the arithmetic-geometric inequality.

One motivation for defining this class of functions (1.3) is that it allows to capture nonnegativity of polynomials on $\mathbb{R}^n$ and nonnegativity of polynomials on the nonnegative orthant $\mathbb{R}^n_+$ within a uniform setting. Moreover, global nonnegativity of the summand $\sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha$ is equivalent to global nonnegativity of the exponential sum $y \mapsto \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, y \rangle}$.

Both from the geometric and from the optimization point of view, it is of prominent interest to understand how the different classes of cones relate to each other and whether techniques for different cones can be combined fruitfully. We remind the reader that, concerning relations between the various cones, Averkov has shown that the SONC-cone can be represented as a projection of a spectrahedron [Ave19]. In fact, his proof applies the techniques from [BN01], which reveal that the SONC-cone is even second-order representable. Wang and Magron gave an alternative proof based on binomial squares and $\mathcal{A}$-mediated sets [WM20b]. Both approaches only consider the primal SONC-cone.

We study the $\mathcal{S}$-cone and its dual from the viewpoint of second-order representability — and thus, also study its specilizations. Extending results of Averkov and of Wang and Magron on the primal SONC cone, we provide explicit generalized second-order descriptions for rational $\mathcal{S}$-cones and their duals and also take into account extremality results from Chapter 3 to reduce the size of these problems, see Corollaries 5.3.18 and 5.3.19. Our proof combines the techniques for the second-order representations from [BN01] with the concepts and the duality theory of Chapter 3 (the corresponding results in the language of the $\mathcal{S}$-cone were obtained in [KNT21]). Our derivation is different from the approach of Wang and Magron and it does not need binomial squares or $\mathcal{A}$-mediated sets. Moreover, our second-order representation prevents the involvement of redundant circuits by using a characterization of the extreme rays of the $\mathcal{S}$-cone from Chapter 3 (in the language of the $\mathcal{S}$-cone, again, see [KNT21]).

1.3 Sublinear Circuits and the Conditional SAGE-Cone

In this part of the thesis, we examine the constrained case, i.e., for a convex and non-empty set X, we consider the *constrained optimization problem*

$$f_X^* = \inf\{f(x) : x \in X\} = \sup\{\lambda : f - \lambda \geq 0 \text{ on } X\}$$

for an exponential sum f with sparse support $\mathcal{A} \subseteq \mathbb{R}^n$.

Historically, the conditional SONC relaxation of a constrained polynomial optimization problem examined decompositions of $f - \lambda$ into sums of nonnegative polynomials supported on classical $\mathbb{R}^n$-circuits [DKW21; DIW19]. However, in 2019, Murray, Chandrasekaran, and Wierman proposed a different approach, namely, examining when an exponential sum with at most one negative coefficient — not necessarily supported on a circuit — is nonnegative over X [MCW21b]. They used the fact that if we are interested in nonnegativity of an exponential sum f over X which has at most one negative coefficient c_β, i.e.,

$$f \ = \sum_{\alpha \in \mathcal{A}\setminus\{\beta\}} c_\alpha e^{\langle \alpha, x\rangle} + c_\beta e^{\langle \beta, x\rangle} \quad \text{with} \quad c_\alpha \geq 0 \ \text{ for all } \alpha \text{ in } \mathcal{A} \setminus \{\beta\},$$

then we can equivalently examine nonnegativity of $\sum_{\alpha\in\mathcal{A}} c_\alpha e^{\langle(\alpha-\beta),x\rangle}$. This is a convex function by construction. Hence, X-nonnegativity of this function can be exactly characterized by applying the principle of strong duality in convex optimization leading to a relative entropy program in a dual variable $\nu = (\nu_\alpha)_{\alpha\in\mathcal{A}}$, the exponential's coefficients, and involving the support function of X (see Proposition 2.4.12 for a precise statement). The same of course holds for the optimization formulation.

Following [MCW21b], this approach is called *conditional SAGE*. Whenever X is fixed, X-nonnegative exponential sums with at most one negative coefficient are called *X-AGE exponentials*, and the exponential sums which decompose into a sum of such functions are called *X-SAGE* in this thesis. Similarly to the unconstrained case, deciding whether a given function f is X-SAGE can also be determined in terms of a relative entropy program.

Even though this relative entropy program works perfectly without considering some circuit-like notion, it does not reveal anything about the structure of the conditional SAGE-cone. Similar to many results for the unconstrained case, partially covered in Chapter 3, see also [FW19; KNT21; MCW21a], we seek to understand extremality, duality and numerical issues of the cone. To do so, we establish the concept of *sublinear circuits*. Whenever the considered support $\mathcal{A}$ or constraint set X plays an important role, we also call these objects the X*-circuits of* $\mathcal{A}$. X-circuits of $\mathcal{A}$ are nonzero vectors $\nu^\star \in \mathbb{R}^{\mathcal{A}}$ at which the support function $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ exhibits a strict sublinearity condition (see Definition 6.1.2). This construction ensures that the special case of $\mathbb{R}^n$-circuits reduces to the simplicial circuits of the affine-linear matroid induced by $\mathcal{A}$, and even for real subsets $X \subsetneq \mathbb{R}^n$, sublinear circuits have affinely independent positive supports, compare Proposition 6.1.5.

Connecting the theory of sublinear circuits to X-nonnegative AGE exponentials, Theorem 6.2.2 reveals that for every X-AGE exponential, there exists a sublinear circuit serving as the dual variable in the relative entropy formulation. Moreover, every normalized X-circuit λ induces a λ*-witnessed* X*-AGE-cone* $C_X(\mathcal{A}, \lambda)$ satisfying some generalized variant of the circuit number condition ([IW16a]) from the unconstrained case. The generalization again involves the support function of X. The union of all these cones is the whole cone of X-AGE exponentials supported on $\mathcal{A}$, see Theorem 6.2.4. However, beyond these similarities to the unconstrained case, Example 6.2.9 reveals that we cannot recognize a circuit solely by its support.

Section 6.3 deals with the question of which X-circuits in fact are necessary for the representation of the X-AGE-cone as well as the X-SAGE-cone. To do so, we develop the notion of *reduced* X*-circuits* — namely those X-circuits ν for which $(\nu, \sigma_X(-\mathcal{A}\nu))$ generates an extreme ray of the circuit graph

$$\operatorname{pos}\left(\{(\lambda, \sigma_X(-\mathcal{A}\lambda)) \,:\, \lambda \text{ normalized } X\text{-circuit of } \mathcal{A}\} \cup \{(\mathbf{0}, 1)\}\right).$$

In fact, we can construct the conditional SAGE-cone using only λ-witnessed AGE-cones for reduced normalized X-circuits λ.

Sublinear circuits for polyhedral constraint sets

In the case of polyhedral constraint sets, the X-nonnegativity question simplifies substantially and yields interesting results. Hence, we put a particular focus on this situation. Here, the sublinear circuits can be characterized exactly in terms of the normal fan of a certain polyhedron, see Theorem 6.1.8. This induces a rich polyhedral-combinatorial structure and makes these sublinear circuits accessable for effective computations. For polyhedral X, the number of sublinear circuits is finite. This gives a decomposition of the X-SAGE-cone into *finitely* many X-AGE-cones, each induced by a sublinear circuit, see Theorem 6.3.6, paraphrased below.

> *Let X be a polyhedron. Denote by $\Lambda^\star_X(\mathcal{A})$ the set of normalized reduced X-circuits of $\mathcal{A}$. Assume $\lambda^\star_X(\mathcal{A})$ is non-empty and let the cone of X-SAGE exponentials consist of at least one non-positive term over X. Then, the cone equals the sum*
>
> $$\sum_{\lambda \in \Lambda^\star_X(\mathcal{A})} C_X(\mathcal{A}, \lambda).$$
>
> *Moreover, there is no real subset $\Lambda \subsetneq \Lambda^\star_X(\mathcal{A})$ of the set of normalized reduced X-circuits for which $\sum_{\lambda \in \Lambda^\star_X(\mathcal{A})} C_X(\mathcal{A}, \lambda) = \sum_{\lambda \in \Lambda} C_X(\mathcal{A}, \lambda)$.*

Theorem 6.3.6 provides the most efficient description possible of the X-SAGE-cone in terms of power cone inequalities.

Among the class of polyhedra, polyhedral cones exhibit particularly nice properties. Note that, as a very particular case, the unconstrained setting $X = \mathbb{R}^n$, which is treated in Chapter 3, also falls into the class of polyhedral cones. Every univariate case can be transformed to one of the two conic cases $\mathbb{R}$ (unconstrained case), $\mathbb{R}_+$ (one-sided infinity interval), or to the non-conic case $[-1, 1]$ (compact interval). In the multivariate case, the sets $\mathbb{R}^n$, $\mathbb{R}^n_+$ (nonnegative orthant), and the cube $[-1, 1]^n$ provide prominent examples for polyhedra. In contrast to the unconstrained case and to the nonnegative orthant, the cube $[-1, 1]^n$ provides a non-conic case.

Throughout Chapter 6, we illustrate key concepts with the help of the univariate compact interval $X = [-1, 1]$ and the half-line $X = [0, \infty)$. We examine both the X-circuits and the *reduced* X-circuits of a generic point set $\mathcal{A} \subseteq \mathbb{R}$ for both sets X. This culminates in a complete characterization of the extreme rays of the X-SAGE-cone for $X = [-1, 1]$ and $X = [0, \infty)$ with respect to $\mathcal{A} \subseteq \mathbb{R}$ (Theorem 6.4.1).

Constrained AM/GM-Based Optimization Approaches

To strengthen the connection to our optimization results, we finish this thesis by showing that certain optimization and decomposition results for the unconstrained setting also hold for the constrained case.

In particular, whenever the constraint set X is G-symmetric for a group G, one can find a symmetric decomposition of an X-AGE exponential leading to a relative entropy program with substantially reduced size, see Theorem 7.1.1 and Corollary 7.1.2.

In Section 7.2, we examine optimization over some cone induced by the dual of the X-SAGE-cone. Here, we need to make different restrictions on the constraint set X, namely, that it is polyhedral and conic. The necessity of being conic stems from the fact that in this case, the support function $\sup_{x \in X} -(\mathcal{A}\nu)^T x$ evaluates to 0 whenever its value is finite. This ensures that X-AGE-like exponentials with coefficients in the dual X-SAGE-cone are contained in the primal. The restriction of X being polyhedral enables us to use the result on the finiteness of the number of X-circuits – and, hence, allows us to describe the X-SAGE-cone in terms of power-cone-inequalities. It also ensures that the resulting optimization program is linear: If X was not polyhedral, one might not be able to certify containment in the set X by solving linear constraints.

We also provide a second-order-cone representation for the conditional SAGE-cone and its dual in the case that X is polyhedral and $\mathcal{A}^T X$ is rational, see Section 7.3. The latter already was a restriction in the unconstrained case. We need the former to again exploit the finiteness of the set of X-circuits and the fact that in this situation, the X-SAGE-cone is power-cone representable.

Chapter 2

Preliminaries

In this Chapter, we collect the background and state of the art.

We start by examining convex optimization and the duality theory of convex optimization problems in Subsection 2.1.1, with a particular emphasis on the Lagrange function, see Definition 2.1.7.

After that, we take a short detour and look at the basics of group theory, see Section 2.2. We need to do this, as part of this thesis deals with the representation and nonnegativity of *symmetric* exponential sums.

Section 2.3 focuses on polynomial optimization and nonnegativity. We start by examining how these two aspects are related, namely, we show how to transform a polynomial optimization problem into a problem of polynomial nonnegativity, (2.3). This procedure not only works for polynomials but for general functions on $\mathbb{R}^n$. We examine the nonnegativity cone (2.4), i.e., the cone of all nonnegative polynomials. For the univariate case, there exist several efficient approaches for certifying membership in this cone, but in general, finding the optimal value of a given polynomial is NP-hard.

Hence, we turn to a prominent approximation method in Subsection 2.3.1 — sums of squares (SOS). This method is based on the fact that every sum of squares of polynomials is trivially nonnegative. It does, however, not cover the whole nonnegativity cone. A prominent counterexample is the *Motzkin* polynomial. This polynomial is also interesting, as its exponential version provides a member of the class this thesis is interested in — the cone of sums of arithmetic-geometric exponentials.

Before turning to examining the state of the art of this particular cone, we take a short look at the duality theory for the cone of sums of squares, see Theorem 2.3.4, constraint optimization approaches using sums of squares, see Theorems 2.3.5 and 2.3.6 as well as Proposition 2.3.7, and the existing results for symmetric polynomials, see Theorems 2.3.8 and 2.3.9.

In the last part of this chapter, Section 2.4, we finally introduce sums of nonnegative circuit polynomials (SONC) and sums of AGE exponentials (SAGE). The former were first introduced by Iliman and de Wolff [IW16a], extending results from Reznick [Rez89]. The latter were originally introduced by Chandrasekaran and Shah [CS16]. In 2018, Wang and Murray, Chandrasekaran, and Wierman independently managed to prove the de facto equivalence of these two classes of functions [Wan18a; MCW21a] with respect to the given class of polynomials or exponential sums.

Both classes build upon sparse nonnegativity certificates. The underlying motivation for the cone of sums of nonnegative circuit polynomials was the concept of so-called *circuits*, which is closely related to the concept of circuits in matroid theory. The motivation for introducing — and naming — the cone of sums of arithmetic-geometric exponentials stems from the weighted *AM/GM inequality.* This inequality proves nonnegativity of an element in the SAGE-cone.

Besides containing elements that are nonnegative but not contained in the SOS-cone, the SONC and SAGE-cones have the huge advantage that membership in any of those cones can be certified by efficient convex optimization methods: Relative entropy programming [CS16] and for sub classes also geometric programming [DIW17]. We introduce the relative entropy formulation in Theorem 2.4.2, and in Theorem 2.4.5, we introduce the so-called *circuit number*, which yields a nonnegativity certificate for SONC polynomials and provides the theory for the geometric programs. In Subsection 2.4.1, we have a look at the relation of the SONC and SAGE-cone, in Subsection 2.4.2, we give an overview over decomposition results, and in Subsection 2.4.3, we examine constrained optimization, i.e., optimization over some subset $X \subsetneq \mathbb{R}^n$.

2.1 Notation and Convexity Theory

Throughout this article, we denote by $\mathbb{N}$ the set $\{0, 1, 2, 3, \ldots\}$; for $m \in \mathbb{N}$ let $[m]$ abbreviate the set $\{1, \ldots, m\}$. Moreover, let $\mathbf{1}$ denote the all-one vector, $\mathbf{0}$ the all-zero vector and $\delta^{(i)}$ the i-th unit vector in $\mathbb{R}^n$ for some $n \in \mathbb{N}$ and $i \leq n$.

We write $\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$ and $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Moreover, for $\mathcal{A} \subseteq \mathbb{R}^n$, denote by $\mathbb{R}^{\mathcal{A}}$ the set of $|\mathcal{A}|$-dimensional vectors whose components are indexed by the set $\mathcal{A}$, and for $\beta \in \mathcal{A}$ and $\nu \in \mathbb{R}^{\mathcal{A}}$ denote by $\nu_{\setminus\beta}$ the vector we obtain by deleting the β-th entry from ν. For $\nu \in \mathbb{R}^{\mathcal{A}}$ we denote its *support* by $\operatorname{supp}(\nu) := \{\alpha \in \mathcal{A} : \nu_\alpha \neq 0\}$, and we sometimes use $\mathcal{A}$ as a linear operator $\mathcal{A} : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}^n, \nu \mapsto \mathcal{A}\nu = \sum_{\alpha\in\mathcal{A}} \alpha\nu_\alpha$. We denote by $\langle \cdot, \cdot \rangle$ the standard Euclidean inner product in $\mathbb{R}^n$.

Given a set $\mathcal{A} \subseteq \mathbb{R}^n$ we denote by $\operatorname{relint}(\mathcal{A})$ its relative interior, by $\operatorname{pos}(\mathcal{A})$ its conic hull and by $\operatorname{conv}(\mathcal{A})$ its convex hull. We refer to the vertices of $\operatorname{conv}(\mathcal{A})$ as $\operatorname{Vert}(\operatorname{conv}(\mathcal{A}))$.

For a given linear space L, we denote by L^* its dual, for a given cone C we denote by C^* its dual cone, and for a set P by $\operatorname{rec}(P)^\circ$ its polar. The *extreme rays* of a cone C are rays that cannot be written as a sum of other rays in C, and *edge generators* are those elements inducing an extreme ray. A convex cone C is *pointed* if it contains no lines. For two vector spaces V and W and a linear map $L : V \to W$, we denote by $\ker L := \{v \in V : L(v) = 0\}$ the kernel of L, and we denote the *trace* of an $n \times n$ square matrix M by $\operatorname{tr}(M) := \sum_{i\leq n} M_{i,i}$, where $M_{i,i}$ denotes the entry of M with index (i, i).

A subset $F \subseteq S$ of some convex set S is called a *face*, denoted by $F \trianglelefteq S$, if for every $x \in F$ and $y, z \in S$ with $x \in [y, z] := \{\lambda y + (1 - \lambda)z : \lambda \in [0, 1]\}$, we have $y, z \in F$. The *recession cone* of a set S is $\operatorname{rec}(S) := \{t : \exists s \in S : s + \lambda t \in S \text{ for all } \lambda \geq 0\}$. All logarithms are base e, where e is Euler's number, and for the logarithmic function, we use the conventions $0 \ln(\frac{0}{y}) = 0$, $\ln(\frac{y}{0}) = \infty$ if $y > 0$, $\ln(\frac{0}{0}) = 0$, $\ln(0) = -\infty$ and in addition $\frac{0}{0} = 1$ and $(\frac{y}{0})^0 = 1$ for $y \in \mathbb{R}$.

For any linear operator $F : \mathbb{R}^k \to \mathbb{R}^l$ with $k, l \in \mathbb{N}$, the corresponding adjoint operator is $F^\# : \mathbb{R}^l \to \mathbb{R}^k$ with

$$\langle Fx, y\rangle = \langle x, F^\# y\rangle.$$

2.1.1 Convex Optimization

In this subsection, we collect the basic convex optimization principles starting with the easiest case of linear optimization and ending with relative entropy programming and geometric programming, both of which are relevant in the context of sparse polynomial and exponential optimization.

We put a particular focus on the duality theory of convex optimization and examine when the primal and dual optimal values coincide. In the end, we cover interior point methods. Those are known efficient algorithms to solve convex optimization problems relying on so-called barrier functions. In our context, those are of particular interest as both relative entropy programs and geometric programs have computationally tractable barrier functions and, hence, can be efficiently solved using interior point algorithms.

Let $S \subseteq \mathbb{R}^n$. A function $f : S \to \mathbb{R}$ is called *convex* on S if for all $x, y \in S$ and for all $\lambda \in [0, 1]$, we have

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Convex functions provide a class of functions where the study of extreme points is particularly easy thanks to the following theorem.

Theorem 2.1.1 (see, e.g., [BV04, p. 95])**.** *Let $S \subseteq \mathbb{R}^n$ be open and convex, and $f : S \to \mathbb{R}$ differentiable and convex. A point $y \in S$ minimizes $f(x)$ if and only if $\nabla f(y) = 0$.*

As a direct consequence of convexity, we obtain *Jensen's inequality*, which we use frequently in the course of this thesis.

Theorem 2.1.2 (Jensen's Inequality, see, e.g., [BV04, p. 77])**.** *For $S \subseteq \mathbb{R}^n$ and a real convex function $f : S \to \mathbb{R}$, $x^{(1)}, \ldots, x^{(m)} \in S$, and positive weights $(a_i)_{i \in [m]}$, we have*

$$f\left(\frac{\sum_i a_i x^{(i)}}{\sum_i a_i}\right) \leq \frac{\sum_i a_i f(x^{(i)})}{\sum_i a_i}.$$

Equality holds if and only if f is a linear function on S or $x^{(i)} = x^{(j)}$ for all $i, j \in [m]$.

Having collected these basic statements, we introduce convex optimization problems as the main subject of study in this subsection. A *convex optimization problem* is a problem of the form

$$\inf\{f(x) : x \in C\},$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and $C \subseteq \mathbb{R}^n$ a convex set.

The simplest example of a convex optimization problem is a linear program, where f is a linear function and C a polyhedron. Linear programs have a long history. Orignally, linear programming was introduced by Fourier (1827) and Kantorowitsch (1939), while 1947 Dantzig developed the famous simplex method as an efficient way to solve linear programs. In 1979 the first method for solving linear programs in polynomial time was proposed by Khachiyan — the ellipsoid method — and in 1984 Karmakar provided an interior point algorithm to solve linear programs.

Generalizing linear programs leads to another prominent convex optimization problem involving matrix inequalities, namely *semidefinite programming*.

An $n \times n$ symmetric real matrix M is *positive definite* if $x^T M x > 0$ for all $x \in \mathbb{R}^n \setminus \mathbf{0}$, and *positive semidefinite* if $x^T M x \geq 0$ for all $x \in \mathbb{R}^n$. We denote this by $M \succcurlyeq 0$, and we denote by $\mathcal{S}^n$ the space of all real symmetric $n \times n$ matrices together with the inner product

$$\langle A, B\rangle = \mathrm{tr}(A^T B) = \mathrm{tr}(A \cdot B).$$

Definition 2.1.3 (SDP)**.** A *semidefinite program (SDP)* is an optimization problem of the form

$$\begin{aligned}\inf_{X \in \mathbb{R}^{n\times n}} & \langle C, X\rangle \\ & \langle A_i, X\rangle = b_i \text{ for all } i \in [m] \\ & X \succcurlyeq 0,\end{aligned}$$

where $A_i \in \mathcal{S}_n$, $b_i \in \mathbb{R}^n$ and $m \in \mathbb{N}$.

In [Ave19], Averkov defined the *semidefinite extension degree*, which has a close relation to semidefinite programming. The definition is in terms of *linear matrix inequalities* of size k for some $k \in \mathbb{N}$ and $n \in \mathbb{N}$:

$$A_0 + x_1 A_1 + \ldots + x_n A_n \succcurlyeq 0$$

with symmetric $k \times k$ matrices A_i, $0 \leq i \leq n$, and $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$.

Definition 2.1.4 (Semidefinite Extension Degree [Ave19])**.** Let $S \subseteq \mathbb{R}^n$. The minimal $k \in \mathbb{N}$ such that S can be described by a linear matrix inequality of size k is called the *semidefinite extension complexity* of S (and is set to ∞ if such a k does not exist).

The *semidefinite extension degree* of S is defined as the smallest $k \in \mathbb{N}$ such that S can be described by m linear matrix inequalities of size k for some finite $m \in \mathbb{N}$ (and as ∞ if S has no semidefinite extended formulation).

Note that any set which can be represented as a projection of the feasible set of an SDP has finite semidefinite extension degree. A related class of optimization programs are programs optimizing over sets with semidefinite extension degree 2, namely *second-order representable sets*:

Definition 2.1.5 (Second-Order-Cone Programs (SOCP))**.** For the Euclidean norm $\|\cdot\|$, a *second-order-cone program (SOCP)* is an optimization problem of the form

$$\min\left\{c^T x \,:\, ||A_i x + b_i||_2 \leq c_i^T b + d_i \text{ for all } i \in [m]\right\}$$

with real symmetric matrices A_i, vectors b_i, c_i, d_i and a vector c. A subset of $\mathbb{R}^n$ is called *second-order representable* if it can be represented as a projection of the feasible set of a second-order-cone program.

Second-order-cone programs are related to semidefinite programs since for a symmetric 2×2-matrix, positive semidefiniteness can be formulated as a second-order condition:

Lemma 2.1.6. (See, e.g., [NN94, §6.4.3.8], [WM20a, Lemma 4.3].) *A symmetric 2×2 matrix $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is positive semidefinite if and only if the second-order condition*

$$\left\| \begin{pmatrix} 2b \\ a - c \end{pmatrix} \right\|_2 \leq a + c$$

is satisfied.

Before we examine other convex optimization problems that are particularly useful in the optimization of sparse polynomials and exponential sums, we take a detour to duality theory.

To do so, we examine convex optimization problems of the form

$$\begin{aligned} &\inf f_0(x) \\ &\quad f_j(x) \le 0 \text{ for all } j \in [m], x \in \mathbb{R}^n \end{aligned} \tag{2.1}$$

where f_j are convex functions for all $j \in \{0, \dots, m\}$.

We start by introducing the *Lagrange function*, which reduces optimization of a convex function over a convex constraint set to optimization over a single convex function.

Definition 2.1.7 (Lagrange Function)**.** The *Lagrange function* of the problem (2.1) is the function $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ defined as

$$L(x, \mu) = f_0(x) + \sum_{j \in [m]} \mu_j f_j(x).$$

Following [BV04, Section 5], the infimum

$$\inf_{x \in \mathbb{R}^n} \sup_{\mu \in \mathbb{R}^m_+} L(x, \mu)$$

is finite on the set of feasible $x \in \mathbb{R}^n$ for (2.1) and coincides with the optimal value of (2.1).

The *Lagrange dual problem* to (2.1) is the problem

$$\sup_{\mu \in \mathbb{R}^m_+} \inf_{x \in \mathbb{R}^n} L(x, \mu). \tag{2.2}$$

Whenever we consider linear programs, i.e., for all $j \in \{0, \dots, m\}$, f_j is a linear function, the optimal values of the primal and dual problems coincide as soon as both are finite and there exists a feasible point. For non-linear functions, we can only guarantee some weak duality argument.

Theorem 2.1.8 (Weak Duality for Convex Optimization, see, e.g., [BV04, p. 225])**.** *If x is a feasible solution of* (2.1) *and μ is a feasible solution of* (2.2)*, then*

$$\sup_{\mu \in \mathbb{R}^m_+} \inf_{x \in \mathbb{R}^n} L(x, \mu) \le \inf_{x \in \mathbb{R}^n} \sup_{\mu \in \mathbb{R}^m_+} L(x, \mu).$$

There are, however, criteria guaranteeing that there is no duality gap — the *Slater Conditions.*

Theorem 2.1.9 (Strong Duality for Convex Optimization [Sla14])**.** *For some $m \in \mathbb{N}$ and all $i \in \{0, \dots, m\}$ let $f_i : \mathbb{R}^n \to \mathbb{R}$ be convex and assume that the* Slater constraint qualification

$$\textit{there exists } \hat{x} \in \mathbb{R}^n \textit{ with } f_i(\hat{x}) < 0 \textit{ for all } i \in [m]$$

is satisfied. Then,

$$\sup_{\mu \in \mathbb{R}^m_+} \inf_{x \in \mathbb{R}^n} L(x, \mu) = \inf_{x \in \mathbb{R}^n} \sup_{\mu \in \mathbb{R}^m_+} L(x, \mu).$$

So instead of the precondition of a feasible point — which is sufficient in the case of linear constraints –, we need a *strictly feasible* point to guarantee coincidence of primal and dual optimal values.

The Fenchel conjugate (or convex conjugate / Legendre conjugate) captures the special case of the Lagrange function with linear constraint sets (see, e.g., [BV04, Section 3.3])

Definition 2.1.10 (Fenchel Conjugate)**.** The *Fenchel conjugate* of a given function $f : \mathbb{R}^n \to \mathbb{R}$ is the function f^* defined as

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (y^T x - f(x))$$

where only the domain set $\text{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}$ is considered. The function f^* is convex on its domain even if f itself is not convex.

To see the connection to the Lagrange function, consider the problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad Ax \leq b,\ Cx = d$$

for matrizes $A \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{l \times n}$ and $m, l \in \mathbb{N}$. The Lagrange dual optimization problem then is

$$\begin{aligned} \sup_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^l_+} L(\lambda, \mu) &= \sup_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^l_+} \inf_{x \in \mathbb{R}^n} f(x) + \lambda^T (Cx - d) + \mu^T (Ax - b) \\ &= \sup_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^l_+} - d^T \lambda - b^T \mu - f^*(-C^T \lambda - A^T \mu). \end{aligned}$$

We use $\text{dom}(L) = \{(\lambda, \mu) \in \mathbb{R}^l \times \mathbb{R}^m_+ : -C^T \lambda - A^T \mu \in \text{dom}(f^*)\}$, as the Lagrange dual has an extended domain.

Having established the duality theory, we introduce the concept of *primal-dual interior point algorithms*, which show that the optimization techniques used later for polynomial and exponential optimization are indeed efficiently solvable. Interior point methods are algorithms solving convex optimization problems. The idea is to transform any convex optimization problem into a problem where a linear function is optimized subject to convex constraints, using a *barrier function*.

Definition 2.1.11 (Barrier Function)**.** Let $C \subseteq \mathbb{R}^n$ be convex, closed and such that $\text{int}\, C \neq \emptyset$. A continuous function $F : \mathbb{R}^n \to \mathbb{R}$ is called *barrier function* for C if $\text{dom}\, F = \text{int}\, C$ and $F(x) \to \infty$ for $x \to \partial F$.

For an optimization problem of the form (2.1), one usually uses the logarithmic barrier function $-\sum_{j \in [m]} \ln(f_j(x))$.

There exist a lot of solvers for convex optimization programs, especially for the special case of linear programming. For the computational experiments in Chapter 4, we want to point out the solvers *MOSEK* [DA21] and *ECOS* [DCB13], which can both be used with the *CVXPY* package [DB16; Agr+18; ADB19].

As mentioned before, for polynomial or exponential optimization, i.e., for the optimization of a given polynomial $f = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$ or exponential sum $f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle}$ with $c_\alpha \in \mathbb{R}$, $\mathcal{A} \subseteq \mathbb{R}^n$ and $x^\alpha = \prod_{i \in [n]} x_i^{\alpha_i}$, the classes of geometric and relative entropy programs are of particular interest.

Definition 2.1.12 (Geometric Program)**.** A *geometric program (GP)* is an optimization problem of the form

$$\begin{aligned}\min\, & p_0(x),\\ & p_i(x) \leq 1 \text{ for all } i \in [m],\\ & q_j(x) = 1 \text{ for all } j \in [r], x \in \mathbb{R}^n_{>0}\end{aligned}$$

where q_j are *monomials*, i.e., $q_j = d_j x^{\alpha^{(j)}}$ with $d_j > 0, \alpha^{(j)} \in \mathbb{R}^n$ for all $j \in [r]$ and p_i are *posynomials*, i.e., sums of monomials $p_i = \sum_{l \in [k_i]} c_l^{(i)} x^{\alpha^{(i_l)}}$ with $c_l^{(i)} > 0$ for all $l \in [k_i]$ and $k_i \in \mathbb{N}$ for all $i \in \{0, \ldots, m\}$.

In the form defined above, however, geometric programs are not convex in general. Using the substitution $y_i = \ln(x_i)$, we can transform the program into its *convex version*:

$$\begin{aligned}\min\, & \sum_{l \in [k_0]} c_l^{(0)} e^{\langle \alpha^{(i_l)}, y\rangle},\\ & \sum_{l \in [k_i]} c_l^{(i)} e^{\langle \alpha^{(i_l)}, y\rangle} \leq 1 \text{ for all } i \in [m],\\ & d_j e^{\langle \alpha^{(j)}, y\rangle} = 1 \text{ for all } j \in [r].\end{aligned}$$

Since

$$-\sum_{i \in [m]} \ln\left(-\ln\left(\sum_{l \in [k_i]} c_l^{(i)} e^{\langle y, \alpha^{(i_l)}\rangle}\right)\right)$$

is a barrier function for the non-linear constraints

$$-\ln\left(\sum_{l \in [k_i]} c_l^{(i)} e^{\langle y, \alpha^{(i_l)}\rangle}\right) \geq 0,$$

we can solve geometric programs using interior point methods. Geometric programs have the advantage that — even in high dimension and with many variables — they can be solved efficiently and remain stable [Boy+07].

We conclude this section by looking at *relative entropy programming*.

Definition 2.1.13 (Relative Entropy Program)**.** Relative entropy programs (REPs) are conic optimization problems in which a linear function under conic constraints specified by the *relative entropy cone* and linear constraints is minimized:

$$\begin{aligned}\min\, & f_0(x),\\ & f_i(x) \geq 0 \text{ for } i \in [m],\\ & x \in \mathcal{RE}_n,\end{aligned}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$ are linear functions for all $i \in \{0, \ldots, m\}$ and the relative entropy cone is the convex cone

$$\mathcal{RE}_n = \{(\nu, \lambda, \delta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n : \nu_i \ln(\nu_i/\lambda_i) \leq \delta_i \text{ for all } i \in [n]\}.$$

The analysis in [CS17] reveals that optimizing over the relative entropy function $D : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ with $(\nu, \lambda) \mapsto \sum_{i \leq n} \nu_i \ln(\nu_i/\lambda_i)$ can be done by a relative entropy program:

$$D(\nu, \lambda) \leq t \Leftrightarrow \text{ there exists } \delta \in \mathbb{R}^n : (\nu, \lambda, \delta) \in \mathcal{RE}_n \text{ and } \mathbf{1}^T \delta = t.$$

Relative entropy programs naturally are convex, and, moreover, there exist efficiently computable barrier functions [NN94]. Hence, they can be solved via interior point methods. Note that both SOCPs as well as GPs are special cases of relative entropy programs [Boy+07].

2.2 Group Theory

For the results in Chapter 4, we collect the basics of group theory, compare [Bos20].

We start by reminding the reader of the definition of two prominent examples of finite groups on $\mathbb{R}^n$. Let $n \in \mathbb{N}$ with $n \geq 1$. The *Symmetric Group* $\mathcal{S}_n$ is the set of all permutations of the set $[n]$, having the composition of permutations as the group operation. The *General Linear Group* $\mathrm{GL}_n(\mathbb{R})$ consists of all invertible $n \times n$-matrices with real entries, together with matrix multiplication.

In the following chapters, we are also interested in how a group acts on a set. Thus, we define the *left and right group action*.

Definition 2.2.1 (Group Action on a Set)**.** Let G be a group with identity element e and let $\mathcal{A}$ be a set.

1. A *left action* σ of G on $\mathcal{A}$ is a function

$$\sigma : G \times \mathcal{A} \to \mathcal{A},$$

such that $e \cdot \alpha = \alpha$ and $g \cdot (h \cdot \alpha) = (gh) \cdot \alpha$ for all $\alpha \in \mathcal{A}$ and $g, h \in G$. We say, the group G acts on $\mathcal{A}$ from the left.

2. Similarly, a *right action* σ of G on $\mathcal{A}$ is a function

$$\sigma : \mathcal{A} \times G \to \mathcal{A},$$

such that $\alpha \cdot e = \alpha$ and $(\alpha \cdot g) \cdot h = \alpha \cdot (gh)$ for all $\alpha \in \mathcal{A}$ and $g, h \in G$. We say, the group G acts on $\mathcal{A}$ from the right.

Note that a subgroup S of a group G acts on G via translation from the left, as follows

$$S \times G \to G, \ (s, g) \mapsto sg$$

and via translation from the right as follows

$$G \times S \to G, \ (g, s) \mapsto gs.$$

Having this definition, we can now define left and right cosets as well as orbits.

Definition 2.2.2 (Cosets and Orbits)**.** Let G be a group and S be a subgroup of G acting on G via a *right action* $G \times H \to G$, $(g, h) \mapsto g \cdot h$ or a *left action* $H \times G \to G$, $(h, g) \mapsto h \cdot g$. Then, for $g \in G$, a *left coset* is the *orbit* of g under a right action, i.e., a set of the form

$$gS = \{h \in G : \text{ there exists } s \in S \text{ with } h = g \cdot s\},$$

and a *right coset* is the *orbit* of g under a left action, i.e., a set of the form

$$Sg = \{h \in G : \text{ there exists } s \in S \text{ with } h = s \cdot g\}.$$

For subgroups S and H of G, we denote by G/S the set of left cosets and by $S\backslash G$ the set of right cosets. Moreover, we say $S\backslash G/H$ for the set of *double cosets*

$$S\backslash G/H = \{SgH : g \in G\}.$$

Over the course of this thesis, we solely use groups acting on $\mathbb{R}^n$ from the left. Hence, we assume from here on that any group G acts from the left on any subset $\mathcal{A} \subseteq \mathbb{R}^n$.

Definition 2.2.3. Let G be a group acting on a set $\mathcal{A}$ from the left.

1. For every $\alpha \in \mathcal{A}$, the *orbit* of α is the set

$$G \cdot \alpha = \{g \cdot \alpha : g \in G\}.$$

   We call a set $\hat{\mathcal{A}} \subseteq \mathcal{A}$ a *set of orbit representatives* if it is inclusion-minimal and with the property that $G \cdot \hat{\mathcal{A}} = \mathcal{A}$.

2. For every $\beta \in \mathcal{A}$, the *stabilizer* of β is the set

$$\operatorname{Stab} \beta = \{g \in G : g \cdot \beta = \beta\}.$$

3. The set of all orbits of $\mathcal{A}$ under the action of G

$$\mathcal{A}/G := \{G \cdot \alpha : \alpha \in \mathcal{A}\}.$$

   is called *orbit space*.

A subset $\mathcal{B} \subseteq \mathcal{A}$ is called *G-invariant* if

$$G \cdot \mathcal{B} = \mathcal{B}.$$

The same holds for elements: An element $\alpha \in \mathcal{A}$ is called *G-invariant* if $g \cdot \alpha = \alpha$ for all $g \in G$.

A prominent result here is the *Orbit Counting Theorem* (sometimes also called Burnside's Lemma or Cauchy-Frobenius Lemma):

Lemma 2.2.4 (see, e.g., [Sta99, Lemma 7.24.5])**.** *Let G be a finite group acting on a set $\mathcal{A}$. Then,*

$$|\mathcal{A}/G| = \frac{1}{|G|} \sum_{\alpha \in \mathcal{A}} |\operatorname{Stab} \alpha|.$$

For double cosets, we can use a variation: Assume S, H are subgroups of G. Then,

$$|S\backslash G/H| = \frac{1}{|S||H|} \sum_{(s,h) \in S \times H} |\{g \in G : s \cdot g \cdot h = g\}|.$$

We close this subsection by stating the definition of a *linear representation* ρ of a finite group G over a field $\mathbb{F}$, as well as its dual. We assume that $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ in general. A *linear representation* of a finite group G over a field $\mathbb{F}$ is a homomorphism $\rho : G \to \mathrm{GL}(V)$, where V is a vector space over $\mathbb{F}$ and $\mathrm{GL}(V)$ is the general linear group of V. Given a basis of V, its representation can be expressed as a group homomorphism into the group $\mathrm{GL}_n(\mathbb{F})$, where $n := \dim(V)$; this is a matrix representation. Its *dual representation* is the linear representation ρ^* of the dual space V^* defined via

$$\rho^*(g) = \rho(g^{-1})^T \text{ for all } g \in G.$$

2.3 Polynomial and Exponential Optimization and Nonnegativity

As already mentioned in the introduction, finding the minimal value of a given real function — either over $\mathbb{R}^n$ or any subset of it — is a fundamental question in real algebraic geometry. Framing this question in terms of the subclass of polynomials or exponential sums has various applications, for example in control theory and robotics [HG05; AM19].

For $x \in \mathbb{R}^n$, a polynomial is a function

$$f(x) = \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} \prod_{j \leq n} x_j^{i_j}$$

with coefficients $c_{i_1,\ldots,i_n} \in \mathbb{R}$ for $n \in \mathbb{N}$ and $i_j \in \mathbb{N}$ for all $j \leq n$. For a fixed exponent vector $\alpha \in \mathbb{N}^n$, we denote

$$x^\alpha = \prod_{i \in [n]} x_i^{\alpha_i}.$$

We denote by $\mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]$ the space of polynomials in n variables with real coefficients . A *sparse polynomial* with support in $\mathcal{A} \subseteq \mathbb{N}^n$ is a function

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$$

and, for $\mathcal{A} \subseteq \mathbb{R}^n$, a *sparse exponential sum* (sometimes also called signomial) is a function

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle x, \alpha \rangle}.$$

In a polynomial or exponential optimization problem, we optimize a polynomial or exponential sum f, respectively, given some set of constraints, here referred to as X. In the case that $X \subsetneq \mathbb{R}^n$, we speak of a *constrained* polynomial or exponential optimization problem, and when $X = \mathbb{R}^n$, we speak of a *global* polynomial or exponential optimization problem.

Optimization and nonnegativity of a given real function are tightly related. Consider for a polynomial or exponential sum f the following problem:

$$\min_{x \in X} f(x)$$

This problem is equivalent to finding the maximal real number that we can subtract from f to obtain a nonnegative function:

$$\sup_{\lambda \in \mathbb{R}} \lambda \text{ such that } f(x) - \lambda \geq 0 \text{ for all } x \in X. \tag{2.3}$$

This way, we transfer an optimization problem into a problem of nonnegativity. Hence, it is natural to examine the cone of *nonnegative polynomials* in n variables, defined as the set

$$\mathcal{P} := \{f \in \mathbb{R}[x_1, \ldots, x_n] : f(x) \geq 0 \text{ for all } x \in \mathbb{R}^n\}. \tag{2.4}$$

For a given subset $X \subseteq \mathbb{R}^n$, we can analogously define the cone of *nonnegative polynomials over* X:

$$\mathcal{P}_X := \{f \in \mathbb{R}[x_1, \ldots, x_n] : f(x) \geq 0 \text{ for all } x \in X\}.$$

More generally, in this thesis, we consider the cone of nonnegative generalized polynomial (or signomial) functions; compare Chapter 5. We also use the notations $\mathcal{P}$ and $\mathcal{P}_X$ for the analogous cones of (X-)nonnegative exponential sums.

At least for the univariate case, there exist algorithms providing nonnegativity certificates for polynomials, like Sturm sequences, which can be used to compute the number of roots of a given polynomial [BPR06]. But for the multivariate case, even computing whether a given polynomial with degree at least 4 is globally nonnegative was proven to be NP-hard [DG08]. Hence, the idea is to find *subclasses* of the class of nonnegative polynomials which approximate the question. A prominent subclass is the cone of *sums of squares of polynomials*, which we examine in the next subsection. Before doing so, we introduce the basic notion and the duality theory of polynomials and exponential sums.

For $\mathcal{A} \subseteq \mathbb{R}^n$ we consider the set $\mathbb{R}^{\mathcal{A}}$ of vectors with components indexed by $\alpha \in \mathcal{A}$, which allows us to identify a polynomial or exponential sum with the vector of coefficients $c \in \mathbb{R}^{\mathcal{A}}$; hence, we sometimes write $f \in \mathbb{R}^{\mathcal{A}}$ for polynomials or exponential sums f with coefficient vectors $c \in \mathbb{R}^{\mathcal{A}}$. The *support* of a polynomial $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$ or exponential sum $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle x, \alpha \rangle}$ is $\operatorname{supp}(f) = \{\alpha \in \mathcal{A} : c_\alpha \neq 0\}$ and the *degree* of f is $\deg(f) = \max_{\alpha \in \operatorname{supp}(f)} |\alpha|$.

Duality is a strong aspect in this thesis. In Chapters 3, 4 and 5, we examine — among others — the dual cones of a subcone of the nonnegativity cone $\mathcal{P}$, and in Chapters 6 and 7, we examine — again, among others — the dual cone of a subcone of the *constrained* nonnegativity cone $\mathcal{P}_X$. Hence, it is natural to establish the duality theory for polynomials.

Definition 2.3.1 (Natural Duality Pairing)**.** Given a polynomial

$$p = \sum_{\alpha \in \operatorname{supp}(p)} c_\alpha x^\alpha \in \mathbb{R}[x_1, \ldots, x_n]$$

with coefficients $c_\alpha \in \mathbb{R} \setminus \{0\}$ for all $\alpha \in \operatorname{supp}(p)$ and an element v in the dual space, the *natural duality pairing* is

$$v(p) = \sum_{\alpha \in \operatorname{supp}(p)} c_\alpha v_\alpha.$$

The equivalent holds for exponential sums $\sum_\alpha c_\alpha e^{\langle x, \alpha \rangle}$ or polynomial-like functions.

Then, the dual cone of nonnegative polynomials is the set

$$\mathcal{P}^* = \{v : v(p) \geq 0 \text{ for all } p \in \mathcal{P}\}.$$

Again, the equivalent holds for exponential sums, polynomial-like functions, and the constrained situation, i.e., $\mathcal{P}_X$.

2.3.1 Sums of Squares

We finally turn to examining the subclass of sums of squares.

The SOS-cone $\Sigma[x]$ for $x \in \mathbb{R}^n$ is the cone of *sums of squares* of polynomials in x. Naturally, this is a subcone of the cone of nonnegative polynomials, and in the univariate case the converse also holds, i.e., every univariate nonnegative polynomial of even degree is a sum of squares.

In 1888, Hilbert showed that the cones of nonnegative polynomials and sums of squares coincide in exactly three cases:

Theorem 2.3.2 (Hilbert)**.** *Let $n \geq 2$ and d be even. The cones $\Sigma_{n,d}$ of sums of squares and $\mathcal{P}_{n.d}$ of nonnegative polynomials, both in n variables and of degree d, coincide if and only if a) $n = 2$, b) $d = 2$, or c) $n = 3$ and $d = 4$.*

However, as the cone of sums of squares is a subcone of the nonnegativity cone in *any* case, considering the (global or constrained) polynomial optimization problem introduced previously, a natural relaxation for this problem is the following:

$$\begin{aligned} p^{\text{sos}} := \min\ & \lambda \\ \text{s.t. } & p - \lambda \in \Sigma[x]. \end{aligned} \tag{2.5}$$

This problem provides a lower bound to the original problem: Let λ^* be the optimal value of (2.3) and λ^{sos} the optimal value of (2.5). Since the feasible set of the second problem is a subset of the feasible set of the first problem, we clearly have $\lambda^* \geq \lambda^{\text{sos}}$.

The decomposition of a given polynomial into a sum of squares can be computed by a semidefinite program:

Theorem 2.3.3 (see [Las15, Proposition 2.1])**.** *Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be of even degree $2d$ and let Y be the vector of monomials in $x_1, \ldots, x_n$ of degree at most d. Then, $p \in \Sigma[x_1, \ldots, x_n]$ if and only if there exists $Q \in \mathcal{S}_+^{|Y|}$ with*

$$p = Y^T Q Y.$$

The proof of this theorem relies on the fact that a polynomial p of degree $2d$ is a sum of squares $p = \sum (s_j(x))^2$ with polynomials $s_j(x)$ of degree at most d if and only if the polynomial's coefficient vectors s_j and the vector Y of all monomials in n variables of degree at most d satisfy

$$p = Y^T (\sum s_j s_j^T) Y,$$

i.e., if and only if the matrix $\sum s_j s_j^T$ is positive semidefinite.

As explained above, in general, this constructed SDP leads to a lower bound. To examine when this bound is *sharp*, i.e., when $p^* = p^{\text{sos}}$, we take a look at the *dual* of this semidefinite program, building upon the duality theory for polynomials established in the previous subsection. In particular, this provides a method to capture any duality gap when using the SOS relaxation.

We consider a convexified version of the global optimization problem. Let $\mathcal{P}(\mathbb{R}^n)$ denote the set of probability measures on $\mathbb{R}^n$. In case the minimum is finite, we have the equality

$$p^* = \min_{x \in \mathbb{R}^n} p(x) = \min_{\mu \in \mathcal{P}(\mathbb{R}^n)} \int_{\mathbb{R}^n} p(x) d\mu(x) = \sum_\alpha p_\alpha y_\alpha$$

for $p = \sum_\alpha p_\alpha y_\alpha$ and where $y = (y_\alpha)_\alpha$ is the *moment sequence*

$$y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu.$$

Given a moment sequence y, the *moment problem* is the question whether there exists some representing measure μ on $\mathbb{R}^n$, i.e., some measure μ with $y_\alpha = \int x^\alpha d\mu$. For a real subset $X \subsetneq \mathbb{R}^n$, the X*-moment problem* is the analogous question of the existence of some representing measure on X; we discuss the constrained case in detail at the end of this subsection.

We can interpret the moment sequence y as a *linear functional* $L : \mathbb{R}[x] \to \mathbb{R}$ being the integration on $\mathbb{R}^n$ with respect to μ, i.e., with $x^\alpha \mapsto y_\alpha = \int x^\alpha d\mu$. Moreover, we often work with a bilinear form $\mathcal{L} : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$ with $(p, q) \mapsto L(p \cdot q)$ — the *associated moment form*.

For a given moment sequence, the *moment matrix* is the matrix

$$M(y) = (y_{\alpha+\beta})_{\alpha,\beta} \text{ for } \alpha, \beta \in \mathbb{N}^n$$

and the *truncated moment matrix* is

$$M_d(y) = (y_{\alpha+\beta})_{\alpha,\beta} \text{ for } \alpha, \beta \in \mathbb{N}^n \text{ with } |\alpha| \leq d, |\beta| \leq d,$$

i.e., where the degrees of the exponent vectors α and β are bounded by d. Observe that in particular $(M_d)_{\alpha,\beta} = \mathcal{L}(x^\alpha, x^\beta) = L(x^{\alpha+\beta})$ for all $d \geq 0$ and $\alpha, \beta \in \mathbb{N}^n$ with $|\alpha| \leq d$ and $|\beta| \leq d$. As $\mathcal{L}(p, p) = L(p^2) = \int p(x)^2 d\mu \geq 0$, whenever there exists a measure μ such that L indeed is the integration with respect to μ, we have $L(1) = 1$ and $\mathcal{L}$ is positive semidefinite as well as $M_d(y) \succcurlyeq 0$ for all $d \geq 0$.

The problem, however, is that the underlying matrices might have infinite dimension as the degree of the polynomials in the decomposition is not bounded above in general. Thus, without a *degree bound*, this problem cannot be solved efficiently. Hence, for a given polynomial $p = \sum_\alpha p_\alpha y_\alpha$ of degree $2d$, one often uses the *moment relaxation*

$$p^{\text{mom}} = \inf\{\sum_\alpha p_\alpha y_\alpha : M_d(y) \succcurlyeq 0\},$$

i.e., we use the truncated moment matrix bounded in size by d.

The following theorem collects an important statement concerning the relation of the SOS relaxation and the moment relaxation in terms of *duality*. It also shows how to extract an optimal point from an optimal solution of the SOS relaxation in case this relaxation actually provides the optimal value.

Theorem 2.3.4 (see [Las15; Lau09])**.** *Let $p \in \mathbb{R}[x]$ with $p^* = \min_{x \in \mathbb{R}^n} p(x)$ and degree $2d$ for $d \in \mathbb{N}$.*

1. *Then, $p^{sos} = p^{mom}$. If, moreover, $p^{mom} > -\infty$, then $p^* = p^{sos} = p^{mom}$. We say the SOS program and the moment relaxation are* dual *programs.*

2. *If $p - p^*$ admits an SOS decomposition, then $p^* = p^{sos} = p^{mom}$. If x^* is a minimal point for p in $\mathbb{R}^n$, then $y^* = ((x^*)^\alpha)_{|\alpha| \leq 2d}$ is a minimal point for the moment relaxation.*

Moreover, the cone $\mathcal{M} = \{y : y \text{ has a representing measure}\}$ is the dual of the cone of nonnegative polynomials $\mathcal{P}$. Also, the cone $\mathcal{M}_+ = \{y : M(y) \geq 0\}$ and the cone Σ of sums of squares are duals of each other.

We turn to examining the constrained SOS relaxation starting with Putinar's Positivstellensatz [Put93], see also [Las10, Theorem 2.14]. This Positivstellensatz provides the starting point for examining constrained optimization problems of the form

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} p(x) \\ &\qquad g_i(x) \geq 0 \text{ for } i \in [m] \end{aligned}$$

with $g_1, \ldots, g_m \in \mathbb{R}[x]$. We sometimes write $X = \{x \in \mathbb{R}^n : g_i(x) \geq 0 \text{ for all } i \leq m\}$ to denote the constraint set.

Before stating the Positivstellensatz, recall that a *quadratic module* is a subset $M \subseteq \mathbb{R}[x]$ which contains 1 and is closed both under addition and multiplication by squares.

Theorem 2.3.5 (Putinar's Positivstellensatz)**.** *Let $p, g_1, \ldots, g_m \in \mathbb{R}[x]$ and assume that there exists some $N \geq 1$ such that $N - \sum_{i\leq n} x_i^2$ is contained in the quadratic module*

$$Q(g_1, \ldots, g_m) := \left\{ p_0 + \sum_{j=1}^{m} p_j g_j \text{ with } p_0, \ldots, p_m \in \Sigma[x] \right\}.$$

If $p \in \mathbb{R}[x]$ is strictly positive on the set X, then $p = p_0 + \sum_{j=1}^{m} p_j g_j$ with sum of squares polynomials $p_0, \ldots, p_m$.

The precondition in the previous theorem is equivalent to $Q(g_1, \ldots, g_m)$ being *Archimedean*, and it is well known that $Q(g_1, \ldots, g_m)$ is Archimedean if $g_1, \ldots, g_m$ are affine, see, e.g., [Las10].

The first well-known Positivstellensatz on basic closed semi-algebraic sets goes back to Krivine [Kri64] and was then further developed by Stengle [Ste74] — hence it is sometimes called the Krivine-Stengle Positivstellensatz. They stated how to represent a given function which is positive over a set of constraints and, hence, provides a *certificate of nonnegativity* for this function. The assumptions made in Putinar's Positivstellensatz are stronger than those in the ones by Krivine and Stengle, but in contrast to the Krivine-Stengle Positivstellensatz, a solution can be computed efficiently.

As a Positivstellensatz provides a natural certificate of nonnegativity, building upon Theorem 2.3.5, we construct the following SOS relaxation for the constrained optimization problem:

$$p^{\text{sos}}(g_1, \ldots, g_m) = \sup\{\gamma \in \mathbb{R} : p - \gamma \in Q(g_1, \ldots, g_m)\}.$$

Analogously to the unconstrained situation, we can solve this relaxation via semidefinite programming. Similarly, however, the underlying matrices might have infinite dimension.

Before we consider a relaxation of the problem similar to the one in the unconstrained case, we take a brief look at the dual problem — again via *moments*. Similar to the unconstrained situation, there exists a statement by Putinar, concerning the linear functional L and the corresponding bilinear form $\mathcal{L}$.

Theorem 2.3.6 (Putinar [Put93])**.** *Suppose $Q(g_1, \ldots, g_m)$ is Archimedean. Then, a linear map $L : \mathbb{R}[x] \to \mathbb{R}$ is the integration with respect to some measure μ on K if and only if $L(1) = 1$ and $L(p_o + \sum_{j\leq m} p_j g_j) \geq 0$ for sums of squares polynomials p_j for all $j \leq m$.*

An approach for working around the possibly infinite sizes of the SDPs are *Lasserre relaxations*. For degree $2d \geq \deg(p), \deg(g_i)$ for all $i \leq m$, the *d-th Lasserre relaxation* [Las00] is the problem $p_d^{\text{sos}}(g_1, \ldots, g_m)$ defined as

$$\sup\{\gamma : p - \gamma = p_0 + \sum_{j=1}^{m} p_j g_j \text{ with } p_i \in \Sigma[x] \text{ and } \deg(p_j g_j) \leq 2d \text{ for all } j \leq m\}.$$

With $d_{g_i} = \lceil \deg(g_i)/2 \rceil$, the analogous relaxation based on moments is the problem

$$p_d^{\text{mom}}(g_1, \ldots, g_m) = \inf_y \{\sum_\alpha p_\alpha y_\alpha : M_d(y) \succcurlyeq 0, M_{d-d_{g_i}}(y) g_i \succcurlyeq 0 \text{ for all } i \leq m\}.$$

Those two problems are dual semidefinite problems which bound the optimal value $p_X^* = \min\{p(x) : x \in X\}$ below, and both problems are polynomial in their input sizes. We state a convergence result of the two relaxations under the same preconditions as for Putinar's Positivstellensatz.

Proposition 2.3.7 ([Las00])**.** *If $Q(g_1, \ldots, g_m)$ is Archimedean, then*

$$\lim_{t\to\infty} p_t^{sos} = \lim_{t\to\infty} p_t^{mom} = \min_{x\in X} p(x).$$

To conclude this section, we have a look at what happens in situations involving symmetries, i.e., we consider a polynomial invariant under the action of some group G. In 2013, Riener, Theobald, Jansson-Andrén and Lasserre provided a symmetric version of Theorem 2.3.6 stated above.

Theorem 2.3.8 ([Rie+13])**.** *For a group G, let $g_1, \ldots, g_m \in \mathbb{R}[x]$ be G-invariant, and assume the set $Q(g_1, \ldots, g_m)$ is Archimedean. Setting $g_0 := 1$, a G-linear map $L^G : \mathbb{R}[x] \to \mathbb{R}$ is the integration with respect to some G-invariant measure on K if and only if the bilinear forms*

$$\begin{aligned} &\mathcal{L}_{g_j}^G : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}, \\ &(p, q) \mapsto L^G\Big(\frac{1}{|G|} \sum_{\sigma\in G} (p \cdot q)^\sigma \cdot g_j\Big) \end{aligned}$$

are positive semidefinite for all $0 \leq j \leq m$.

In the same paper, they also provided a symmetric version of Putinar's Positivstellensatz as well as a symmetry-adapted relaxation of the moment problem. We will not examine those two theorems here in full detail as they require a lot of linear representation theory that is irrelevant for the main results of this thesis. Nevertheless, we collect the rough version of the Positivstellensatz below.

Theorem 2.3.9 ([Rie+13], rough version)**.** *Let G be a group, $f, g_1, \ldots, g_m \in \mathbb{R}[x]$ be G-invariant and $Q(g_1, \ldots, g_m)$, as defined above, Archimedean. If f is strictly positive on K, then*

$$f = \rho^G\Big(\sum_{i\leq h} q_0^i + \sum_{j\leq m} g_j \sum_{i\leq h} q_j^i\Big),$$

where $\rho^G(p) = \frac{1}{|G|}\sum_{\sigma\in G} p^\sigma$ and q_j^i are sums of squares in a set $\mathcal{S}^i$ basically consisting of the basis of the isotypic components of a real decomposition of $\mathbb{R}[x]$.

2.4 Sums of Arithmetic-Geometric Exponentials and Sums of Nonnegative Circuit Polynomials

Over the last years, there has been a strong interest in sparse nonnegativity certificates. The earliest results in this area are due to Reznick [Rez89], with a recent resurgence marked by the work of Pantea, Koeppl, and Craciun [PKC12]. In 2016, Iliman and de Wolff [IW16a], and Chandrasekaran and Shah [CS16] proposed specifically defined approaches to this topic, speaking either in the language of polynomials or exponential sums. They called their approach *sums of nonnegative circuit polynomials* (SONC) and *sums of arithmetic-geometric exponentials* (SAGE), respectively. Both notations, however, capture the same cone in the respective language of polynomials or exponentials [Wan18a; MCW21a].

Recent results (see the explanations in Equations (2.7) and (2.8) below) show that the approximation approach for the problem of polynomial nonnegativity using the SONC-cone can be reduced to the simpler problem of *exponential nonnegativity*. Hence, it suffices to examine a subcone of the cone of nonnegative exponential sums, which is our reason to only talk about the *SAGE-cone* in this thesis. Most results immediately follow for the polynomial case as well. Various results we review in this section — and use later on in the thesis –, however, have initially been stated for polynomials and under the name *SONC*. For the sake of completeness, we introduce SONC polynomials in this section as well, even though our setup is with respect to exponential sums, and hence, SAGE.

2.4.1 AGE Exponentials, Circuits and Nonnegative Circuit Exponentials

We start by introducing AGE exponentials. For $\mathcal{A} \subseteq \mathbb{R}^n$, an exponential sum supported on $\mathcal{A}$ is a function of the form

$$y \mapsto \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, y \rangle} \tag{2.6}$$

with real coefficients c_α. We are interested in *sparse* exponential sums, i.e., we consider a fixed *non-empty and finite* support set $\{\alpha : c_\alpha \neq 0\}$. This way, an exponential sum can be uniquely identified by its coefficient vector. Hence, we identify $\mathbb{R}^{\mathcal{A}}$, the space of coefficient vectors indexed by $\mathcal{A}$, with the space of exponential sums supported on $\mathcal{A}$, for $\mathcal{A}$ non-empty and finite. We denote both exponential sums f supported on $\mathcal{A}$ as well as the corresponding coefficient vector c as elements in $\mathbb{R}^{\mathcal{A}}$. The space of exponential sums with support $\mathcal{A}$ contained in $(2\mathbb{N})^n$ can be viewed as the space of polynomials supported on $\mathcal{A}$ and evaluated on the positive orthant by means of the substitution $|x_i| = e^{y_i}$.

We define AGE exponentials (sometimes also called arithmetic-geometric exponentials or AM/GM exponentials) to be nonnegative sums of exponentials with at most one negative term, i.e., for $\mathcal{A} \subseteq \mathbb{R}^n$ non-empty and finite nonnegative sums

$$\sum_{\alpha \in \mathcal{A} \setminus \{\beta\}} c_\alpha e^{\langle \alpha, x \rangle} + c_\beta e^{\langle \beta, x \rangle}$$

with $c_\alpha \in \mathbb{R}_+$ for all $\alpha \in \mathcal{A} \setminus \{\beta\}$ and $c_\beta \in \mathbb{R}$. Note that the *nonnegativity* of such a sum of exponentials is already captured by the name AGE exponential.

The name *AGE exponential* is chosen since nonnegativity of these exponentials can be certified using the *weighted AM/GM-inequality*:

Theorem 2.4.1 (Weighted AM/GM-inequality)**.** *Let $x_1, x_2, \ldots, x_n$ as well as the weights $w_1, w_2, \ldots, w_n$ be nonnegative and set $w = \sum_{i \le n} w_i$. If $w > 0$, then*

$$\frac{\sum_{i \le n} w_i x_i}{w} \ge \sqrt[w]{\prod_{i \le n} x_i^{w_i}}.$$

Equality holds if and only if $x_k = x_j$ for all $k, j \le n$ with $w_k > 0, w_j > 0$.

As stated in the introduction, for a set of support points $A \subseteq \mathbb{R}^n$, a vector $\beta \in \mathbb{R}^n$ and coefficients $\lambda \in \mathbb{R}_+^A$ satisfying $\mathbf{1}^T \lambda = 1$ and $\sum_{\alpha \in A} \lambda_\alpha \alpha = \beta$, the weighted AM/GM inequality yields

$$\sum_{\alpha \in A} \lambda_\alpha e^{\langle \alpha, x \rangle} \geqslant \prod_{\alpha \in A} e^{\lambda_\alpha \langle \alpha, x \rangle} = e^{\langle \beta, x \rangle}$$

and, hence, implies nonnegativity on $\mathbb{R}^n$ of exponential sums of the form

$$\sum_{\alpha \in A} \lambda_\alpha e^{\langle \alpha, x \rangle} - e^{\langle \beta, x \rangle}.$$

For a proof on the connection of nonnegativity of general exponential sums with at most one negative term and the arithmetic-geometric inequality see [CS16].

In the previous subsections, we mainly focused on *polynomial* optimization. The reason to switch to the language of exponential sums in this subsection and in the rest of this thesis is the fact that for $\mathcal{A} \subseteq \mathbb{N}^n$ non-empty and finite and $\beta \in \mathcal{A}$, the polynomial

$$\sum_{\alpha \in \mathcal{A} \setminus \{\beta\} \cap (2\mathbb{N})^n} c_\alpha x^\alpha + c_\beta x^\beta, \tag{2.7}$$

with $c_\alpha \ge 0$ for all $\alpha \in \mathcal{A} \setminus \{\beta\}$ is nonnegative on $\mathbb{R}^n$ if and only if

$$\sum_{\alpha \in \mathcal{A} \setminus \{\beta\} \cap (2\mathbb{N})^n} c_\alpha e^{\langle x, \alpha \rangle} - |c_\beta| e^{\langle x, \beta \rangle} \tag{2.8}$$

is nonnegative on $\mathbb{R}^n$. This way, we transform the problem of polynomial nonnegativity into a problem of exponential nonnegativity. For the special case of $\mathcal{A} \setminus \{\beta\}$ affinely independent and $\beta \in \operatorname{relint} \operatorname{conv}(\mathcal{A})$ (a *circuit* — see below for a formal definition), this was shown by Iliman and de Wolff [IW16a], and for general polynomials with at most one nonpositive coefficient by Murray, Chandrasekaran, and Wierman [MCW21a].

We denote the *cone of AGE exponentials* supported on $\mathcal{A} \subseteq \mathbb{R}^n$ with possibly negative term $\beta \in \mathcal{A}$ by $C_{\mathrm{AGE}}(\mathcal{A}, \beta)$. By a slight abuse of notation, we also denote the *cone of coefficients* of AGE exponentials by $C_{\mathrm{AGE}}(\mathcal{A}, \beta)$.

A prominent example of a polynomial which is nonnegative on $\mathbb{R}^n$ but not a sum of squares of other polynomials is the *Motzkin polynomial*

$$1 + x^2 y^4 + x^4 y^2 - 3x^2 y^2.$$

The term "$-3x^2y^2$" is the only negative term in this polynomial — hence, its exponential version is an AGE exponential and motivates the study of AGE exponentials.

Another important aspect when choosing the approximation in terms of AGE exponentials is the fact that containment in the cone of AGE exponentials (and, hence, certification of nonnegativity of an exponential sum with at most one negative coefficient) can be solved via relative entropy programming [CS16]: For an exponential sum $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle x, \alpha \rangle}$ with $c_\alpha \in \mathbb{R}$ for all $\alpha \in \mathcal{A}$, we can check membership in the AGE-cone using the *relative entropy function* known from Subsection 2.1.1.

Then, the following theorem examines nonnegativity of an exponential sum with at most one negative term.

Theorem 2.4.2 ([CS16])**.** *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite and fix $\beta \in \mathcal{A}$. The exponential $f(x) = \sum_{\alpha \in \mathcal{A} \setminus \{\beta\}} c_\alpha e^{\langle \alpha, x \rangle} + c_\beta e^{\langle \beta, x \rangle}$ with $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A} \setminus \{\beta\}$ is nonnegative if and only if there exists some $\nu \in \mathbb{R}_+^{\mathcal{A} \setminus \{\beta\}}$ such that $\sum_{\alpha \in \mathcal{A} \setminus \{\beta\}} \nu_\alpha (\alpha - \beta) = 0$ and*

$$D(\nu, ec) \leq c_\beta.$$

One direction of this theorem can be proven immediately using the arithmetic-geometric inequality.

An AGE exponential f with $\mathcal{A} := \mathrm{supp}(f) \subseteq \mathbb{R}^n$ is nonnegative only if all coefficients associated to vertices of $\mathrm{conv}(\mathcal{A})$ are positive; see e.g., [Fel+20] for a detailed proof in the language of polynomials. Thus, we make the assumption

$$\alpha \in \mathrm{Vert}\left(\mathrm{conv}(\mathcal{A})\right) \;\Rightarrow\; c_\alpha > 0. \tag{2.9}$$

Wang and independently also Murray, Chandrasekaran, and Wierman [Wan18a; MCW21a] proved that the extreme rays of the AGE-cone are supported on circuits:

Definition 2.4.3 (Circuit)**.** Let $A \subseteq \mathbb{R}^n$ and $\beta \in \mathbb{R}^n$. The tuple (A, β) is a *circuit* if A consists of affinely independent vectors and $\beta \in \mathrm{relint}\,\mathrm{conv}(A)$.

We address polynomials, exponential sums and other functions supported on circuits later on. Note that whenever we talk about polynomials supported on circuits, we additionally need $A \subseteq (2\mathbb{N})^n$ and $\beta \in \mathbb{N}^n$. The latter is obvious, the former follows as for polynomials, the prerequisites on the outer exponents need to be changed to

$$\alpha \in \mathrm{Vert}\left(\mathrm{conv}(A)\right) \;\Rightarrow\; c_\alpha x^\alpha > 0 \text{ for all } x \in \mathbb{R}^n \setminus \{\mathbf{0}\}.$$

Let $\mathcal{A} \subseteq \mathbb{R}^n$ be non-empty and finite, and

$$I(\mathcal{A}) = \Big\{ (A, \beta) \;:\; A \subseteq \mathcal{A} \text{ affinely independent}, \; \beta \in \mathrm{relint}(\mathrm{conv}\, A) \cap \mathcal{A} \Big\} \tag{2.10}$$

be the set of circuits (A, β) contained in $\mathcal{A}$.

For singleton sets $A = \{\alpha\}$, the tuples (A, β) are formally of the form $(\{\alpha\}, \alpha)$. By convention, we write these circuits simply as (α), and with this convention, the set $\{(\alpha) \,:\, \alpha \in \mathcal{A}\}$ is contained in $I(\mathcal{A})$.

With this notation, we can now introduce *nonnegative circuit exponentials.*

Definition 2.4.4 (Nonnegative Circuit Exponentials and Polynomials)**.** Let $\mathcal{A} \subseteq \mathbb{R}^n$.

1. For $(A, \beta) \in I(\mathcal{A})$, a *nonnegative circuit exponential* is an exponential sum in $\mathbb{R}^{\mathcal{A}}$ whose support equals $A \cup \{\beta\}$ and which is nonnegative on $\mathbb{R}^n$.

2. For $(A, \beta) \in I(\mathcal{A})$ with $A \subseteq (2\mathbb{N})^n$ and $\beta \in \mathbb{N}^n$, a *nonnegative circuit polynomial* is a polynomial in $\mathbb{R}[x_1, \ldots, x_n]$ supported on $A \cup \{\beta\}$ and which is nonnegative on $\mathbb{R}^n$. We denote the cone of *nonnegative circuit polynomials* supported on the circuit (A, β) by $P_{A,\beta}$.

Note that by (2.9), $c_\alpha \geq 0$ for all $\alpha \in A$ for any nonnegative circuit polynomial or exponential supported on the circuit (A, β); otherwise the corresponding function cannot be nonnegative.

The previously considered Motzkin polynomial is not only a polynomial version of an AGE exponential, but the Motzkin polynomial is also supported on the circuit $A = \{(0,0)^T, (2,4)^T, (4,2)^T\}$ and $\beta = (2,2)^T \in \operatorname{relint} \operatorname{conv}(A)$.

A crucial fact about a circuit exponential f is that its nonnegativity can be decided by an object Θ_f called the *circuit number* alone.

Theorem 2.4.5 ([IW16a, Theorem 1.1])**.** *Let $f(x) = \sum_{\alpha \in A} c_\alpha e^{\langle \alpha, x\rangle} + c_\beta e^{\langle \beta, x\rangle}$ supported on a circuit (A, β), and let $\lambda \in \mathbb{R}^A_{>0}$ denote the vector of barycentric coordinates of β in terms of A. Then, f is nonnegative if and only if $c_\alpha > 0$ for all $\alpha \in A$ and*

$$-c_\beta \;\leq\; \Theta_f \;=\; \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha}.$$

Remark 2.4.6. The polynomial equivalent for this theorem has the slight difference that we need "$|c_\beta|$" instead of "$-c_\beta$" whenever $\beta \notin (2\mathbb{N})^n$, i.e., for $A \subseteq (2\mathbb{N})^n$ and $\beta \in \mathbb{N}^n$, a polynomial $f(x) = \sum_{\alpha \in A} c_\alpha x^\alpha + c_\beta x^\beta$ supported on the circuit (A, β) with barycentric coordinates $\lambda \in \mathbb{R}^A_{>0}$ of β in terms of A is nonnegative if and only if $c_\alpha > 0$ for all $\alpha \in A$ and

$$|c_\beta| \;\leq\; \Theta_f \;=\; \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha}$$

or f is a sum of monomial squares.

Forsgård and de Wolff proposed a slightly different view on circuits [FW19]: They define a circuit as a minimally-supported nonzero vector $\lambda \in \ker \mathcal{A} \subseteq \mathbb{R}^{\mathcal{A}}$ for which $\sum_{\alpha \in \mathcal{A}} \lambda_\alpha = 0$. They call those circuits *simplicial* where λ contains a single negative component, say $\lambda_\beta < 0$. We usually canonically rescale $\lambda_\beta = -1$, so that simplicial circuits satisfy $\sum_{\alpha \in \operatorname{supp} \lambda \setminus \{\beta\}} \alpha \lambda_\alpha = \beta(-\lambda_\beta) = \beta$. This way, an associated normalized circuit λ coincides with the vector of barycentric coordinates for its support with respect to β, together with an additional coordinate $\lambda_\beta = -1$. Simplicial circuits therefore certify that β belongs to the relative interior of $A = \operatorname{supp} \lambda \setminus \{\beta\} \subseteq \mathcal{A}$. For the course of this thesis, we will only talk about simplicial circuits. Hence, we drop the term *simplicial* when speaking about circuits — even when using the terminology proposed by Forsgård and de Wolff.

We mention this slightly different view on the concept of circuits here because we need it for the results of *constrained optimization and nonnegativity*, examined in Chapters 6 and 7. Whenever switching to the constrained case, the concept of circuits is not as easy anymore, because — as we will see in Chapter 6 — we cannot determine the constrained version of a circuit in the sense of Forsgård and de Wolff solely by its support.

Now, as extreme rays of the cone of AGE exponentials are supported on circuits [Wan18a; MCW21a], we know that

$$\sum_{\beta \in \mathcal{A}} C_{\mathrm{AGE}}(\mathcal{A}, \beta) = \sum_{\beta \in \mathcal{A}} \sum_{(A,\beta) \in \mathcal{I}(\mathcal{A})} C_{\mathrm{AGE}}(A \cup \{\beta\}, \beta).$$

This sum contains exponential monomials by the definition of $\mathcal{I}(\mathcal{A})$ in particular. Hence, the cone of sums of AGE exponentials essentially is a cone of sums of nonnegative circuit exponentials, motivating the next subsection.

2.4.2 The SAGE-Cone and the SONC-Cone

The Minkowski sum

$$C_{\mathrm{SAGE}}(\mathcal{A}) \ = \sum_{\beta \in \mathcal{A}} C_{\mathrm{AGE}}(\mathcal{A}, \beta) \tag{2.11}$$

defines the *cone of sums of arithmetic-geometric exponentials (SAGE) with support set* $\mathcal{A}$ [CS16] and the Minkowski sum

$$C_{\mathrm{SONC}}(\mathcal{A}) \ = \sum_{\beta \in \mathcal{A}} \sum_{(A,\beta) \in \mathcal{I}(\mathcal{A})} P_{\mathcal{A},\beta}$$

defines the *cone of sums of nonnegative circuit polynomials (SONC)* with support set $\mathcal{A} \subseteq \mathbb{N}^n$ [IW16a; Ave19]. Both cones — as well as their subcones of AGE exponentials and nonnegative circuit polynomials — are closed convex cones in $\mathbb{R}^{\mathcal{A}}$ and $\mathbb{R}[x_1, \dots, x_n]$, respectively.

As in the case of the subcones of AGE exponentials, we often overload notation and write $f \in C_{\mathrm{SAGE}}(\mathcal{A})$ for an exponential sum f with coefficients in $C_{\mathrm{SAGE}}(\mathcal{A})$ as well as $c \in C_{\mathrm{SAGE}}(\mathcal{A})$ for the corresponding vector of coefficients.

The definition of both the SAGE-cone and the SONC-cone might seem a bit restrictive: per definition we do not allow support points in the support of the AGE exponentials or nonnegative circuit polynomials appearing in the decomposition which lie *outside* of the ground set $\mathcal{A} \subseteq \mathbb{R}^n$ of support points of the original SAGE exponential or the ground set $\mathcal{A} \subseteq \mathbb{N}^n$ of the original SONC polynomial. This could happen due to cancellation. Consider for example the two exponential sums

$$f^{(1)} := 1 + e^{3x+y} + e^{x+4y} - e^{x+2y} \text{ and } f^{(2)} := e^{3x} + e^{x+2y} - e^{3x+y} + e^{4x+3y}.$$

In the sum $f := f^{(1)} + f^{(2)}$ the terms $\pm e^{3x+y}$ and $\pm e^{x+2y}$ cancel, hence, the overall support of f does not contain $(3,1)^T$ and $(1,2)^T$ and f would not be captured in the set of SAGE exponentials with support in $\mathcal{A} := \mathrm{supp}(f)$ even though it *is* a sum of arithmetic-geometric exponentials and has $\mathrm{supp}(f) \subseteq \mathcal{A}$ (per definition of $\mathcal{A}$), compare Figures 2.1 and 2.2.

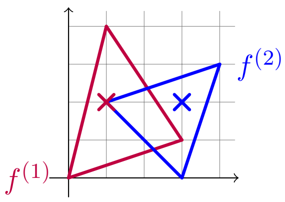

FIGURE 2.1: The supports of $f^{(1)}$ and $f^{(2)}$.

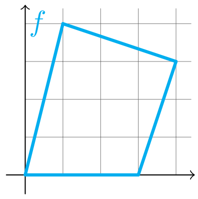

FIGURE 2.2: The support of f

In fact, such a phenomen cannot happen for our setup: In 2018, Wang proved a theorem in the language of polynomials, stating that for every SONC polynomial there always exists a decomposition without additional support points:

Theorem 2.4.7 ([Wan18b]). *Let $f = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha$. If f is a SONC polynomial, then f can be decomposed into a sum of nonnegative circuit polynomials $f^{(i)}$ with* $\operatorname{supp}(f^{(i)}) \subseteq \operatorname{supp}(f)$.

Similarly, Murray, Chandrasekaran, and Wierman proved a statement concerning the support of SAGE exponentials [MCW21a]. It is even stronger than the one observed by Wang because it additionally states that per negative term in the original SAGE exponential, there appears only one single AGE exponential in the decomposition, and in the support of this AGE exponential all exponents corresponding to the *other* negative coefficients in the original exponential sum do not appear.

Theorem 2.4.8 ([MCW21a]). *If c is contained in $C_{SAGE}(\mathcal{A})$ with non-empty set $N := \{\beta : c_\beta < 0\}$, then there exist vectors $(c^{(\beta)})_{\beta \in N}$ satisfying*

1. $c = \sum_{\beta \in N} c^{(\beta)}$,
2. $c^{(\beta)} \in C_{AGE}(\mathcal{A}, \beta)$ *for all* $\beta \in N$,
3. $c_\alpha^{(\beta)} = 0$ *for all* $\alpha \in N \setminus \{\beta\}$.

Combining Theorems 2.4.8 and 2.4.2 leads to the following nonnegativity certificate using the SAGE-cone.

Theorem 2.4.9. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite and $f \in \mathbb{R}^{\mathcal{A}}$ with coefficients $c \in \mathbb{R}^{\mathcal{A}}$ and non-empty set $N := \{\beta : c_\beta < 0\}$. Then, $f \in C_{\mathrm{SAGE}}$ if and only if for every $\beta \in N$ there exist $c^{(\beta)} \in \mathbb{R}_+^{\mathcal{A} \setminus N}$ and $\nu^{(\beta)} \in \mathbb{R}_+^{\mathcal{A} \setminus N}$ such that*

$$c_\alpha = \sum_{\beta \in N} c_\alpha^{(\beta)} \text{ for all } \alpha \in \mathcal{A} \setminus N \text{ and}$$

$$c^{(\beta)}_{|N \setminus \{\beta\}} = \mathbf{0}, \quad \sum_{\alpha \in \mathcal{A} \setminus N} \nu_\alpha^{(\beta)} (\alpha - \beta) = \mathbf{0} \text{ and } D(\nu^{(\beta)}_{\setminus N}, ec^{(\beta)}_{\setminus N}) \leq c_\beta \text{ for all } \beta \in N.$$

Here, $c^{(\beta)}_{|N \setminus \{\beta\}}$ denotes the restriction of the vector $c^{(\beta)}$ to elements in $N \setminus \{\beta\}$.

There has been a lot of interest in optimization techniques for SAGE exponentials, i.e., for the relaxation

$$\sup_{\gamma \in \mathbb{R}} \gamma \quad \text{s.t.} \quad f - \gamma \in C_{\mathrm{SAGE}}(\mathcal{A})$$

for some non-empty, finite set $\mathcal{A} \subseteq \mathbb{R}^n$. Iliman and de Wolff showed that for polynomials whose positive terms are affinely independent, containment in the SONC-cone can be determined by a geometric program [IW16b]. To do so, they computed a *circuit decomposition* and certified nonnegativity of the polynomial via the *circuit number*, i.e., they showed that each circuit polynomial in the decomposition of $f-\gamma$ then fulfills the conditions of Theorem 2.4.5. Another algorithm using the circuit number approach tries to compute the best possible covering of circuits, see the results by Magron, Seidler and de Wolff in [MSW19]. For algorithms implemented in POEM, see [Sd19].

Averkov stated that the SONC-cone can be written as the projection of a spectrahedron with semidefinite extension degree 2 [Ave19], and Wang and Magron provided a result for second-order representability of the primal SONC-cone [WM20a].

In contrast to the results using geometric programming, which only return the exact value for the described special case and an approximation of the SONC-cone in other cases, the SAGE relaxation can be computed efficiently via relative entropy programming for general cases [CS16; MCW21b; Mur20]. The approach using an approximated circuit decomposition, however, is sometimes faster and provides lower bounds in cases where the relative entropy computation fails. For a comparison of the various optimization approaches, see [Sd18].

Before turning to constrained optimization and constrained nonnegativity problems, we cover the dual cone of SAGE exponentials as well as the duals of its subcones of AGE exponentials briefly. Following the previous section, for an exponential sum $f \in \mathbb{R}^{\mathcal{A}}$ with coefficient vector $c \in \mathbb{R}^{\mathcal{A}}$ and $v(\cdot) \in (\mathbb{R}^{\mathcal{A}})^*$, the natural duality pairing is

$$v(f) = \sum_{\alpha \in \mathcal{A}} c_\alpha v_\alpha, \tag{2.12}$$

and the dual space $(\mathbb{R}^{\mathcal{A}})^*$ of $\mathbb{R}^{\mathcal{A}}$ can be identified with $\mathbb{R}^{\mathcal{A}}$. Then, the dual cone of SAGE exponentials is defined as

$$C_{\mathrm{SAGE}}(\mathcal{A})^* = \{v \in \mathbb{R}^{\mathcal{A}} : v(f) \geq 0 \text{ for all } f \in C_{\mathrm{SAGE}}(\mathcal{A})\}$$

(and analogously the dual cone of SONC polynomials).

The duality theory of the SAGE-cone was examined first by Chandrasekaran and Shah [CS16] in the language of exponential sums and later in [DNT21] in the language of polynomials and hence, for the SONC-cone, leading to the following results.

Theorem 2.4.10 ([CS16; DNT21]). *Let $\mathcal{A} \subseteq \mathbb{R}^n$ be non-empty and finite.*

1. *The dual cone of SAGE exponentials is the set*

$$\{\nu \in \mathbb{R}_+^{\mathcal{A}} : \textit{ for all } \beta \in \mathcal{A} \textit{ there exists } \tau^{(\beta)} \in \mathbb{R}^n \textit{ with}$$
$$\nu_\alpha \ln\left(\frac{\nu_\alpha}{\nu_\beta}\right) \leq (\alpha-\beta)^T \tau^{(\beta)} \textit{ for all } \alpha \in \mathcal{A} \setminus \{\beta\}\}.$$

2. *Let $\mathcal{A} \subseteq \mathbb{N}^n$ additionally. The dual cone of SONC-polynomials is the set*

$$\{\nu \in \mathbb{R}^{\mathcal{A}} : \nu_\alpha \geq 0 \textit{ for } \alpha \in \mathcal{A} \cap (2\mathbb{N})^n, \textit{ and for all } (A,\beta) \in I(\mathcal{A})$$
$$\textit{there exist } \nu^* \geq |\nu_\beta|, \tau \in \mathbb{R}^n : \nu^* \ln\left(\frac{\nu^*}{\nu_\alpha}\right) \leq (\beta-\alpha)^T \tau \textit{ for all } \alpha \in A\}.$$

Both of these representations use projections to describe the dual of the SAGE- and SONC-cone.

2.4.3 The Conditional SAGE-Cone

Over the course of the years, there have been several approaches to constrained SONC relaxations, i.e., to minimizing a SONC polynomial over some set of constraints.

The earliest approaches try to determine whether a sum of nonnegative circuit polynomials, defined in the same manner as in the unconstrained case, is nonnegative over some set X.

A disadvantage of this approach (and other approaches) compared to SOS is that there does not exist a Putinar-like Positivstellensatz. A multivariate counter-example was given in [DKW21] and a univariate one can be found in [KNT21]. In [DIW17], Dressler, Iliman and de Wolff provided a SONC-specific Positivstellensatz, which essentially follows from Krivine's Positivstellensatz, see [Dre18].

In this thesis, however, we take a different approach for the conditional SAGE-cone using a setting proposed by Murray, Chandrasekaran, and Wierman [MCW21b].

For a finite non-empty set of support points $\mathcal{A} \subseteq \mathbb{R}^n$ and some convex and non-empty set $X \subseteq \mathbb{R}^n$, we define *X-AGE* exponentials as exponential sums with at most one negative term being nonnegative on X. The cone of X-AGE exponentials with support contained in $\mathcal{A}$ and distinguished exponent $\beta \in \mathcal{A}$ is denoted $C_X(\mathcal{A}, \beta)$. As in the unconstrained case, we often overload notation and consider $C_X(\mathcal{A}, \beta)$ as a cone of coefficient vectors in $\mathbb{R}^{\mathcal{A}}$.

In this thesis, we only consider sets $\mathcal{A}$ and X where the functions $\{x \mapsto e^{\langle x, \alpha \rangle}\}_{\alpha \in \mathcal{A}}$ are linearly independent on X; this assumption is necessary to prevent the nonnegativity cone from containing a lineality space. A direct consequence of this assumption is that the moment cone $\operatorname{pos}\{e^{\mathcal{A}^T x} \in \mathbb{R}^{\mathcal{A}} : x \in X\}$ is full-dimensional.

The *conditional SAGE-cone* again is defined analogously to the unconstrained situation.

Definition 2.4.11. Let $\emptyset \neq X \subseteq \mathbb{R}^n$ be a convex set. The X-SAGE-cone with respect to exponents $\mathcal{A}$ is the Minkowski sum

$$C_X(\mathcal{A}) = \sum_{\beta \in \mathcal{A}} C_X(\mathcal{A}, \beta).$$

At this point, there does not exist a proper connection to SONC polynomials (or circuits at all). Whereas in the unconstrained situation, it is necessary for the exponent corresponding to the non-positive term to lie in the convex hull of all other exponents, this is not necessary anymore for a real subset $X \subsetneq \mathbb{R}^n$. Hence, a circuit structure does not exist per se. However, in Chapter 6 of this thesis, we establish the notion of *sublinear circuits* as a generalization of $\mathbb{R}^n$-circuits and show that extremal rays of the X-SAGE-cone are defined with respect to these objects. Hence, the polynomial version of this X-SAGE-cone can be seen as a conditional equivalent to the SONC-cone.

When Murray, Chandrasekaran, and Wierman introduced this concept of conditional SAGE [MCW21b], they showed how to efficiently check membership in an X-SAGE-cone whenever X is a tractable convex set using relative entropy programming again and an object called the *support function* of the constraint set X:

$$\sigma_X(y) = \sup_{x \in X} y^T x$$

With the notation $N_\beta = \{\nu \in \mathbb{R}_+^{\mathcal{A} \setminus \{\beta\}} \times \mathbb{R} : 1^T \nu = 0\}$ for vectors $\beta \in \mathcal{A}$, we can phrase the conditional nonnegativity certificates for SAGE exponentials by Murray, Chandrasekaran, and Wierman as follows.

Proposition 2.4.12 ([MCW21b], Theorem 6)**.** *Let $X \subseteq \mathbb{R}^n$ be a convex set, $\mathcal{A} \subseteq \mathbb{R}^n$ non-empty and finite and $\beta \in \mathcal{A}$. An exponential sum $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle}$ belongs to $C_X(\mathcal{A}, \beta)$ if and only if $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A} \setminus \{\beta\}$ and some $\nu \in \mathbb{R}^{\mathcal{A}}$ satisfies*

$$\nu \in N_\beta \quad \text{and} \quad \sigma_X(-\mathcal{A}\nu) + D(\nu_{\setminus\beta}, ec_{\setminus\beta}) \leq c_\beta. \tag{2.13}$$

For the constrained case, we can take advantage of a decomposition result analogous to Theorem 2.4.8. Again, there always exists a decomposition of an X-SAGE exponential into X-AGE exponentials. There do not appear additional support points. For every negative term, there exists exactly one X-AGE exponential and the negative term does not appear positively in any other exponential sum in the decomposition.

Theorem 2.4.13. *If $X \subseteq \mathbb{R}^n$ is a convex set and c is a vector in $C_X(\mathcal{A})$ with non-empty set $N := \{\beta : c_\beta < 0\}$, then there exist vectors $(c^{(\beta)})_{\beta \in N}$ satisfying*

1. $c = \sum_{\beta \in N} c^{(\beta)}$,
2. $c^{(\beta)} \in C_X(\mathcal{A}, \beta)$ *for all* $\beta \in N$,
3. $c_\alpha^{(\beta)} = 0$ *for all* $\alpha \in N \setminus \{\beta\}$.

An implementation of the relative entropy program certifying nonnegativity of an exponential sum over some convex and non-empty set of constraints X using the X-SAGE approach can be found in [Mur20]. The results can also be applied to polynomial optimization, which was also covered by Murray, Chandrasekaran, and Wierman in the same article.

We conclude this section with a brief look at the duality theory for constrained SAGE. For an exponential sum f with a vector of coefficients c and some v in the dual space, we again use the duality pairing $v(f) = \sum_{\alpha \in \mathcal{A}} v_\alpha c_\alpha$. With this, the dual of the X-SAGE-cone

$$\{v \in \mathbb{R}^{\mathcal{A}} : v(f) \geq 0 \text{ for all } f \in C_X(\mathcal{A})\}$$

can be represented as follows.

Theorem 2.4.14. *Let $\mathcal{A} \subseteq \mathbb{R}^n$ non-empty and finite and $X \subseteq \mathbb{R}^n$ a convex set. The dual cone of X-SAGE exponentials is the set*

$$C_X^*(\mathcal{A}) = \left\{ v \in \mathbb{R}_+^{\mathcal{A}} : v_\beta \ln\left(\frac{v_\alpha}{v_\beta}\right) \geq (\alpha - \beta)^T z : z/v_\beta \in X \textit{ for all } \beta \in \mathcal{A}, \alpha \in \mathcal{A} \setminus \{\beta\} \right\}.$$

Chapter 3

Extremality and Duality Theory of the SAGE-Cone

In this chapter, we investigate the cone of SAGE exponentials as well as its dual. We are interested in structural questions such as the form of extremal rays of the SAGE-cone and how these can be deduced from duality theory.

We start in Section 3.1 with developing some generalized form of nonnegativity certificate for SAGE exponentials, see Theorem 3.1.1. The statement contains certificates of nonnegativity we already know from the earlier works of Iliman and de Wolff [IW16a] and Chandrasekaran and Shah [CS16] and partially extends those. Then, we introduce the notion of a *reduced* circuit, see Definition 3.1.3. The main theorem of this section is Theorem 3.1.5: Here, we derive projection-free representations of the dual of the SAGE-cone. We compare them to known representations from [CS16] and [DNT21] in Proposition 3.1.8.

Building upon the dual representations derived in the first section, in Section 3.2, we turn to examining extremality of the SAGE-cone. In Theorem 3.2.1, we review the proof for sparsity preservation of SAGE exponentials. In this particular setting, the statement was already observed in [MCW21a]. Here, however, we prove it using the dual representation from Theorem 3.1.5. The second part of Theorem 3.2.1 is novel. Here, we show that any SAGE exponential can be decomposed into a sum of AGE exponentials supported on reduced circuits. This leads to the main theorem of this second section, namely, Theorem 3.2.4, which provides a complete characterization of the extreme rays of the SAGE-cone.

3.1 Reduced Circuits and Duality Theory

Before we start by introducing reduced circuits, we revisit the existing certificates for nonnegativity of SONC polynomials and SAGE exponentials. Particularly, we provide a variant of the *circuit number* certificate, which extends to general support sets. To do so, it is useful to introduce the following notation:

For a non-empty finite set $\mathcal{A} \subseteq \mathbb{R}^n$ and $\beta \in \mathbb{R}^n \setminus \mathcal{A}$ let $\Lambda(\mathcal{A}, \beta)$ be the polytope

$$\Lambda(\mathcal{A}, \beta) := \left\{ \lambda \in \mathbb{R}_+^{\mathcal{A}} \;\middle|\; \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \alpha = \beta, \ \sum_{\alpha \in \mathcal{A}} \lambda_\alpha = 1 \right\} \tag{3.1}$$

and let $\Lambda(\mathcal{A})$ be the union $\bigcup_{\beta \in \mathcal{A}} \Lambda(\mathcal{A} \setminus \{\beta\}, \beta)$ Note that $\Lambda(\mathcal{A}, \beta) \neq \emptyset$ if and only if β is contained in the convex hull of $\mathcal{A}$, which is our reason for sometimes referring to the set $\mathcal{A}$ as the set of outer exponents and to β as the inner exponent. In the special case that $\mathcal{A}$ is affinely independent, $\Lambda(\mathcal{A}, \beta)$ consists of a single element, which we denote by $\lambda(\mathcal{A}, \beta)$. In particular, the tuple $(\mathcal{A}, \beta)$ defines a circuit in this situation and λ defines the barycentric coordinates of the set $\mathcal{A}$ with respect to β.

Theorem 3.1.1. *Let $\mathcal{A} \subseteq \mathbb{R}^n$ be a non-empty finite set, $\beta \in \mathbb{R}^n \setminus \mathcal{A}$ and f be an AGE exponential of the form*

$$f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} + c_\beta e^{\langle \beta, x \rangle}$$

where $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$ and $c_\beta \in \mathbb{R}$. Then, the following statements are equivalent:

1. *$f(x) \geq 0$ for all $x \in \mathbb{R}^n$.*
2. *There exists a $\nu \in \mathbb{R}_+^{\mathcal{A}}$ such that $\sum_{\alpha \in \mathcal{A}} \nu_\alpha \alpha = (\sum_{\alpha \in \mathcal{A}} \nu_\alpha)\beta$ and*

$$D(\nu, e \cdot c) \leq c_\beta.$$

3. *There exists a $\lambda \in \Lambda(\mathcal{A}, \beta)$ such that*

$$\prod_{\alpha \in \mathcal{A}} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq -c_\beta.$$

A vector $\lambda \in \Lambda(\mathcal{A}, \beta)$ as in this theorem is called an *AGE witness*. Note that with the convention $\left(\frac{y}{0}\right)^0 = 1$ for $y \in \mathbb{R}$ we introduced in the previous chapter, this is well-defined even if $\lambda_\alpha = 0$ for some $\alpha \in \mathcal{A}$.

Proof of Theorem 3.1.1. Observe that the equivalence of the first two statements is precisely Theorem 2.4.9.

For the implication (2) $\implies$ (3), set $\lambda := (\sum_{\alpha \in \mathcal{A}} \nu_\alpha)^{-1} \nu$. It is clear from the properties of ν that $\lambda \in \Lambda(\mathcal{A}, \beta)$. The discussion in [CS16, p. 1151] shows that

$$\prod_{\alpha \in \mathcal{A}} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq -D(\nu, e \cdot c),$$

and thus, this λ has the desired properties. The implication (3) $\implies$ (1) is a direct consequence of the weighted arithmetic-geometric mean inequality:

$$\sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} \overset{\text{AM/GM-inequality}}{\geq} \prod_{\alpha \in \mathcal{A}} \left(\frac{c_\alpha}{\lambda_\alpha} e^{\langle \alpha, x \rangle} \right)^{\lambda_\alpha} = \prod_{\alpha \in \mathcal{A}} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} e^{\langle \beta, x \rangle}.$$

Using (3), we obtain

$$\sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} + c_\beta e^{\langle \beta, x \rangle} \geq e^{\langle \beta, x \rangle} \cdot (-c_\beta + c_\beta) \geq 0.$$

As we already know that (1) $\iff$ (2), we obtain the desired statement. □

Example 3.1.2. Let $\mathcal{A} = \{1, 2, 3\} \subseteq \mathbb{N}$ and consider

$$g(x) = e^x + c_2 e^{2x} + e^{3x}.$$

We seek to find the smallest c_2 for which g is a nonnegative function on $\mathbb{R}$. Since the equality condition in statement (2) of Theorem 3.1.1 is satisfied for all $\nu = \delta \mathbf{1}$ for $\delta \geq 0$, we have $g(x) \geq 0$ for all $x \in \mathbb{R}$ if and only if there exists a $\delta \in \mathbb{R}_+$ with

$$2\delta \ln \left(\frac{\delta}{e} \right) \leq c_2. \tag{3.2}$$

Since the function $x \ln\left(\frac{x}{e}\right)$ attains its minimum at $x = 1$ yielding $x \ln\left(\frac{x}{e}\right) \geq -1$, we obtain $c_2 \geq -2$; hence, the minimal value for c_2 is -2.

We now introduce the concept of *reduced* circuits, which we use to determine a variant of the projection-free representation of the dual cone. In particular, in the following section, we determine the extreme rays of the SAGE-cone using these types of circuits. Recall that a *circuit* is a tuple (A, β) with $A \subseteq \mathbb{R}^n$ being affinely independent and $\beta \in \operatorname{relint}\operatorname{conv}(A)$. The *set of circuits* supported on some ground set $\mathcal{A}$ is denoted $\mathcal{I}(\mathcal{A})$.

Definition 3.1.3. For a circuit $(A, \beta) \in \mathcal{I}(\mathcal{A})$ with $\mathcal{A} \subseteq \mathbb{R}^n$ let

$$r(A, \beta) := |\,(\operatorname{conv}(A) \setminus A) \cap \mathcal{A}|.$$

A circuit (A, β) is called *reduced* if $r(A, \beta) = 0$.

In other words, reduced circuits contain no elements of the ground set $\mathcal{A}$ in their convex hull except those which are trivially there.

Example 3.1.4. Whether a circuit is reduced or not depends on the ground set $\mathcal{A}$. For example, the circuit (A, β) with $A = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$ and $\beta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is *reduced* for the ground set $\mathcal{A} = A \cup \{\beta\} \cup \left\{ \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right\}$ (compare Figure 3.1 below) but *not reduced* for $\mathcal{A} = A \cup \{\beta\} \cup \left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}$ (compare Figure 3.2).

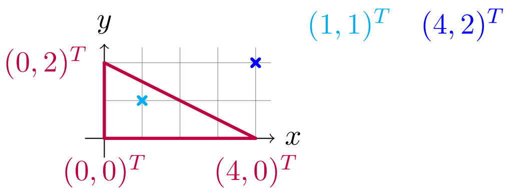

FIGURE 3.1: The circuit is reduced, as $(4,2)^T \notin \operatorname{conv}(A)$.

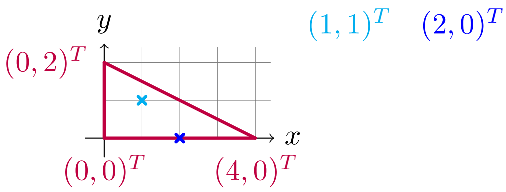

FIGURE 3.2: The circuit is not reduced, as $(2,0)^T \in \operatorname{conv}(A)$.

We can now provide the following characterization of the dual SAGE-cone. Again, we use the convention that $0\ln(0) = 0$ and $\ln(0) = -\infty$.

Theorem 3.1.5. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite and let $v \in \mathbb{R}^{\mathcal{A}}$.*

(1) If $v \in C_{\mathrm{SAGE}}(\mathcal{A})^*$, then $v_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$.

(2) If the condition of part (1) is satisfied, then the following are equivalent:

(a) v lies in the dual cone $C_{\mathrm{SAGE}}(\mathcal{A})^*$.

(b) For all $\beta \in \mathcal{A}$ and all $\lambda \in \Lambda(\mathcal{A}, \beta)$, it holds that

$$\ln(v_\beta) \leq \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \ln(v_\alpha).$$

(c) For every circuit $(A, \beta) \in I(\mathcal{A})$ and $\lambda = \lambda(A, \beta)$, it holds that

$$\ln(v_\beta) \leq \sum_{\alpha \in A} \lambda_\alpha \ln(v_\alpha).$$

(d) For every reduced circuit $(A, \beta) \in I(\mathcal{A})$ and $\lambda = \lambda(A, \beta)$, it holds that

$$\ln(v_\beta) \leq \sum_{\alpha \in A} \lambda_\alpha \ln(v_\alpha).$$

Formally, we included elements $(\{\alpha\}, \alpha)$ in the set of circuits, i.e., elements where outer and inner exponent coincide. Such a circuit is always reduced, and in fact, the statements hold for this kind of circuit trivially.

Before we prove Theorem 3.1.5, we consider the dual of the sub-cone $C_{\text{AGE}}(\mathcal{A}, \beta)$ of $C_{\text{SAGE}}(\mathcal{A})$. Note that one direction of this statement can be found in [MCW21a], followed by a partial — but non-complete — statement addressing the second inclusion.

Lemma 3.1.6. *Let $\mathcal{A} \subseteq \mathbb{R}^n$ be a non-empty finite set. For $\beta \in \mathcal{A}$, the dual cone of $C_{\text{AGE}}(\mathcal{A}, \beta)$ consists of those $v \in \mathbb{R}^{\mathcal{A}}$ where*

1. *$v_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$ and*
2. *$\ln(v_\beta) \leq \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \ln(v_\alpha)$ for all $\lambda \in \Lambda(\mathcal{A}, \beta)$.*

Proof. Let $v \in (C_{\text{AGE}}(\mathcal{A}, \beta))^*$. We show that it satisfies the claimed conditions.

1. For every $\alpha \in \mathcal{A}$, it holds that $e^{\langle \alpha, x \rangle} \in C_{\text{AGE}}(\mathcal{A}, \beta)$. Thus, $0 \leq v(e^{\langle \alpha, x \rangle}) = v_\alpha$, as claimed.
2. Fix a $\lambda \in \Lambda(\mathcal{A}, \beta)$. First assume that $v_\alpha \neq 0$ for all $\alpha \in \mathcal{A}$. Then

$$f := \sum_{\alpha \in \mathcal{A}} \left(\prod_{\alpha' \in \mathcal{A}} v_{\alpha'}^{\lambda_{\alpha'}} \right) \frac{\lambda_\alpha}{v_\alpha} e^{\langle \alpha, x \rangle} - e^{\langle \beta, x \rangle}$$

is an AGE exponential and a straightforward computation shows that f satisfies the condition (3) of Theorem 3.1.1 (with the given λ), hence, f is nonnegative. Thus,

$$0 \leq v(f) = \prod_{\alpha \in \mathcal{A}} v_\alpha^{\lambda_\alpha} - v_\beta$$

which is equivalent to property *(2)*. Since the mapping (2.12) is continuous in v, the statements also hold if $v_\alpha = 0$ for some $\alpha \in \mathcal{A}$.

For the converse implication, assume that v satisfies conditions *(1)* and *(2)*.

We need to show that every AGE exponential $f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle x, \alpha \rangle} + c_\beta e^{\langle x, \beta \rangle}$ satisfies $v(f) \geq 0$. Let $\lambda \in \Lambda(\mathcal{A}, \beta)$ be an AGE witness for f as in Theorem 3.1.1. Using the AM/GM-inequality once again, observe that

$$\begin{aligned} \sum_{\alpha \in \mathcal{A}} v_\alpha c_\alpha &= \sum_{\alpha \in \mathcal{A}, \lambda_\alpha > 0} \lambda_\alpha \left(\frac{v_\alpha c_\alpha}{\lambda_\alpha} \right) \geq \prod_{\alpha \in \mathcal{A}, \lambda_\alpha > 0} \left(\frac{v_\alpha c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \\ &= \prod_{\alpha \in \mathcal{A}, \lambda_\alpha > 0} v_\alpha^{\lambda_\alpha} \cdot \prod_{\alpha \in \mathcal{A}, \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq v_\beta \prod_{\alpha \in \mathcal{A}} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha}, \end{aligned}$$

which implies

$$v(f) = \sum_{\alpha \in \mathcal{A} \setminus \{\beta\}} v_\alpha c_\alpha + v_\beta c_\beta \geq v_\beta \left(\prod_{\alpha \in \mathcal{A} \setminus \{\beta\}} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} + c_\beta \right). \tag{3.3}$$

The right expression in (3.3) is nonnegative, because f is an AGE exponential. $\square$

In addition, we need the following lemma for the proof of Theorem 3.1.5.

Lemma 3.1.7 (Essentially [MCW21a], Lemma 8)**.** *Let $\mathcal{A} \subseteq \mathbb{R}^n$ be a non-empty finite set and $\beta \in \operatorname{conv}(\mathcal{A})$. Then, every $\lambda \in \Lambda(\mathcal{A}, \beta)$ can be written as a sum*

$$\lambda = \sum_{j=1}^{k} \mu_j \lambda^{(j)}$$

with $k \geq 1$, $\mu \in \mathbb{R}_+^k, \sum_{j=1}^k \mu_j = 1$ and $\lambda^{(j)} \in \Lambda(\mathcal{A}, \beta)$ for all j such that the support of each $\lambda^{(j)}$ is affinely independent.

Proof. Since the polytope $\Lambda(\mathcal{A}, \beta)$ is the convex hull of its vertices, it suffices to show that the support of every vertex of $\Lambda(\mathcal{A}, \beta)$ is an affinely independent set.

Let λ be a vertex of $\Lambda(\mathcal{A}, \beta)$ and $\mathcal{A}' := \{\alpha \mid \lambda_\alpha > 0\}$ be its support. Assume to the contrary that $\mathcal{A}'$ is affinely dependent. Then there exists $\mu \in \mathbb{R}^{\mathcal{A}} \setminus \{\mathbf{0}\}$ with $\sum_{\alpha \in \mathcal{A}'} \mu_\alpha = 0$, $\sum_{\alpha \in \mathcal{A}'} \mu_\alpha \alpha = 0$ and $\mu_\alpha = 0$ for $\alpha \notin \mathcal{A}'$. Since $\lambda_\alpha > 0$ for all $\alpha \in \mathcal{A}'$, for sufficiently small $\epsilon > 0$ both $\lambda + \epsilon\mu$ and $\lambda - \epsilon\mu$ are contained in $\Lambda(\mathcal{A}, \beta)$. But this implies that $\lambda = \frac{1}{2}(\lambda + \epsilon\mu) + \frac{1}{2}(\lambda - \epsilon\mu)$ is not a vertex of $\Lambda(\mathcal{A}, \beta)$, a contradiction. ☐

Proof of Theorem 3.1.5. (1): Since $e^{\langle \alpha, x \rangle} \in C_{\mathrm{SAGE}}(\mathcal{A})$ for every $\alpha \in \mathcal{A}$, we have that every $v \in C_{\mathrm{SAGE}}(\mathcal{A})^*$ satisfies

$$0 \leq v(e^{\langle \alpha, x \rangle}) = v_\alpha.$$

(2): The implications (b) $\implies$ (c) $\implies$ (d) are trivial. For the equivalence of (a) and (b) note that

$$C_{\mathrm{SAGE}}(\mathcal{A})^* = \bigcap_{\alpha \in \mathcal{A}} (C_{\mathrm{AGE}}(\mathcal{A} \setminus \{\alpha\}, \alpha))^*,$$

because Minkowski sum and intersection are dual operations, see, e.g., [Sch14], Theorem 1.6.3. Hence, the claim follows with Lemma 3.1.6.

It remains to show (c) $\implies$ (b) and (d) $\implies$ (c).

(c) $\implies$ (b):
Now let $\beta \in \mathcal{A}$ and $\lambda \in \Lambda(\mathcal{A}, \beta)$. Then, applying Lemma 3.1.7, we can decompose λ as $\lambda = \sum_{j=1}^k \mu_j \lambda^{(j)}$ with $k \geq 1$, $\mu \in \mathbb{R}_+^k, \sum_{j=1}^k \mu_j = 1$ and $\lambda^{(1)}, \ldots, \lambda^{(k)} \in \Lambda(\mathcal{A}, \beta)$ so that the support of each $\lambda^{(j)}$ is affinely independent. Now the claim follows from

$$\begin{aligned} \ln(v_\beta) &= \sum_{j=1}^{k} \mu_j \ln(v_\beta) \overset{(c)}{\leq} \sum_{j=1}^{k} \mu_j \sum_{\alpha} \lambda_\alpha^{(j)} \ln v_\alpha \\ &= \sum_{\alpha} \ln v_\alpha \sum_{j=1}^{k} \mu_j \lambda_\alpha^{(j)} = \sum_{\alpha} \lambda_\alpha \ln v_\alpha. \end{aligned}$$

(d) $\implies$ (c):
We proceed by induction on $r = r(A, \beta)$. Since the base case $r = 0$ captures exactly the reduced circuits, there is nothing to prove in this case.

Now consider a circuit $(A, \beta) \in I(\mathcal{A})$ with $r(A, \beta) > 0$. Then, there exists a $\beta' \in \operatorname{conv}(A) \cap \mathcal{A}$ with $\beta' \notin A$ and $\beta' \neq \beta$.

Set $\lambda := \lambda(A, \beta)$ and $\lambda' := \lambda(A, \beta')$. Let $\tau \geq 0$ be the maximal real number with $\widetilde{\lambda} := \lambda - \tau\lambda' \in \mathbb{R}_+^A$. This number clearly exists, and we have $\tau \leq 1$ because the coordinate sums of λ and λ' are equal. Further, it holds that $\tau > 0$ because all components of λ are positive.
Similarly, let τ' be the maximal real number with $\widetilde{\lambda}' := \lambda' - \tau'\lambda \in \mathbb{R}_+^A$. As above, it holds that $0 \leq \tau' \leq 1$. Moreover, note that $\beta \neq \beta'$ implies $\tau, \tau' < 1$. The construction gives

$$\begin{aligned} \beta &= \sum_{\alpha\in A} \widetilde{\lambda}_\alpha \alpha + \tau\beta', & \sum_{\alpha\in A} \widetilde{\lambda}_\alpha + \tau &= 1, \\ \beta' &= \sum_{\alpha\in A} \widetilde{\lambda}'_\alpha \alpha + \tau'\beta \quad \text{and} & \sum_{\alpha\in A} \widetilde{\lambda}'_\alpha + \tau' &= 1. \end{aligned}$$

Note that at least one of the entries of $\widetilde{\lambda}$ is zero, and, moreover, τ' or at least one of the entries of $\widetilde{\lambda}'$ is zero. Define two new circuits (A_1, β) and (A_2, β') with $A_1 := \operatorname{supp}(\widetilde{\lambda}) \cup \{\beta'\}$ and

$$A_2 := \begin{cases} \operatorname{supp}(\widetilde{\lambda}') \cup \{\beta\} & \text{if } \tau' > 0, \\ \operatorname{supp}(\widetilde{\lambda}') & \text{if } \tau' = 0. \end{cases}$$

We observe $\operatorname{conv}(A_1) \subsetneq \operatorname{conv}(A)$, and since β' is not counted towards $r(A_1, \beta)$, it follows that $r(A_1, \beta) < r(A, \beta)$. Similarly, since $\operatorname{conv}(A_2) \subseteq \operatorname{conv}(A)$ and β' is not counted towards $r(A_2, \beta')$, we obtain $r(A_2, \beta') < r(A, \beta)$. Hence, by induction,

$$\ln(v_\beta) \leq \sum_{\alpha\in A} \widetilde{\lambda}_\alpha \ln(v_\alpha) + \tau \ln(v_{\beta'}) \qquad \text{and} \tag{3.4}$$

$$\ln(|v_{\beta'}|) \leq \sum_{\alpha\in A} \widetilde{\lambda}'_\alpha \ln(v_\alpha) + \tau' \ln(v_\beta). \tag{3.5}$$

Note that $v_{\beta'} \geq 0$ and $v_\beta \geq 0$. Adding τ times (3.5) to (3.4) gives, due to $\widetilde{\lambda} + \tau\widetilde{\lambda}' = (1 - \tau\tau')\lambda$, the uniform inequality

$$0 \leq (1 - \tau\tau')\Big(\sum_{\alpha\in A} \lambda_\alpha \ln v_\alpha - \ln |v_\beta| \Big).$$

Since $1 - \tau\tau' > 0$, this proves the claim. □

The descriptions of the dual of the SONC-cone and of the dual of the SAGE-cone in Theorem 2.4.10 are based on projections and differ from the one in Theorem 3.1.5. For completeness, here we show that they are in fact equivalent.

Proposition 3.1.8. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be a finite set and $\beta \in \mathcal{A} \cap \operatorname{conv}(\mathcal{A} \setminus \{\beta\})$. For $v \in \mathbb{R}_{>0}^{\mathcal{A}}$, the following are equivalent:*

(1) $\forall \lambda \in \Lambda(\mathcal{A}, \beta)\colon\ \ln(v_\beta) \leq \sum_{\alpha\in\mathcal{A}} \lambda_\alpha \ln(v_\alpha)$.

(2) $\exists \tau \in \mathbb{R}^n,\ \forall \alpha \in \mathcal{A}\colon v_\beta \ln\left(\frac{v_\beta}{v_\alpha}\right) \leq (\beta - \alpha)^T\tau$.

(3) $\exists v^* \geq v_\beta, \exists \tau \in \mathbb{R}^n, \forall \alpha \in \mathcal{A}\colon v^* \ln\left(\frac{v^*}{v_\alpha}\right) \leq (\beta - \alpha)^T\tau$.

In this proposition, statement (1) is the one we used earlier, statement (2) is the description of the dual SAGE-cone used in [CS16], and statement (3) in conjunction with Theorem 3.1.1(c) is the description of the dual SONC-cone used in [DNT21].

Note that for general $v \in \mathbb{R}_+^{\mathcal{A}}$, the statement is not necessarily true:

Example 3.1.9. Let $\mathcal{A} = \{(0,0)^T, (0,1)^T, (0,2)^T, (1,1)^T\}$ and consider the AGE-cone $C_{\text{AGE}}(\mathcal{A}, \beta)$ with respect to $\beta = (0,1)^T$. Then, the unique $\lambda \in \Lambda(\mathcal{A}, \beta)$ has $\lambda_{(1,1)} = 0$.

According to the first formulation of Proposition 3.1.8, some element $v \in \mathbb{R}^{\mathcal{A}}$ with $v_{0,1} = \sqrt{v_{0,0}v_{0,2}}, v_{0,0} \neq 0, v_{0,2} \neq 0$, and $v_{1,1} = 0$ is contained in the dual cone of AGE exponentials. But according to (2), it is not, as $v_{1,1} = 0$ yields $v_{0,1} = 0$ in this formulation.

This can be solved by expressing $C^*_{\text{SAGE}}(\mathcal{A})$ for $\mathcal{A} \subseteq \mathbb{R}^n$ non-empty and finite with the cleaner formulation

$$\text{cl}\{\nu \in \mathbb{R}^{\mathcal{A}}_{>0} : \ \forall\, \beta \in \mathcal{A}\, \exists\, \tau^{(\beta)} \in \mathbb{R}^n : \nu_\alpha \ln\left(\frac{\nu_\alpha}{\nu_\beta}\right) \leq (\alpha - \beta)^T \tau^{(\beta)}\ \forall\, \alpha \in \mathcal{A} \setminus \{\beta\}\}$$

instead of the one used in Theorem 2.4.10. Indeed, for $v_{1,1} \to 0$, the witness $\tau \in \mathbb{R}^2$ for the above example set $\mathcal{A}$ can then be obtained by choosing $\tau_2 = v_{0,1} \ln\left(\frac{\sqrt{v_{0,2}}}{\sqrt{v_{0,0}}}\right)$ and $\tau_1 = v_{0,1} \ln(\frac{v_{1,1}}{v_{0,1}}) - \tau_2$.

Remark 3.1.10. We show the equivalence via the following variant of statement (2):

(2') $\exists \tau \in \mathbb{R}^n,\ \forall \alpha \in \mathcal{A}\colon \ln\left(\frac{v_\beta}{v_\alpha}\right) \leq (\alpha - \beta)^T \tau$.

Proof. By the precondition, $v_\beta \neq 0$ and $v_\alpha \neq 0$ for all $\alpha \in \mathcal{A}$.

(1) $\iff$ (2'):

Consider (2') as the feasibility set of a system of linear inequalities in τ. (2') is satisfied if and only if its Farkas alternative system (in the version of Proposition 1.7 of [Zie95])

$$\exists \lambda \in \mathbb{R}_+^{\mathcal{A}} : \quad \sum_{\alpha \in \mathcal{A}} \lambda_\alpha(-\alpha + \beta) = 0 \quad \text{and} \quad \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \cdot \left(-\ln\left(\frac{v_\beta}{v_\alpha}\right)\right) < 0$$

does not have a solution.

We can normalize λ so that all its components sum to 1. Hence, the alternative system simplifies to

$$\sum_{\alpha \in \mathcal{A}} \lambda_\alpha \ln v_\beta > \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \ln v_\alpha > 0,$$

i.e., to $\ln(v_\beta) > \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \ln v_\alpha$. Since this is the opposite of (1), the equivalence of (1) and (2') follows.

(2') $\implies$ (2):

We obtain (2) from (2') by multiplying with v_β and replacing $v_\beta \tau$ with $-\tau$.

(2) $\implies$ (3):

This follows by setting $v^* := v_\beta$.

(3) $\implies$ (2'):

We have that $v^* \geq v_\beta > 0$, and thus, we may divide the inequality in (3) by v^* to obtain

$$\exists \tau' \in \mathbb{R}^n, \forall \alpha \in \mathcal{A}\colon \ln\left(\frac{v^*}{v_\alpha}\right) \leq (\beta - \alpha)^T \tau',$$

where $\tau' = \tau/v^*$. Note that the left-hand side of the inequality is monotonous in v^*, and hence,

$$\ln\left(\frac{v_\beta}{v_\alpha}\right) \leq \ln\left(\frac{v^*}{v_\alpha}\right) \leq (\beta-\alpha)^T\tau'.$$

We further replace τ' with $-\tau'$ to obtain (2'). □

3.2 Extreme Rays of the SAGE-Cone

As a first application of our description of the dual cone, we prove the following theorem. Note that the first statement in the theorem was already observed by Wang for SONC polynomials [Wan18b] and by Murray, Chandrasekaran an Wierman for SAGE exponentials [MCW21a, Theorem 4].

Theorem 3.2.1. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite . For every $f \in C_{\mathrm{SAGE}}(\mathcal{A})$, the following statements hold:*

1. *f can be written as a sum of nonnegative circuit exponentials whose supports are contained in* $\operatorname{supp} f$.
2. *f can be written as a sum of nonnegative circuit exponentials supported on reduced circuits in $C_{\mathrm{SAGE}}(\mathcal{A})$.*

Note that in statement (2) of this Theorem, the supports of the reduced circuits do not need to be contained in the support of f. The following example shows a situation in which this phenomenon happens.

Example 3.2.2. Let $\mathcal{A} := \{0, 3/4, 4\}$. Consider the nonnegative circuit exponential $f = 1 - 4\cdot 3^{-3/4}e^x + e^{4x}$. Its support $(\{0,4\},1)$ is not reduced with respect to $\mathcal{A}$, and indeed, we can write f as sum

$$f = \left(\frac{2}{3} - 4\cdot 3^{-3/4}e^x + \frac{2}{3}\sqrt{3}e^{2x}\right) + \left(\frac{1}{3} - \frac{2}{3}\sqrt{3}e^{2x} + e^{4x}\right)$$

of nonnegative circuit exponentials, whose supports $(\{0,2\},1)$ and $(\{0,4\},2)$ are reduced. Note that the coefficient of e^{2x} cancels in the sum.

Proof of Theorem 3.2.1. By Lemma 3.1.6 and part (c) of Theorem 3.1.5, the dual of the SAGE-cone is

$$C_{\mathrm{SAGE}}(\mathcal{A})^* = \bigcap_{(A,\beta)\in I(\mathcal{A})} C_{\mathrm{AGE}}(A,\beta)^*. \tag{3.6}$$

Let $f \in C_{\mathrm{SAGE}}(\mathcal{A})$ and assume that the support of f is given by $\mathcal{A}' \subseteq \mathcal{A}$. By Theorem 2.4.8, $f \in C_{\mathrm{SAGE}}(\mathcal{A}')$. Apply (3.6) on the sub-cone $C_{\mathrm{SAGE}}(\mathcal{A}')$ and dualize that identity. Using that $C_{\mathrm{SAGE}}(\mathcal{A}')^{**} = C_{\mathrm{SAGE}}(\mathcal{A}')$ (because the cone is closed) then yields

$$f \in \sum_{(A,\beta)\in I(\mathcal{A})} C_{\mathrm{AGE}}(A,\beta).$$

This shows part (1).

Part (2) then follows from part (d) of Theorem 3.1.5. Note that in this case we cannot restrict the sets of exponents to $\mathcal{A}'$ as it depends on the choice of $\mathcal{A}$ whether a circuit is reduced or not. □

Remark 3.2.3. If we demand $\operatorname{supp}(f) = \mathcal{A}$, we obtain the same statement about the support in (2) as in (1).

Our next application of our description of the dual cone is a precise characterization of the extreme rays of $C_{\text{SAGE}}(\mathcal{A})$. This sharpens the result in [Wan18a] and [MCW21a, Theorem 4], where the necessary condition is that every extreme ray of the SAGE-cone is supported on a single coordinate or on a circuit. The essential concept for this characterization is provided by the reduced circuits.

Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite. For a circuit $(A, \beta) \in I(\mathcal{A})$, i.e., $A \subseteq \mathcal{A}$ affinely independent and $\beta \in \mathcal{A} \cap \operatorname{relint}\operatorname{conv}(A)$, write shortly $\lambda = \lambda(A, \beta)$. Let

$$E(A, \beta) := \left\{ \sum_{\alpha \in A} c_\alpha e^{\langle \alpha, x \rangle} - \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} e^{\langle \beta, x \rangle} \,\middle|\, c \in \mathbb{R}^A_{>0} \right\},$$

and for $\beta \in \mathcal{A}$ let

$$E_1(\beta) := \mathbb{R}_+ \cdot e^{\langle \beta, x \rangle}.$$

The sets $E(A, \beta)$ for circuits $(A, \beta) \in \mathcal{I}(\mathcal{A})$ contain those nonnegative circuit exponentials for which the inequality from Theorem 2.4.5 on the circuit number holds with equality. The sets $E_1(\beta), \beta \in \mathcal{A}$ provide the special case for circuits supported on a single element.

Theorem 3.2.4. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be a finite set and write $\lambda = \lambda(A, \beta)$. The set $\mathcal{E}(\mathcal{A})$ of extreme rays of the cone of SAGE exponentials with support in $\mathcal{A}$ is*

$$\begin{aligned}
\mathcal{E}(\mathcal{A}) &= \bigcup_{\substack{(A,\beta) \in I(\mathcal{A}), \\ r(A,\beta)=0,\, |A|>1}} E(A, \beta) \;\cup\; \bigcup_{\beta \in \mathcal{A}} E_1(\beta) \\
&= \bigcup_{\substack{(A,\beta) \in I(\mathcal{A}), \\ r(A,\beta)=0,\, |A|>1}} \left\{ \sum_{\alpha \in A} c_\alpha e^{\langle \alpha, x \rangle} - \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} e^{\langle \beta, x \rangle} \,\middle|\, c \in \mathbb{R}^A_{>0} \right\} \\
&\quad \cup \bigcup_{\beta \in \mathcal{A}} \left\{ c e^{\langle \beta, x \rangle} \,\middle|\, c \in \mathbb{R}_+ \right\}.
\end{aligned}$$

This statement was later also partially observed by Forsgård and de Wolff [FW19]. As mentioned in Chapter 2, in their language, circuits are minimally supported elements $\nu \in \ker(\mathcal{A})$ with a single negative entry, i.e., multiples of the vector of barycentric coordinates with respect to the inner term. As we refer to the statement in the following chapters again, we denote for *their* definition of a circuit $\nu^+ := \{\alpha \in \operatorname{supp}(\nu) : \nu_\alpha > 0\}$ and by ν^- the single support point with negative entry and frame the statement as follows:

Proposition 3.2.5 ([FW19, Theorem 3.2])**.** *A vector ν is an edge generator of the SAGE-cone (i.e., it is contained in an extremal ray) if and only if*

$$\mathcal{A} \cap \operatorname{relint}\operatorname{conv} \nu^+ = \{\nu^-\}.$$

This, in particular, captures our concept of reduced circuits, Definition 3.1.3: Proposition 3.2.5 states that a vector ν is an edge generator of the SAGE-cone if and only if it is a reduced circuit.

We have a look at the extreme rays from Example 3.1.4.

Example 3.2.6. Let $\mathcal{A} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$. The set of reduced circuits $(A, \beta) \in \mathcal{I}(\mathcal{A})$ with $|A| > 1$ is pictured in Figure 3.3. Extreme rays

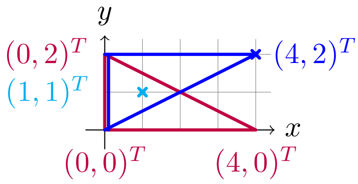

FIGURE 3.3: Reduced circuits of $\mathcal{A}$ from Example 3.2.6.

of the SAGE-cone with fixed support set $\mathcal{A}$ are

$$
\begin{aligned}
&\left\{ c_0 + c_4 e^{4x} + c_2 e^{2y} - 4{c_0}^{1/4} 4{c_4}^{1/4} 2{c_2}^{1/2} e^{x+y} \,\middle|\, c_\alpha > 0 \,\forall\, \alpha \in \{0,4,2\} \right\} \\
\cup &\left\{ c_0 + c_{(4,2)} e^{4x+2y} + c_2 e^{2y} - 2{c_0}^{1/2} 4{c_{(4,2)}}^{1/4} 4{c_2}^{1/4} e^{x+y} \,\middle|\, c_\alpha > 0 \,\forall\, \alpha \in \{0,(4,2),2\} \right\} \\
\cup \;& \mathbb{R}_+ \cup \mathbb{R}_+ e^{4x} \cup \mathbb{R}_+ e^{2y} \cup \mathbb{R}_+ e^{4x+2y} \cup \mathbb{R}_+ e^{x+y}.
\end{aligned}
$$

For the cone of AGE exponentials, every circuit supports a family of extreme rays in the AGE-cone. This is not correct for the full SAGE-cone anymore, as shown by the following example.

Example 3.2.7. Let $\mathcal{A} = \{0,1,2,4\}, \mathcal{B} = \emptyset$ and $f := 1 - 4 \cdot 3^{-3/4} e^x + e^{4x}$ be a nonnegative circuit exponential with non-reduced support in terms of $\mathcal{A}$. We write f as a sum

$$f = \left(1 - 2 \cdot 3^{1/4} e^x + \sqrt{3} e^{2x}\right) + \left(\frac{2}{3} 3^{1/4} e^x - \sqrt{3} e^{(2x} + e^{4x}\right)$$

of circuit exponentials, whose supports $\{0,1,2\}$ and $\{1,2,4\}$ are reduced.

As the cone of SONC polynomials can be deduced from the cone of SAGE exponentials, we obtain a corollary for the special case of SONC polynomials.

Corollary 3.2.8. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{N}^n$ be a finite set and write shortly $\lambda = \lambda(A,\beta)$. The set $\mathcal{E}(\mathcal{A})$ of extreme rays of the cone of SONC-polynomials with support in $\mathcal{A}$ is*

$$
\begin{aligned}
\mathcal{E}(\mathcal{A}) = & \bigcup_{\substack{(A,\beta) \in I(\mathcal{A}), \\ r(A,\beta)=0,\, |A|>1}} \left\{ \sum_{\alpha \in A} c_\alpha x^\alpha - \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} x^\beta \,\middle|\, c \in \mathbb{R}^A_{>0} \right\} \\
& \cup \bigcup_{\substack{(A,\beta) \in I(\mathcal{A} \cap (2\mathbb{N})^n, \mathcal{A}), \\ r(A,\beta)=0,\, |A|>1,\, \beta \in \mathcal{A} \setminus (2\mathbb{N})^n}} \left\{ \sum_{\alpha \in A} c_\alpha x^\alpha + \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} x^\beta \,\middle|\, c \in \mathbb{R}^A_{>0} \right\} \\
& \cup \bigcup_{\beta \in \mathcal{A} \cap (2\mathbb{N})^n} \mathbb{R}_+ \cdot x^\beta.
\end{aligned}
$$

For the proof of Theorem 3.2.4, we use a variant of Hölder's inequality.

Theorem 3.2.9 ([HLP52, Theorem 11, p. 22])**.** *Let $n, m \in \mathbb{N}$. Let $(a_{ij}) \in \mathbb{R}^{n \times m}$ be a matrix and let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_{>0}$ with $\sum_{i=1}^n \lambda_i = 1$. Then*

$$\sum_{j=1}^m \prod_{i=1}^n a_{ij}^{\lambda_i} \leq \prod_{i=1}^n \left(\sum_{j=1}^m a_{ij}\right)^{\lambda_i}$$

and equality holds if and only if either (1) for some i, $a_{i1} = \cdots = a_{im} = 0$, or (2) the matrix (a_{ij}) has rank 1.

Note that in case (1) both sides of the inequality are zero.

Proof of Theorem 3.2.4. By Theorem 3.2.1, every exponential sum $f \in C_{\text{SAGE}}(\mathcal{A})$ can be written as a sum of nonnegative circuit exponentials supported on reduced circuits. Hence, it suffices to show the following two statements:

(a) Every nonnegative circuit exponential can be written as a sum of nonnegative circuit exponentials with the same support whose circuit condition is satisfied with equality.

(b) Every function in $\mathcal{E}(\mathcal{A})$ is indeed an extreme ray, i.e., it cannot be written as a sum of other AGE exponentials supported on $\mathcal{A}$.

(a) Let f be a nonnegative circuit exponential supported on the circuit (A, β), whose coefficients are denoted by $(c_\alpha)_{\alpha \in A}$ and c_β.

Then f can be written as the sum of nonnegative circuit exponentials with the same support whose inner coefficient , i.e., coefficient corresponding to the inner exponent, equals the negative of the circuit number and of some function $de^{\langle \beta, x \rangle}$ for $d > 0$. The latter is contained in $E_1(\beta)$.

(b) Let $f \in \mathcal{E}(\mathcal{A})$ with coefficients $(c_\alpha)_{\alpha \in \mathcal{A}}$. Assume that f can be decomposed into $f = \sum_{i=1}^k f_i$ with AGE exponentials $f_1, \ldots, f_k \in \mathbb{R}^{\mathcal{A}}$. Denote the coefficients of f_i by $(c_\alpha^{(i)})_{\alpha \in \mathcal{A}}$.

Set $\tilde{A} := \bigcup_i \operatorname{supp}_{\mathrm{e}}(f_i) = \{\alpha \in \mathcal{A} \,\big|\, \exists i \colon c_\alpha^{(i)} \neq 0\}$. We claim that $\tilde{A} \subseteq \operatorname{conv} \operatorname{supp}(f)$. To show this, we consider a vertex $\tilde{\alpha}$ of $\operatorname{conv} \tilde{A}$. Since $\tilde{\alpha}$ must be an outer exponent of each f_i with $c_{\tilde{\alpha}}^{(i)} \neq 0$, we have $c_{\tilde{\alpha}}^{(i)} \geq 0$ for all i. It follows that $\sum_{i=1}^k c_{\tilde{\alpha}}^{(i)} > 0$, and thus, $\tilde{\alpha} \in \operatorname{supp}(f)$. As this holds for every vertex of $\operatorname{conv} \tilde{A}$, we obtain that $\tilde{A} \subseteq \operatorname{conv} \operatorname{supp}(f)$.

Next, we fix $\beta \in \mathcal{A}$ and a circuit $(A, \beta) \in \mathcal{I}(\mathcal{A})$ and distinguish two cases depending on whether $f \in E_1(\beta)$ or $f \in E(A, \beta)$.

(1) *Case $f \in E_1(\beta)$, $\beta \in \mathcal{A}$*: In this case, $\operatorname{supp}(f_i) = \{\beta\}$ for each i. Thus, each f_i is a multiple of $e^{\langle \beta, x \rangle}$, and thus, a multiple of f.

(2) *Case $f \in E(A, \beta)$ for $(A, \beta) \in I(\mathcal{A})$ with $r(A, \beta) = 0$ and $|A| > 1$*: In this case, our initial considerations imply that $\bigcup_i \operatorname{supp}(f_i) \subseteq \operatorname{conv}(A)$. Since (A, β) is reduced we can also conclude that $\bigcup_i \operatorname{supp}(f_i) \subseteq A \cup \{\beta\}$. Hence, each f_i is of the form

$$f_i = \sum_{\alpha \in A} c_\alpha^{(i)} e^{\langle \alpha, x \rangle} + c_\beta^{(i)} e^{\langle \beta, x \rangle}. \tag{3.7}$$

It follows that $c_\alpha^{(i)} \geq 0$ for all i and $\alpha \in A$, because otherwise the f_i cannot be nonnegative. Moreover, as each f_i is an AGE exponential,

$$-c_\beta^{(i)} \leq \prod_{\alpha \in A} \left(\frac{c_\alpha^{(i)}}{\lambda_\alpha} \right)^{\lambda_\alpha} \tag{3.8}$$

for all i, where again, we write $\lambda = \lambda(A, \beta)$.

In the next step, we derive

$$\begin{aligned}-c_\beta = -\sum_{i=1}^k c_\beta^{(i)} &\overset{\text{(a)}}{\leq} \sum_{i=1}^k \prod_{\alpha\in A}\left(\frac{c_\alpha^{(i)}}{\lambda_\alpha}\right)^{\lambda_\alpha} \overset{\text{(b)}}{\leq} \prod_{\alpha\in A}\left(\sum_{i=1}^k \frac{c_\alpha^{(i)}}{\lambda_\alpha}\right)^{\lambda_\alpha} \\ &\overset{\text{(c)}}{=} \prod_{\alpha\in A}\left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} = -c_\beta,\end{aligned} \tag{3.9}$$

where in (a) we use (3.8), (b) follows from Theorem 3.2.9, and (c) uses the fact that $\sum_{i=1}^k c_\alpha^{(i)} = c_\alpha$. Moreover, by Theorem 3.2.9, equality in (b) implies that either (1) there exists an $\alpha \in A$ such that $c_\alpha^{(i)}$ vanishes for all i, or (2) the $|A| \times k$ matrix with entries $c_\alpha^{(i)}/\lambda_\alpha$ has rank one. However, (1) would imply that $c_\beta = 0$ which is impossible, thus, we are in case (2). Hence, there exist scalars $\epsilon_1, \ldots, \epsilon_k \geq 0$ such that $c_\alpha^{(i)} = \epsilon_i c_\alpha$ for all i and all $\alpha \in A$. Further, equality in (a) implies that

$$-c_\beta^{(i)} = \prod_{\alpha\in A}\left(\frac{c_\alpha^{(i)}}{\lambda_\alpha}\right)^{\lambda_\alpha} = \prod_{\alpha\in A}\left(\epsilon_i\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} = -\epsilon_i c_\beta.$$

By (3.7), it follows that every f_i is of the form $f_i = \epsilon_i f$. Thus, $\epsilon_i > 0$ for all i. But this implies that $c_\beta^{(i)} = \epsilon_i c_\beta < 0$. Hence, since the f_i are AGE exponentials, they are all multiples of f. $\square$

Chapter 4

Global Optimization via the SAGE-Cone and its Dual

In this Chapter, we present two optimization methods over the SAGE-cone and its dual. To do so, we start in Section 4.1 by introducing a tool named the *Signed SAGE-cone* which we use for optimization purposes. This cone contains SAGE exponentials with fixed signs of the coefficients, see Definition 4.1.1. This is a common approach in optimization and this restriction to certain sign patterns allows us to use slightly different statements for the SAGE-cone and its dual which are particularly helpful for optimization.

In Section 4.2, we present a method exploiting symmetries in a given exponential sum. We assume G-invariance of an exponential for a group G and start by proving a symmetry-adapted decomposition result, namely Theorem 4.2.1. Bulding upon this decomposition result, we then show that the complexity of the relative entropy programs in Theorem 3.1.1 can be substantially reduced by using this adaption, see Theorem 4.2.3. For computational reasons, we exploit symmetries in this setting even further and obtain a complex but efficient variant of Theorem 3.1.1 (2), which substantially reduces the size of the corresponding relative entropy program and the run time solving it. From this we can see that the improvement is dependent on the orbit structure of the group action.

We close Section 4.2 by presenting some numerical experiments and comparing the computational gain of the symmetrized version of the relative entropy program to that of the conventional one. Whenever we have a strong symmetric structure, this causes the variable, equation and inequality count to decrease a lot, yielding a substantial reduction in computation time. We also present cases where the conventional computation fails but we do obtain a solution using the symmetrized version of the relative entropy program.

In Section 4.3, we show that the dual cone of SAGE exponentials provides a linear programming method to approximate the global optimization problem (1.1):

$$f^* = \inf\{f(x) \ : \ x \in \mathbb{R}^n\} \ = \ \sup\{\lambda \in \mathbb{R} \ : \ f - \lambda \text{ is nonnegative on } \mathbb{R}^n\}.$$

This result is based on the fact that the dual cone of AGE exponentials is contained in the primal cone. Moreover, certain exponential sums whose coefficients can be constructed using the dual SAGE-cone (but do not neccessarily *lie in* the dual SAGE-cone) are contained in the primal SAGE-cone, and, hence, nonnegative; see Proposition 4.3.1. This yields a linear program certifying nonnegativity of a given function, as explained in Proposition 4.3.4. Building upon this certificate, we consider a relaxation of a global optimization problem, (4.24), and provide two optimization programs (LP_{A^+}) and (LP_{A^-}). We then state the main theorem of this section, Theorem 4.3.8, which explains how to obtain a lower bound on the global optimal value of a given function using this developed approach.

4.1 The Signed SAGE-Cone and its Dual

Motivated by this optimization approach, we make a restriction when investigating the SAGE-cone. For a fixed exponential sum f with support $\mathcal{A}$, which we intend to minimize, we have additional information on the signs of the coefficients of f. Using this, we obtain a decomposition

$$\mathcal{A} \;=\; A^+ \cup A^- \tag{4.1}$$

with disjoint sets $\emptyset \neq A^+ \subseteq \mathcal{A}$, corresponding to the set of nonnegative coefficients, and $A^- \subseteq \mathcal{A}$, corresponding to the remaining negative coefficients.

Thus, we represent exponential sums in this case as

$$f = \sum_{\alpha \in A^+} c_\alpha e^{\langle \alpha, x \rangle} + \sum_{\alpha \in A^-} c_\alpha e^{\langle \alpha, x \rangle} \tag{4.2}$$

with $c_\alpha \geq 0$ for all $\alpha \in A^+$ and $c_\alpha < 0$ for all $\alpha \in A^-$, i.e., we explicitely distinguish between nonnegative terms and negative terms. In case we want to emphasize the underlying function f, we sometimes also write $A^+(f)$ and $A^-(f)$.

If we minimize a given exponential sum f using the SAGE-cone approach, we restrict to exponential sums respecting the sign-pattern indicated by f. This is the common, tractable approach used by various authors, e.g. in [DIW19; IW16b; MCW21a; MCW21b]; it motivates the following definition.

Definition 4.1.1 (Signed SAGE-Cone)**.** Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be a finite set joint with a decomposition $\mathcal{A} = A^+ \cup A^-$ in the sense of (4.1). Then, the *signed SAGE-cone* $\mathcal{S}_{A^+,A^-}$ is the cone of all exponential sums that can be written as a sum of AGE exponentials of the form (4.2) or as elements $c_\alpha e^{\langle \alpha, x \rangle}$ with $\alpha \in A^+$ and $c_\alpha > 0$. We denote the special case $A^- = \{\beta\}$ by $\mathcal{S}_{A^+,\beta}$.

Following the convention from previous chapters, elements $c_\alpha e^{\langle \alpha, x \rangle}$ with $\alpha \in A^+$ and $c_\alpha > 0$ are formally supported on circuits $(\{\alpha\}, \alpha)$. Note that we formally include elements with 0-coefficients in the set A^+. This ensures that for $A^+ \subseteq A'$, we have $\mathcal{S}_{A^+,\beta} \subseteq \mathcal{S}_{A',\beta}$. In reality, for $\mathcal{A} = A^+ \cup A^-$, we mostly assume $A^+ \cup A^- = \operatorname{supp}(f)$ whenever considering the signed SAGE-cone and elements therein.

In fact, by using a generalization of the circuit number and the subsequent notation, we can refine the representation of $\mathcal{S}_{A^+,\beta}$. To do so, we use a refined definition of the polytope $\Lambda(\mathcal{A}, \beta)$ from (3.1), namely,

$$\Lambda(A^+, \beta) \;=\; \left\{ \lambda \in \mathbb{R}_+^{A^+} \;:\; \sum_{\alpha \in A^+} \lambda_\alpha \alpha = \beta, \; \sum_{\alpha \in A^+} \lambda_\alpha = 1 \right\} \tag{4.3}$$

for $\beta \in A^-$.

Using this definition, we may restate one part of Theorem 3.1.1 as follows:

Theorem 4.1.2 (Essentially Theorem 3.1.1)**.** *Let $\mathcal{A} = A^+ \cup \{\beta\}$ be defined as in* (4.2)*. The signed cone of AGE exponentials $\mathcal{S}_{A^+,\beta}$ is the set*

$$\left\{ f = \sum_{\alpha \in A^+} c_\alpha e^{\langle \alpha, x\rangle} + c_\beta e^{\langle \beta, x\rangle} \ : \ c_\alpha \geq 0 \text{ for } \alpha \in A^+, c_\beta < 0 \text{ and} \right.$$
$$\left. \exists\, \lambda \in \Lambda(A^+, \beta) \text{ s.t. } \prod_{\alpha \in A^+ : \lambda_\alpha > 0} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} \geq -c_\beta \right\}.$$

Note here that nonnegativity of an AGE exponential can be certified directly by using the generalized circuit number above. There is no need to decompose it into a sum of nonnegative circuit exponentials.

We can adapt Theorem 2.4.8 to our setting, which basically states that an exponential sum in the SAGE-cone supported on $\mathcal{A}$ can be decomposed into a sum of AGE exponentials supported on $A^+ \cup \{\beta\}$, $\beta \in A^-$, i.e., the decomposition only uses the support $\mathcal{A}$ and there is only one summand per element in A^-.

Theorem 4.1.3 ([MCW21a, Theorem 2])**.** *Let $f \in \mathcal{S}_{A^+,A^-}$ with a vector of coefficients c and $\operatorname{supp}(f) \cap A^+ = A^+$. Let $A^- \neq \emptyset$. Then, there exist exponential sums $\{f^{(\beta)} : \beta \in A^-\}$ with coefficient vectors $\{c^{(\beta)} : \beta \in A^-\}$ satisfying*

1. $c = \sum_{\beta \in A^-} c^{(\beta)}$,
2. $f^{(\beta)} \in \mathcal{S}_{A+,\beta}$ *and*
3. $c^{(\beta)}_{\beta'} = 0$ *for all $\beta' \neq \beta$ in A^-.*

This leads to the following variant of Theorem 3.1.1 (2) for the signed SAGE-cone.

Theorem 4.1.4. *Let $\mathcal{A} = A^+ \cup A^-$ as defined before and f be an exponential sum respecting the sign pattern* (4.2)*. The function f is contained in $\mathcal{S}_{A^+,A^-}$ if and only if for every $\beta \in A^-$ there exist $c^{(\beta)} \in \mathbb{R}_+^{A^+}$ and $\nu^{(\beta)} \in \mathbb{R}_+^{A^+}$ such that*

$$\sum_{\alpha \in A^+} \nu_\alpha^{(\beta)} (\alpha - \beta) \ = \ 0 \quad \text{for every } \beta \in A^-, \tag{4.4}$$
$$D(\nu^{(\beta)}, e \cdot c^{(\beta)}) \ \leq \ c_\beta \quad \text{for every } \beta \in A^- \text{ and} \tag{4.5}$$
$$\sum_{\beta \in A^-} c_\alpha^{(\beta)} \ \leq \ c_\alpha \quad \text{for every } \alpha \in A^+. \tag{4.6}$$

In what follows, we collect statements for the dual signed SAGE-cone that we need in Subsection 4.3.1 to show containment of the dual in the primal SAGE-cone and in Subsection 4.3.2 to obtain a fast linear approximation for global optimization.

Due to our goals in this chapter, we here discuss duality with respect to the signed SAGE-cone. However, everything generalizes to the full SAGE-cone immediately as we only combine statements from the previous chapters.

We follow the definitions used for the full SAGE-cone.

Definition 4.1.5 (The Dual Signed SAGE-Cone)**.** For an exponential sum f of the form (4.2) with coefficient vectors $c \in \mathbb{R}^{\mathcal{A}}$, we consider the natural duality pairing

$$v(f) \ = \ \sum_{\alpha \in A^+} v_\alpha c_\alpha + \sum_{\alpha \in A^-} v_\alpha c_\alpha \in \ \mathbb{R},$$

where $v(\cdot) \in (\mathbb{R}^{\mathcal{A}})^*$ is canonically identified with its (dual) coefficient vector, hence, we consider v to be an element in $\mathbb{R}^{\mathcal{A}}$. Using this definition, the *dual signed SAGE-cone* is defined as the set

$$\mathcal{S}^*_{A^+,A^-} \;=\; \left\{v \in \mathbb{R}^{\mathcal{A}} : v(f) \geq 0 \text{ for all } f \in \mathcal{S}_{A^+,A^-}\right\}.$$

For brevity, we refer to this cone simply as the *dual SAGE-cone* in this chapter.

The following theorem provides two representations of this cone. We need the first one to show containment of the dual cone of AGE exponentials in the primal one, and the second representation to obtain a linear program approximating the solution of our global optimization problem (1.1).

Theorem 4.1.6. *Let $\mathcal{A} = A^+ \cup A^-$ as in* (4.1)*. The following sets are equal:*

1. $\mathcal{S}^*_{A^+,A^-}$,
2. $\left\{v \in \mathbb{R}^{\mathcal{A}} : \begin{array}{l} \text{for all } \alpha \in A^+, v_\alpha \geq 0; \text{ and for all } \beta \in A^-, \text{ for all} \\ \lambda \in \Lambda(A^+,\beta)\ ,\ \ln(|v_\beta|) \leq \sum\limits_{\alpha \in A^+} \lambda_\alpha \ln(v_\alpha) \end{array}\right\}$,
3. $\mathrm{cl}\left\{v \in \mathbb{R}^{\mathcal{A}}_{\neq 0} : \begin{array}{l} \text{for all } \alpha \in A^+, v_\alpha > 0; \text{ and for all } \beta \in A^- \text{ there exists} \\ \tau \in \mathbb{R}^n, \ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T\tau \text{ for all } \alpha \in A^+ \end{array}\right\}$

Note that, whenever $A^- = \emptyset$, the cone $\mathcal{S}^*_{A^+,A^-}$ equals the set $\mathbb{R}^{A^+}_+$. Hence, the interesting case is whenever $A^- \neq \emptyset$.

In particular, there are two differences to earlier versions of the same statement: First, the conditions " $\ln(|v_\beta|) \leq \sum_{\alpha \in A^+} \lambda_\alpha \ln(v_\alpha)$" and "$\ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T\tau$" only need to hold for every $\beta \in A^-$ and second, the condition "$v_\alpha \geq 0$" only needs to hold for every $\alpha \in A^+$. The reason for the first difference becomes clear with the following statements. The second one is due to the fact that no exponential $c_\alpha e^{\langle \alpha, x\rangle}$ with $\alpha \in A^-$ and *positive* coefficient c_α is contained in $\mathcal{S}_{A^+,A^-}$.

We obtain the following representation of the SAGE-cone and its dual.

Corollary 4.1.7. *Let $\emptyset \neq \mathcal{A} = A^+ \cup A^- \subseteq \mathbb{R}^n$ as in* (4.1)*. The following statements hold.*

1. *The SAGE-cone is the Minkowski sum*

$$\mathcal{S}_{A^+,A^-} = \sum_{\beta \in A^-} \mathcal{S}_{A^+,\beta} + \sum_{\alpha \in A^+} \mathbb{R}_+ \cdot e^{\langle \alpha, x\rangle}$$

2. *The dual SAGE-cone is the set*

$$\mathcal{S}^*_{A^+,A^-} = \bigcap_{\beta \in A^-} \mathcal{S}^*_{A^+,\beta} \cap (\mathbb{R}^{A^+}_+ \times \mathbb{R}^{A^-})$$

Note that the second statement involves a slight abuse of notation: Formally, we need to consider the lifted cones $\{v \in \mathbb{R}^{A^+ \cup A^-} : v_{|A^+ \cup \{\beta\}} \in \mathcal{S}^*_{A^+,\beta}\}$. As $c_\alpha = 0$ for all $\alpha \in A^- \setminus \{\beta\}$ in every exponential sum f with $c \in \mathcal{S}_{A^+,\beta}$, there are no restrictions on v_α for $\alpha \in A^- \setminus \{\beta\}$.

Proof of Corollary 4.1.7. The first statement follows with Theorem 4.1.3. For the second statement, note that Minkowski sum and intersection are dual operations; see, e.g., [Sch14, Theorem 1.6.3]. □

In order to finally prove Theorem 4.1.6, we need another statement, which essentially combines Lemma 3.1.6 and a part of the proof of Proposition 3.1.8 — namely, Remark 3.1.10.

Lemma 4.1.8. *For $\beta \in A^- \neq \emptyset$, the dual of the SAGE-cone $\mathcal{S}^*_{A^+,\beta}$ consists of the closure of the set of those $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$ where the following equivalent conditions hold:*

1. $\ln(|v_\beta|) \leq \sum_{\alpha \in A^+} \lambda_\alpha \ln(v_\alpha)$ *for all* $\lambda \in \Lambda(A^+, \beta)$.
2. *There exists* $\tau \in \mathbb{R}^n$ *such that for all* $\alpha \in A^+$: $\ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T \tau$.

Proof of Theorem 4.1.6. In the case $A^- = \emptyset$, $\mathcal{S}^*_{A^+,\beta}$ only contains sums of nonnegative exponentials with nonnegative coefficient, and the equality of the sets $(1) - (3)$ is clear. For $A^- \neq \emptyset$, the statement follows by Corollary 4.1.7 and Lemma 4.1.8 . Namely, the first representation can be deduced from (1), and the second one from (2). □

4.2 Symmetry Reduction in AM/GM-Based Optimization

We start by examining situations involving symmetries. We provide a specific way of writing SAGE exponentials in the presence of a group symmetry. We study how to characterize and to decide whether a G-symmetric exponential sum is contained in the SAGE-cone with reduced relative entropy programs and provide experimental results of an implementation of the symmetry reduction techniques.

4.2.1 Orbit Decompositions of Symmetric Exponential Sums

In this subsection, we provide a structural result on the decomposition of symmetric exponential sums in the SAGE-cone as sums of orbits of (non-symmetric) AGE exponentials.

Let G be a finite group acting linearly on $\mathbb{R}^n$ on the left, namely, we have a group homomorphism

$$\begin{array}{cccc} \varphi & : & G & \to & \mathrm{GL}_n(\mathbb{R}) \\ & & \sigma & \mapsto & \varphi(\sigma) \end{array} .$$

For $\sigma \in G$ and $x \in \mathbb{R}^n$, we denote by $\sigma \cdot x$ the image of x through $\varphi(\sigma)$. In order to get a left action on the set of functions defined on $\mathbb{R}^n$, we need to take

$$(\sigma * f)(x) = f(\sigma^{-1} \cdot x) = f(\varphi(\sigma^{-1})(x)). \tag{4.7}$$

For an exponential sum $f(x) = \sum_\alpha c_\alpha e^{\langle \alpha, x \rangle}$, we see an exponent vector α as an element of the dual space of $\mathbb{R}^n$. Then, the dual action of G on the exponent vectors is given by

$$\sigma \perp \alpha := \varphi(\sigma^{-1})^\#(\alpha).$$

Recall that $A^\#$ denotes the adjoint operator of A. Note that this is a left action as well. Therefore, even if the exponents and the variables lie in isomorphic spaces, the actions of G on these spaces are different and dual to each other and satisfy

$$\langle \alpha, \sigma \cdot x \rangle = \langle \alpha, \varphi(\sigma)(x) \rangle = \langle \varphi(\sigma)^\#(\alpha), x \rangle = \langle \sigma^{-1} \perp \alpha, x \rangle.$$

Furthermore, for an exponential sum f,

$$(\sigma * f)(x) = f(\sigma^{-1} \cdot x) = \sum_{\alpha} c_\alpha e^{\langle \alpha, \sigma^{-1} \cdot x \rangle} = \sum_{\alpha} c_\alpha e^{\langle \sigma \perp \alpha, x \rangle}. \tag{4.8}$$

From now on, in order to keep notations as light as possible, with a slight abuse of notation, we write $\sigma(x) = \sigma \cdot x$ for the action on the variables, $\sigma f = \sigma * f$ for the action on functions, and $\sigma(\alpha) = \sigma \perp \alpha$ for the dual action. Even if the actions are different, the context should clarify the correspondence.

Recall from Chapter 2 that for a set $\mathcal{S} \subseteq \mathbb{R}^n$ of exponent vectors, the *orbit of* $\mathcal{S}$ under G is

$$G \cdot \mathcal{S} = \{\sigma(s) : s \in \mathcal{S},\, \sigma \in G\}.$$

Moreover, recall that a subset $\hat{\mathcal{S}} \subseteq \mathcal{S}$ is called a set of orbit representatives for $\mathcal{S}$ if $\hat{\mathcal{S}}$ is an inclusion-minimal set with $(G \cdot \hat{\mathcal{S}}) = \mathcal{S}$, and the stabilizer of an exponent vector β is denoted by $\operatorname{Stab} \beta := \{\sigma \in G \;:\; \sigma(\beta) = \beta\}$.

In the following statements, we consider G-invariant exponential sums f. Here , it is convenient to write f in the form

$$f = \sum_{\alpha \in A^+} c_\alpha e^{\langle \alpha, x \rangle} + \sum_{\alpha \in A^-} c_\alpha e^{\langle \alpha, x \rangle} \tag{4.9}$$

with $c_\alpha > 0$ for $\alpha \in A^-$ and $c_\alpha < 0$ for $\alpha \in A^-$, i.e., f is an element of the signed SAGE-cone $\mathcal{S}_{A^+,A^-}$ introduced in the previous section. As already mentioned, in this notation, the overall support set of f is $A^+ \cup A^- = \mathcal{A}$. For the following statements, we assume $\emptyset \neq A^-$: In the case $\emptyset = A^-$, the set $\mathcal{S}_{A^+,A^-}$ coincides with $\mathbb{R}^{A^+}$ and the membership problem of the signed SAGE-cone is trivial.

Theorem 4.2.1. *Let f be a G-invariant exponential sum of the form* (4.9) *and $\emptyset \neq \hat{A}^-$ be a set of orbit representatives for A^-. Then, $f \in \mathcal{S}_{A^+,A^-}$ if and only if for every $\hat{\beta} \in \hat{A}^-$, there exists an AGE exponential $h_{\hat{\beta}} \in \mathcal{S}_{A^+,\beta}$ such that*

$$f = \sum_{\hat{\beta} \in \hat{A}^-} \sum_{\rho \in G/\operatorname{Stab}(\hat{\beta})} \rho h_{\hat{\beta}}. \tag{4.10}$$

The functions $h_{\hat{\beta}}$ can be chosen invariant under the action of $\operatorname{Stab}(\hat{\beta})$*.*

Here, $\rho \in G/\operatorname{Stab}(\hat{\beta})$ shortly denotes that ρ runs over a set of representatives of the left quotient space $G/\operatorname{Stab}(\hat{\beta})$, defined through left cosets $\{\sigma \operatorname{Stab}(\hat{\beta}) \;:\; \sigma \in G\}$. We also use the right quotient space, denoted by $\operatorname{Stab}(\hat{\beta})\backslash G$, further below.

Example 4.2.2. For the symmetric group $\mathcal{S}_n$, every $\sigma \in G$ acts on $x = (x_1, \ldots, x_n)$ through $\sigma(x) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$. On an exponential sum $f(x) = \sum_\alpha c_\alpha e^{\langle \alpha, x \rangle}$, this induces the action

$$\sigma f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Since $\sigma f(x) = \sum_\alpha c_\alpha e^{\langle x, (\alpha_{\sigma^{-1}(1)}, \ldots, \alpha_{\sigma^{-1}(n)}) \rangle}$, the induced action on exponent vectors α is

$$\sigma(\alpha) = (\alpha_{\sigma^{-1}(1)}, \ldots, \alpha_{\sigma^{-1}(n)}). \tag{4.11}$$

Let $f = \frac{4}{3} + \frac{5}{9}\left(e^{9x_1} + e^{9x_2} + ^{9x_3}\right) - e^{3x_1+x_2+x_3} - e^{x_1+3x_2+x_3} - e^{x_1+x_2+3x_3}$. This is a symmetric exponential sum with respect to the symmetric group $\mathcal{S}_3$. In fact, this is also a SAGE exponential. It can be decomposed into the three exponential sums

$$\begin{aligned} f_1 =& \frac{4}{9} + \frac{1}{3}e^{9x_1} + \frac{1}{9}e^{9x_2} + \frac{1}{9}e^{9x_3} - e^{3x_1+x_2+x_3}, \\ f_2 =& \frac{4}{9} + \frac{1}{9}e^{9x_1} + \frac{1}{3}e^{9x_2} + \frac{1}{9}e^{9x_3} - e^{x_1+3x_2+x_3} \text{ and} \\ f_3 =& \frac{4}{9} + \frac{1}{9}e^{9x_1} + \frac{1}{9}e^{9x_2} + \frac{1}{3}e^{9x_3} - e^{x_1+x_2+3x_3}. \end{aligned}$$

Fix $\hat{\beta} := (3,1,1)^T$. The stabilizer is $\operatorname{Stab}(\hat{\beta}) = \{(123),(132)\}$ and, with this, we compute $\mathcal{S}_3/\operatorname{Stab}(\hat{\beta}) = \{(123),(213),(312)\}$. Here, elements in the sets $\operatorname{Stab}(\hat{\beta})$ and $\mathcal{S}_3/\operatorname{Stab}(\hat{\beta})$ are written as tuples. Then, observe that

$$h_{\hat{\beta}} := \frac{1}{6}\sum_{\sigma\in G} \sigma f_{\sigma^{-1}(\beta)} = f_1.$$

Also, $\rho h_{\hat{\beta}} = h_{\rho\hat{\beta}}$ and for the three elements in $\mathcal{S}_3/\operatorname{Stab}(\hat{\beta})$, we have

$$(123)\hat{\beta} = \hat{\beta},\ (213)\hat{\beta} = (1,3,1),\ (312)\hat{\beta} = (1,1,3),$$

hence, $f = \sum_{\rho\in\mathcal{S}_3/\operatorname{Stab}(\hat{\beta})} \rho f_1 = \sum_{\rho\in\mathcal{S}_3/\operatorname{Stab}(\hat{\beta})} \rho h_{\hat{\beta}}$.

Proof of Theorem 4.2.1. Since it is clear that any exponential sum appearing in the decomposition of f as defined in (4.10) is nonnegative and contains a single possibly negative term, we only have to show the converse direction. Let $f \in \mathcal{S}_{A^+,A^-}$. By Theorem 4.1.6, there exist AGE exponentials $f_\beta \in \mathcal{S}_{A^+,\beta}$ for every $\beta \in A^-$, such that $f = \sum_{\beta\in A^-} f_\beta$. The G-invariance of f gives

$$f = \frac{1}{|G|}\sum_{\sigma\in G}\sigma f = \frac{1}{|G|}\sum_{\sigma\in G}\sum_{\beta\in A^-}\sigma f_\beta. \tag{4.12}$$

The idea is to group all the σf_β that have the same *possibly negative* term. According to (4.8), the possibly negative term of σf_β is given by $\sigma(\beta)$. For any $\beta \in A^-$, the exponential sum

$$h_\beta = \frac{1}{|G|}\sum_{\sigma\in G}\sigma f_{\sigma^{-1}(\beta)}$$

is a sum of AGE exponentials in $\mathcal{S}_{A^+,\beta}$, hence, it is contained in $\mathcal{S}_{A^+,\beta}$ as well. Moreover, (4.12) can be expressed as

$$f = \frac{1}{|G|}\sum_{\sigma\in G}\sum_{\beta\in A^-}\sigma f_\beta = \frac{1}{|G|}\sum_{\sigma\in G}\sum_{\gamma\in A^-}\sigma f_{\sigma^{-1}(\gamma)} = \sum_{\gamma\in A^-} h_\gamma.$$

Let $\beta \in A^-$ and $\hat{\beta} \in \hat{A}^-$ be the representative of its orbit in $\hat{A}^-$. If $\sigma,\tau \in G$ are such that $\sigma(\hat{\beta}) = \tau(\hat{\beta}) = \beta$, then $\tau^{-1}\sigma \in \operatorname{Stab}(\hat{\beta})$ and $\tau = \sigma$ in $G/\operatorname{Stab}(\hat{\beta})$. Hence,

$$f = \sum_{\hat{\beta}\in\hat{A}^-}\ \sum_{\rho\in G/\operatorname{Stab}\hat{\beta}} h_{\rho(\hat{\beta})}. \tag{4.13}$$

Now observe that $h_{\rho\beta} = \rho h_\beta$ for every $\beta \in A^-$ and $\rho \in G$ because

$$|G|\rho h_\beta = \sum_{\sigma\in G}\rho\sigma f_{\sigma^{-1}(\beta)} = \sum_{\tau\in G}\tau f_{\tau^{-1}\rho(\beta)} = |G| h_{\rho(\beta)}. \tag{4.14}$$

Substituting (4.14) into (4.13) gives $f = \sum_{\hat{\beta} \in \hat{A}^-} \sum_{\rho \in G/\operatorname{Stab}\hat{\beta}} \rho h_{\hat{\beta}}$ as desired. Moreover, the $\operatorname{Stab}(\hat{\beta})$-invariance of $h_{\hat{\beta}}$ for $\hat{\beta} \in \hat{A}^-$ follows from (4.14). $\square$

4.2.2 Symmetry Reduction in Relative Entropy Programming

Building upon the previous decomposition theorem, we provide a symmetry-adapted relative entropy formulation for containment in the SAGE-cone.

Theorem 4.2.3. *Let $\hat{A}^-$ be a set of orbit representatives for A^-. A G-invariant exponential sum f of the form* (4.9) *is contained in $\mathcal{S}_{A^+,A^-}$ if and only if for every $\hat{\beta} \in \hat{A}^-$ there exist $c^{(\hat{\beta})} \in \mathbb{R}_+^{A^+}$ and $\nu^{(\hat{\beta})} \in \mathbb{R}_+^{A^+}$, invariant under the action of* $\operatorname{Stab}(\hat{\beta})$*, such that*

$$\begin{aligned}
\sum_{\alpha \in A^+} \nu_\alpha^{(\hat{\beta})}(\alpha - \hat{\beta}) \;&=\; 0 \quad \textit{for every } \hat{\beta} \in \hat{A}^-, && (4.15)\\
D(\nu^{(\hat{\beta})}, e \cdot c^{(\hat{\beta})}) \;&\leq\; c_{\hat{\beta}} \quad \textit{for every } \hat{\beta} \in \hat{A}^- \textit{ and} && (4.16)\\
\sum_{\hat{\beta} \in \hat{A}^-} \sum_{\sigma \in \operatorname{Stab}(\hat{\beta}) \backslash G} c_{\sigma(\alpha)}^{(\hat{\beta})} \;&\leq\; c_\alpha \quad \textit{for every } \alpha \in A^+. && (4.17)
\end{aligned}$$

Remark 4.2.4. The right coset condition (4.17) can equivalently be expressed in terms of left cosets,

$$\sum_{\hat{\beta} \in \hat{A}^-} \sum_{\sigma \in G/\operatorname{Stab}\hat{\beta}} c_{\sigma^{-1}(\alpha)}^{(\hat{\beta})} \;\leq\; c_\alpha \quad \text{for every } \alpha \in A^+.$$

Namely, if $\beta \in A^-$, $\hat{\beta} \in \hat{A}^-$ and $\sigma, \tau \in G$ are such that $\sigma^{-1}(\hat{\beta}) = \tau^{-1}(\hat{\beta}) = \beta$, then $\tau\sigma^{-1} \in \operatorname{Stab}(\hat{\beta})$ and $\tau = \sigma$ in the right quotient space $\operatorname{Stab}(\hat{\beta}) \backslash G$.

The following example shows the usefulness of Theorem 4.2.3.

Example 4.2.5. Consider the $\mathcal{S}_3$-symmetric exponential sum

$$f(x_1, x_2, x_3) \;=\; 1 + \sum_{i=1}^{3} e^{8x_i} - \delta \sum_{(i,j,k) \in \mathcal{S}_3} e^{3x_i + 2x_j + x_k} - \delta \sum_{(i,j,k) \in \mathcal{S}_3} e^{2x_i + x_j + x_k},$$

and we ask for the largest δ for which f is SAGE. Let $\delta^{(i)}$ denote the i-th unit vector for $i \in \mathbb{N}$, and let $\mathbf{0}$ denote the three-dimensional zero vector. The conventional relative entropy program from Theorem 4.1.4 has $2 \cdot 4 \cdot 9 + 1 = 73$ variables (including the δ-variable) and $3 \cdot 9 + 9 + 4 = 40$ equations or inequalities. Observing $A^+ = \{\mathbf{0}, 8\delta^{(1)}, 8\delta^{(2)}, 8\delta^{(3)}\}$ as well as

$$|\hat{A^-}| = |\{(3,2,1)^T, (2,1,1)^T\}| = 2,$$

the symmetric relative entropy program is

$$\begin{aligned}
&\min \delta\\
\text{s.t.}\quad \textstyle\sum_{\alpha \in A^+} \nu_\alpha^{(\beta)}(\alpha - \beta) \;&=\; 0 \qquad \text{for } \beta = (3,2,1)^T, (2,1,1)^T,\\
D(\nu^{(\beta)}, e \cdot c^{(\beta)}) \;&\leqslant\; \delta \qquad \text{for } \beta = (3,2,1)^T, (2,1,1)^T,\\
\sum_{\beta = (3,2,1)^T, (2,1,1)^T} 9 c_\alpha^{(\beta)} \;&\leqslant\; 1 \qquad \text{for } \alpha \in A^+,\\
\delta \in \mathbb{R} \text{ and } c^{(\beta)}, \nu^{(\beta)} \;&\in\; \mathbb{R}_+^4 \quad \text{for } \beta = (3,2,1)^T, (2,1,1)^T.
\end{aligned}$$

Therefore, we see that the symmetric relative entropy program from Remark 4.2.4 involves $2 \cdot 4 \cdot 2 + 1 = 17$ variables and at most $2 \cdot 3 + 2 + 4 = 12$ equations or inequalities.

Proof of Theorem 4.2.3. If $A^- = \emptyset$, then the statement is trivially true. Hence, assume $A^- \neq \emptyset$.

If f is G-symmetric, then, by Theorem 4.2.1, there exist $\operatorname{Stab}(\beta)$-invariant AGE exponentials $h_\beta \in \mathcal{S}_{A^+,\beta}$ for every $\beta \in \hat{A}^-$ such that

$$f = \sum_{\beta \in \hat{A}^-} \sum_{\rho \in G/\operatorname{Stab}(\hat{\beta})} \rho h_\beta.$$

Writing h_β in the form

$$h_\beta = \sum_{\alpha \in A^+} c_\alpha^{(\beta)} e^{\langle \alpha, x\rangle} + c_\beta e^{\langle \beta, x\rangle}$$

with coefficients $c_\alpha^{(\beta)}$ and c_β for $\alpha \in A^+$ and $\beta \in \hat{A}^-$, the two conditions (4.15) and (4.16) follow from the property $h_\beta \in \mathcal{S}_{A^+,\beta}$. For (4.17), we observe that for $\alpha \in A^+$, the coefficient of $e^{\langle \alpha, x\rangle}$ in ρh_β is $c_{\rho^{-1}(\alpha)}^{(\beta)}$. We obtain inequality (4.17), even with equality, by setting $\sigma := \rho^{-1}$ and summing over $\beta \in \hat{A}^-$ and over $\sigma \in \operatorname{Stab}(\beta)\backslash G$, following Remark 4.2.4. Moreover, the $\operatorname{Stab}(h_\beta)$-invariance of h_β implies the $\operatorname{Stab}(\beta)$-invariance of $c^{(\beta)}$. In order to make $\nu^{(\beta)}$ invariant under $\operatorname{Stab}(\beta)$, we can replace it with

$$\mu_\alpha^{(\beta)} = \frac{1}{|\operatorname{Stab}(\beta)|} \sum_{\sigma \in \operatorname{Stab}(\beta)} \nu_{\sigma(\alpha)}^{(\beta)}.$$

Obviously, this has no influence on (4.17). For (4.15), we have

$$\begin{aligned}
|\operatorname{Stab}(\beta)| \sum_{\alpha \in A^+} \mu_\alpha^{(\beta)} (\alpha - \beta) &= \sum_{\alpha \in A^+} \sum_{\sigma \in \operatorname{Stab}(\beta)} \nu_{\sigma(\alpha)}^{(\beta)} (\alpha - \beta) \\
&= \sum_{\sigma \in \operatorname{Stab}(\beta)} \sigma^{-1} \sum_{\alpha \in A^+} \nu_{\sigma(\alpha)}^{(\beta)} (\sigma(\alpha) - \sigma(\beta)) \\
&= \sum_{\sigma \in \operatorname{Stab}(\beta)} \sigma^{-1} \sum_{\alpha \in A^+} \nu_\alpha^{(\beta)} (\alpha - \beta)) = 0.
\end{aligned}$$

Finally, for (4.16), using $c_\alpha^{(\beta)} = c_{\sigma(\alpha)}^{(\beta)}$ for $\sigma \in \operatorname{Stab}(\beta)$ and applying Jensen's inequality on the convex function $x \mapsto x \ln x$ gives, for all $\alpha \in A^+$,

$$\begin{aligned}
\mu_\alpha^{(\beta)} \ln \frac{\mu_\alpha^{(\beta)}}{c_\alpha^{(\beta)}} &= \left(\frac{1}{|\operatorname{Stab}(\beta)|} \sum_{\sigma \in \operatorname{Stab}(\beta)} \nu_{\sigma(\alpha)}^{(\beta)} \right) \ln \frac{\frac{1}{|\operatorname{Stab}(\beta)|} \sum_{\sigma \in \operatorname{Stab}(\beta)} \nu_{\sigma(\alpha)}^{(\beta)}}{c_\alpha^{(\beta)}} \\
&= c_\alpha^{(\beta)} \left(\frac{\sum_{\sigma \in \operatorname{Stab}(\beta)} \nu_{\sigma(\alpha)}^{(\beta)} / c_{\sigma(\alpha)}^{(\beta)}}{|\operatorname{Stab}(\beta)|} \ln \frac{\sum_{\sigma \in \operatorname{Stab}(\beta)} \nu_{\sigma(\alpha)}^{(\beta)} / c_{\sigma(\alpha)}^{(\beta)}}{|\operatorname{Stab}(\beta)|} \right) \\
&\leq c_\alpha^{(\beta)} \left(\frac{1}{|\operatorname{Stab}(\beta)|} \sum_{\sigma \in \operatorname{Stab}(\beta)} \frac{\nu_{\sigma(\alpha)}^{(\beta)}}{c_{\sigma(\alpha)}^{(\beta)}} \ln \frac{\nu_{\sigma(\alpha)}^{(\beta)}}{c_{\sigma(\alpha)}^{(\beta)}} \right).
\end{aligned}$$

Using again the $\operatorname{Stab}(\beta)$-invariance of $c^{(\beta)}$ and the precondition then yields

$$\begin{aligned}\sum_{\alpha\in A^+} \mu_\alpha^{(\beta)} \ln \frac{\mu_\alpha^{(\beta)}}{ec_\alpha^{(\beta)}} &\leq \frac{1}{|\operatorname{Stab}(\beta)|} \sum_{\sigma\in\operatorname{Stab}(\beta)} \sum_{\alpha\in A^+} \nu_{\sigma(\alpha)}^{(\beta)} \ln \frac{\nu_{\sigma(\alpha)}^{(\beta)}}{ec_{\sigma(\alpha)}^{(\beta)}} \\ &\leq \frac{1}{|\operatorname{Stab}(\beta)|} \sum_{\sigma\in\operatorname{Stab}(\beta)} c_\beta = c_\beta.\end{aligned}$$

Conversely, assume that $c^{(\hat\beta)}$ and $\nu^{(\hat\beta)}$, invariant under the action of $\operatorname{Stab}(\hat\beta)$, satisfy (4.15)–(4.17). Let $\beta \in A^-$ and $\hat\beta \in \hat{A}^-$ be the representative of its orbit in $\hat{A}^-$. If $\sigma, \tau \in G$ are such that $\sigma(\beta) = \tau(\beta) = \hat\beta$, then $\tau\sigma^{-1} \in \operatorname{Stab}(\hat\beta)$ and $\tau = \sigma$ in $\operatorname{Stab}(\hat\beta)\backslash G$. Since $c^{(\hat\beta)}$ and $\nu^{(\hat\beta)}$ are invariant under $\operatorname{Stab}(\hat\beta)$, we have

$$c_{\tau(\alpha)}^{(\hat\beta)} = c_{\sigma(\alpha)}^{(\hat\beta)}, \quad \nu_{\tau(\alpha)}^{(\hat\beta)} = \nu_{\sigma(\alpha)}^{(\hat\beta)} \quad \text{for } \alpha \in A^+.$$

Thus, we can define

$$c_\alpha^{(\beta)} = c_{\sigma(\alpha)}^{(\hat\beta)}, \quad \nu_\alpha^{(\beta)} = \nu_{\sigma(\alpha)}^{(\hat\beta)} \quad \text{for } \alpha \in A^+,$$

which is independent of σ such that $\sigma(\beta) = \hat\beta$. As a consequence, if $\tau \in \operatorname{Stab}(\hat\beta)\backslash G$, then $c_\alpha^{(\tau^{-1}(\hat\beta))} = c_{\tau(\alpha)}^{(\hat\beta)}$ is well defined.

To see that the first conditions of Theorem 4.1.4 are satisfied, let $\beta \in A^-$ and $\sigma \in G$ such that $\sigma(\beta) = \hat\beta$. Then

$$\begin{aligned}\sum_{\alpha\in A^+} \nu_\alpha^{(\beta)}(\alpha - \beta) &= \sum_{\alpha\in A^+} \nu_{\sigma(\alpha)}^{(\hat\beta)}(\alpha - \sigma^{-1}\hat\beta) = \sigma^{-1} \sum_{\alpha\in\mathcal{A}} \nu_{\sigma(\alpha)}^{(\hat\beta)}(\sigma(\alpha) - \hat\beta) \\ &= \sigma^{-1} \sum_{\alpha\in\mathcal{A}} \nu_\alpha^{(\hat\beta)}(\alpha - \hat\beta) = 0 \text{ and} \\ D(\nu^{(\beta)}, ec^{(\beta)}) &= D(\nu^{(\hat\beta)}, ec^{(\hat\beta)}) \leq c_{\hat\beta} = c_\beta.\end{aligned}$$

For the third condition of Theorem 4.1.4, we obtain

$$\sum_{\beta\in A^-} c_\alpha^{(\beta)} = \sum_{\hat\beta\in\hat{A}^-} \sum_{\tau\in\operatorname{Stab}(\hat\beta)\backslash G} c_\alpha^{(\tau^{-1}\hat\beta)} = \sum_{\hat\beta\in\hat{A}^-} \sum_{\tau\in\operatorname{Stab}(\hat\beta)\backslash G} c_{\tau(\alpha)}^{(\hat\beta)} \leq c_\alpha,$$

which altogether shows that $f \in \mathcal{S}_{A^+,A^-}$. □

The following consequence of Theorem 4.2.3 further reduces the number of variables in the relative entropy program, since a certain number of $c_\alpha^{(\hat\beta)}$ and $\nu_\alpha^{(\hat\beta)}$ are actually equal, and we can take each $c^{(\hat\beta)}, \nu^{(\hat\beta)}$ in the ground set $\mathbb{R}_+^{A^+/\operatorname{Stab}(\hat\beta)}$.

Corollary 4.2.6. *Let $\hat{A}^+$ and $\hat{A}^-$ be a set of orbit representatives for A^+ and A^-. A G-invariant exponential sum f of the form* (4.9) *is contained in $\mathcal{S}_{A^+,A^-}$ if and only* *if for every* $\hat{\beta} \in \hat{A}^-$ *there exist* $c^{(\hat{\beta})} \in \mathbb{R}_+^{A^+ / \operatorname{Stab}(\hat{\beta})}$ *and* $\nu^{(\hat{\beta})} \in \mathbb{R}_+^{A^+ / \operatorname{Stab}(\hat{\beta})}$ *such that*

$$\sum_{\alpha \in A^+ / \operatorname{Stab}(\hat{\beta})} \nu_\alpha^{(\hat{\beta})} \sum_{\alpha' \in \operatorname{Stab}(\hat{\beta}) \cdot \alpha} (\alpha' - \hat{\beta}) \;=\; 0 \quad \forall\, \hat{\beta} \in \hat{A}^-, \tag{4.18}$$

$$\sum_{\alpha \in A^+ / \operatorname{Stab}(\hat{\beta})} \left| \operatorname{Stab}(\hat{\beta}) \cdot \alpha \right| \nu_\alpha^{(\hat{\beta})} \ln \frac{\nu_\alpha^{(\hat{\beta})}}{e c_\alpha^{(\hat{\beta})}} \;\leq\; c_{\hat{\beta}} \quad \forall\, \hat{\beta} \in \hat{A}^-, \tag{4.19}$$

$$\sum_{\hat{\beta} \in \hat{A}^-} \frac{|\operatorname{Stab}(\alpha)|}{|\operatorname{Stab}(\hat{\beta})|} \sum_{\gamma \in (G \cdot \alpha) / \operatorname{Stab}(\hat{\beta})} \left| \operatorname{Stab}(\hat{\beta}) \cdot \gamma \right| c_\gamma^{(\hat{\beta})} \;\leq\; c_\alpha \quad \forall\, \alpha \in \hat{A}^+. \tag{4.20}$$

Proof. For (4.18) and (4.19), equivalence to their versions in Theorem 4.2.3 is straightforward to check. For (4.20), equivalence to (4.17) follows by observing that for every $\alpha \in A^+$

$$\begin{aligned} \sum_{\sigma \in \operatorname{Stab}(\hat{\beta}) \backslash G} c_{\sigma(\alpha)}^{(\hat{\beta})} &= \sum_{\sigma \in \operatorname{Stab}(\hat{\beta}) \backslash G} \frac{1}{|\operatorname{Stab}(\hat{\beta})|} \sum_{\tau \in \operatorname{Stab}(\hat{\beta})} c_{\tau(\sigma(\alpha))}^{(\hat{\beta})} = \frac{1}{|\operatorname{Stab}(\hat{\beta})|} \sum_{\rho \in G} c_{\rho(\alpha)}^{(\hat{\beta})} \\ &= \frac{|\operatorname{Stab}(\alpha)|}{|\operatorname{Stab}(\hat{\beta})|} \sum_{\gamma \in G \cdot \alpha} c_\gamma^{(\hat{\beta})} = \frac{|\operatorname{Stab}(\alpha)|}{|\operatorname{Stab}(\hat{\beta})|} \sum_{\gamma \in (G \cdot \alpha) / \operatorname{Stab}(\hat{\beta})} \left| \operatorname{Stab}(\hat{\beta}) \cdot \gamma \right| c_\gamma^{(\hat{\beta})}, \end{aligned}$$

and the last expression only depends on the orbit $G \cdot \alpha$ rather than on α itself. □

Remark 4.2.7. Note that we cannot simply assume $c_\alpha^{(\beta)} = c_{\alpha'}^{(\beta)}$ for some $\alpha' \in G \cdot \alpha$ and, similarly, we cannot simply assume $\nu_\alpha^{(\beta)} = \nu_{\alpha'}^{(\beta)}$ for some $\alpha' \in G \cdot \alpha$. Namely, if an element β lies in $\operatorname{conv} A^+$ with barycentric coordinates λ, say $\beta = \sum_{\alpha \in A^+} \lambda_\alpha \alpha$, then for any $\sigma \in G$, we have

$$\sigma(\beta) = \sigma\left(\sum_{\alpha \in A^+} \lambda_\alpha \alpha\right) = \sum_{\alpha \in A^+} \sigma(\lambda_\alpha \alpha) = \sum_{\alpha \in A^+} \lambda_\alpha \sigma(\alpha)$$

rather than $\sigma(\beta) = \sum_{\alpha \in A^+} \lambda_\alpha \alpha = \sum_{\alpha \in A^+} \lambda_{\sigma(\alpha)} \sigma(\alpha)$. Of course, this caveat does not occur whenever there is a single inner term.

To close this section, we discuss the resulting complexity reduction:

Recall that the initial relative entropy formulation, which does not take the symmetry into consideration, involves $2|A^-||A^+|$ variables. Furthermore, since every vector equality in (4.18) brings n scalar equalities, it consists of $|A^-|n + |A^-| + |A^+|$ (in)equalities.

In contrast, let us analyze the number of variables and constraints involved in the relative entropy program in Corollary 4.2.6. Observe that $A^+ / \operatorname{Stab}(\hat{\beta})$ is the disjoint union of the $G \cdot \hat{\alpha} / \operatorname{Stab}(\hat{\beta})$ where $\hat{\alpha}$ runs through $\hat{A}^+$. It follows that for every pair $\hat{\beta} \in \hat{A}^-, \hat{\alpha} \in \hat{A}^+$, we have exactly $2|(G \cdot \hat{\alpha}) / \operatorname{Stab}(\hat{\beta})|$ variables $c_\gamma^{(\hat{\beta})}$ and $\nu_\gamma^{(\hat{\beta})}$.

By definition, $|(G \cdot \hat{\alpha}) / \operatorname{Stab}(\hat{\beta})|$ is the number of orbits $\left\{\operatorname{Stab}(\hat{\beta}) \cdot \gamma \,:\, \gamma \in G \cdot \hat{\alpha}\right\}$. Since $G \cdot \hat{\alpha}$ is in bijection with $\operatorname{Stab} \hat{\alpha} \backslash G$, we get a bijection between $(G \cdot \hat{\alpha}) / \operatorname{Stab}(\hat{\beta})$ and the set of double cosets $\operatorname{Stab}(\hat{\alpha}) \backslash G / \operatorname{Stab}(\hat{\beta})$. Therefore, the number of orbits in question equals $|\operatorname{Stab}(\hat{\alpha}) \backslash G / \operatorname{Stab}(\hat{\beta})|$, satisfying, according to the Orbit Counting Theorem (Lemma 2.2.4):

$$|\operatorname{Stab}(\hat{\alpha}) \backslash G / \operatorname{Stab}(\hat{\beta})| = \frac{1}{|\operatorname{Stab}(\hat{\alpha})||\operatorname{Stab}(\hat{\beta})|} \sum_{\substack{\sigma \in \operatorname{Stab}(\hat{\alpha}) \\ \tau \in \operatorname{Stab}(\hat{\beta})}} |G^{\sigma,\tau}|,$$

where $|G^{\sigma,\tau}|$ is the number of elements of G fixed under the action of (σ, τ). From another point of view, this number can be interpreted in terms of representation theory as follows: It is given by the inner product of the two characters corresponding to the representations induced by the trivial representations of $\operatorname{Stab}(\hat{\alpha})$ and $\operatorname{Stab}(\hat{\beta})$ on G (see [Sta99, Exercise 7.77.a] for more details).

Furthermore, (4.18) amounts to $|\hat{A}^+| + |\hat{A}^-|$ inequalities, together with one vector equality for every element of $\hat{A}^-$. We observe that for a given $\hat{\beta}$, this vector is invariant by $\operatorname{Stab}(\hat{\beta})$ and therefore is contained in $(\mathbb{R}^n)^{\operatorname{Stab}(\hat{\beta})}$, the subspace of $\mathbb{R}^n$ of points fixed by $\operatorname{Stab}(\hat{\beta})$. Thus, the number of resulting equations reduces to $\dim\left((\mathbb{R}^n)^{\operatorname{Stab}(\hat{\beta})}\right)$ by projecting onto this subspace. As a conclusion, we state the following theorem from [Mou+21] without proof:

Theorem 4.2.8. *Let $\hat{A}^+$ and $\hat{A}^-$ be a set of orbit representatives for A^+ and A^-. For $\hat{\alpha} \in \hat{A}^+$, $\hat{\beta} \in \hat{A}^-$, denote by ${}_{\hat{\alpha}}G_{\hat{\beta}}$ the cardinality $|\operatorname{Stab}(\hat{\alpha})\backslash G/\operatorname{Stab}(\hat{\beta})|$, and by $n_{\hat{\beta}}$ the dimension of the fixed subspace $(\mathbb{R}^n)^{\operatorname{Stab}(\hat{\beta})}$. Then, the relative entropy program in Corollary 4.2.6 consists of*

$$2\sum_{\substack{\hat{\alpha}\in\hat{A}^+\\ \hat{\beta}\in\hat{A}^-}} {}_{\hat{\alpha}}G_{\hat{\beta}} \text{ variables,} \qquad \sum_{\hat{\beta}\in\hat{A}^-} n_{\hat{\beta}} \text{ scalar equalities, and} \quad |\hat{A}^+| + |\hat{A}^-| \text{ inequalities.}$$

Remark 4.2.9. For special groups such as the symmetric group, the relative entropy programs can be simplified even more using various combinatorial techniques. Those techniques were applied for the numerical experiments in the next subsection; the corresponding statements can be found in [Mou+21]. They are not included here as this would go beyond the line of research examined in this thesis.

4.2.3 Numerical Experiments

To illustrate the previous considerations, we present in this section classes of examples that spotlight the computational gains by the comparison of calculation times in the case of the symmetric group. For these computations, we used the ECOS solver [DCB13] and Python 3.7 on an Intel(R) Xeon(R) Platinum 8168 CPU with 2.7 GHz and 768 GiB of RAM under CentOS Linux release 7.9.2009.

Keeping the previous notation, for the standard method, that is the method that does not exploit the symmetries, the input consists of A^+, A^- as well as the coefficients, while for the symmetry-adapted version, the input is $\hat{A}^+$, $\hat{A}^-$ and the coefficients. This difference of input is mainly due to practical considerations and does not in itself influence the comparison of the time used by the solver. When both methods give an answer, the bounds coincide.

In all the tables in the following, "dim" is the dimension, "V_n" and "C_n" are the number of variables and constraints of the program, while "t_s" and "t_r" denote the solver time and the overall run time (including the building of the optimization program) in seconds. While it might happen that the standard method is slightly faster for very small instances, the growth in size for the program in the standard method makes it quickly practically unsolvable. In that case, this is represented by the symbol "$-$" in the table. The symmetric approach however allows us to go further, and we give all the results until the solver warns about a possible inaccuracy. In this case, we mark the bound with "*".

Example 4.2.10. Consider first the exponential sum

$$f_n^{(1)} = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \sigma e^{\langle \alpha, x \rangle} - e^{\langle \beta, x \rangle}$$

where $\beta = (1, \ldots, 1)$ and $\alpha = (1, 2, \ldots, n)$. The numerical results are shown in Table 4.1.

		Standard method				Symmetric method			
dim	bound	V_n	C_n	t_s	t_r	V_n	C_n	t_s	t_r
2	-0.1481	7	6	0.0113	0.0121	5	4	0.0147	0.0158
3	-0.2499	15	11	0.0148	0.0160	5	4	0.0141	0.0149
4	-0.3257	51	30	0.0304	0.0337	5	4	0.0139	0.0147
5	-0.3849	243	127	–	–	5	4	0.0140	0.0147
6	-0.4327	1443	728	–	–	5	4	0.0136	0.0144
7	-0.4724*	10083	5049	–	–	5	4	0.0211	0.0222

TABLE 4.1: Numerical results for $f_n^{(1)}$.

Example 4.2.11. Consider now the exponential sum

$$f_n^{(2)} = (n-1)! \sum_{i=1}^{n} e^{n^2 x_i} - \sum_{\sigma \in \mathcal{S}_n} \sigma e^{\langle \beta, x \rangle}$$

where $\beta = (1, 2, \ldots, n)^T$ (and $\alpha = (n^2, 0, \ldots, 0)^T$). The numerical results are shown in Table 4.2.

		Standard method				Symmetric method			
dim	bound	V_n	C_n	t_s	t_r	V_n	C_n	t_s	t_r
2	-0.2109	13	9	0.0173	0.0185	7	5	0.0297	0.0311
3	-0.8888	49	28	0.0427	0.0454	9	6	0.0248	0.0264
4	-4.111	241	125	0.152	0.1701	11	7	0.0296	0.0318
5	-22.30	1441	726	0.7888	0.8433	13	8	0.0356	0.0384
6	-141.0	10081	5047	5.422	5.843	15	9	0.0423	0.0458
7	-1024	80641	40328	57.26	66.67	17	10	0.0491	0.0538
8	-8418	725761	362889	1514	2211	19	11	0.0568	0.0626
9	-77355	7257601	3628810	–	–	21	12	0.0661	0.0835
10		79833601	39916811	–	–	23	13	–	–

TABLE 4.2: Numerical results for $f_n^{(2)}$.

Example 4.2.12. Next, we consider the case where both orbits are of maximal size. Let

$$f_n^{(3)} = \frac{1}{n} \sum_{\sigma \in \mathcal{S}_n} e^{\langle \alpha, x \rangle} - \frac{1}{n} \sum_{\sigma \in \mathcal{S}_n} \sigma e^{\langle \beta, x \rangle}$$

where $\beta = (1, 2, \ldots, n)$ and $\alpha = (2, 8, \ldots, 2n^2)$.

The numerical results are shown in Table 4.3.

		Standard method				**Symmetric method**			
dim	bound	V_n	C_n	t_s	t_r	V_n	C_n	t_s	t_r
2	-0.4178	13	9	0.0301	0.0323	7	5	0.0431	0.0465
3	-1.0323	85	31	0.0558	0.0603	15	6	0.0531	0.0569
4	-3.494	1201	145	–	–	51	7	0.1212	0.1301
5	-15.13	29041	841	–	–	243	8	0.5750	0.6215
6		1038241	5761	–	–	1443	9	–	–

TABLE 4.3: Numerical results for $f_n^{(3)}$.

Example 4.2.13. Finally, we consider the case where both orbits are small. Let

$$f_n^{(4)} = \frac{1}{n}\sum_{i=1}^{n} e^{n^2 x_i} - \frac{1}{n}\sum_{i=1}^{n} e^{(n-1)(x_1+\cdots+x_n)+x_i}),$$

($\beta = (n, n-1, n-1, \ldots, n-1)$ and $\alpha = (n^2, 0, \ldots, 0)$). The numerical results are shown in Table 4.4.

		Standard method				**Symmetric method**			
dim	bound	V_n	C_n	t_s	t_r	V_n	C_n	t_s	t_r
2	-0.1054	13	9	0.019 01	0.0204	7	5	0.0213	0.0229
3	-0.092	25	16	0.0268	0.0287	7	5	0.0205	0.0218
4	-0.076	41	25	0.0341	0.0367	7	5	0.0205	0.0218
68	-0.0053	9385	4761	–	–	7	5	0.0475	0.0519
95	-0.0038*	18241	9216	–	–	7	5	0.0267	0.0281

TABLE 4.4: Numerical results for $f_n^{(4)}$.

Example 4.2.14. Finally, we give an example where A^+ and A^- consist of two orbits each:

$$\hat{A}^+ = \{(n^2, 0, \ldots, 0), (1, 4, \ldots, n^2)\} \text{ and } \hat{A}^- = \{(1, \ldots, 1), (1, 2, \ldots, n)\}.$$

In this case, we are still able to compute the number of constraints and the number of variables. With the standard approach,

$$V_n = 2(n! + n + 1)(n! + 1) + 1, \quad C_n = (n! + 1)(n + 2) + n,$$

while using symmetries,

$$V_n = 2n! + 2n + 9, \quad C_n = n + 6.$$

Table 4.5 shows the numerical results for the exponential sums

$$g_n = \frac{1}{n}\sum_{i=1}^{n} e^{n^2 x_i} + \frac{1}{n}\sum_{\sigma \in \mathcal{S}_n} \sigma e^{\langle \alpha, x\rangle} - e^{x_1+\cdots+x_n} - \frac{1}{n}\sum_{\sigma \in \mathcal{S}_n} \sigma e^{\langle \beta, x\rangle}$$

for $\alpha = (1, 4, \ldots, n^2)^T$ and $\beta = (1, 2, \ldots, n)^T$.

		Standard method				Symmetric method			
dim	bound	V_n	C_n	t_s	t_r	V_n	C_n	t_s	t_r
2	-0.1918	31	14	0.0272	0.0292	17	8	0.0738	0.0776
3	-0.5223	141	38	0.0679	0.0727	27	9	0.0623	0.0663
4	-2.118	1451	154	–	–	65	10	0.1436	0.1539
5	-10.45	30493	852	–	–	259	11	0.5856	0.6320
6		1048335	5774	–	–	1461	12	–	–

Table 4.5: Numerical results for g_n.

4.3 An Approximation via the Dual SAGE-Cone and Linear Programming

In this section, we provide a relaxation of the global optimization problem to minimize an exponential sum using the dual of the SAGE-cone, including multivariate real polynomials as special cases. The key idea is to relax the problem (1.1) via optimizing over a variant of the dual signed SAGE-cone $\mathcal{S}^*_{A^+,A^-}$. Our approach is motivated by the recent works [DNT21], [MCW21a], and [KNT21], and builds on two key observations, which are the main theoretical contributions:

1. The dual cone of AGE exponentials is contained in the primal one; see Proposition 4.3.1.
2. Optimizing over the dual cone can be carried out by solving a linear program; see Proposition 4.3.4.

We emphasize that neither the primal nor the dual SAGE-cone is polyhedral; see in this context also the results in [FW19]. The approach works as follows: First, we investigate a lifted version of the dual cone involving additional linear auxiliary variables (Theorem 4.1.6 (3)). Second, we show that the coefficients of a given exponential sum can be interpreted as variables of the dual cone — and sums of these exponentials can be interpreted as having coefficients *induced* by the dual cone; see (4.21). Third, we observe that fixing these coefficient variables yields a linear optimization problem only involving auxiliary variables; see Proposition 4.3.4

4.3.1 The Dual Cone of AGE Exponentials is Contained in the Primal Cone

For $\emptyset \neq \mathcal{A} = A^+ \cup A^-$ defined as in (4.1), we identified the dual space of exponential sums supported on $\mathcal{A}$ with $\mathbb{R}^{\mathcal{A}}$. Now, for every $v \in \mathbb{R}^{\mathcal{A}}$, we associate a function

$$f(x) = |A^-| \sum_{\alpha\in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle}. \tag{4.21}$$

Note that circuit exponentials and AGE exponentials are special cases of these functions, and, moreover, for every $v \in \mathcal{S}^*_{A^+,\beta}$, an f constructed in the sense of (4.21) has coefficients contained in $\mathcal{S}^*_{A^+,\beta}$, as β is the single term not contained in A^+. With this consideration, we set $\mathcal{F}^*_{A^+,\beta}$ to be the cone of exponential sums having exponents in $A^+ \cup \{\beta\}$ and corresponding coefficients in $\mathcal{S}^*_{A^+,\beta} \cap \mathbb{R}^{A^+\cup\{\beta\}}_{\neq 0}$, and $\mathcal{F}^*_{A^+,A^-}$ as the cone of exponential sums $|A^-| \sum_{\alpha\in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle}$ with $v \in \mathcal{S}^*_{A^+,A^-} \cap \mathbb{R}^{\mathcal{A}}_{\neq 0}$. Note that not every element $v \in \mathcal{S}^*_{A^+,A^-}$ induces a function in $\mathcal{F}^*_{A^+,A^-}$.

Proposition 4.3.1. *Let* $\emptyset = A^+ \cup A^-$ *defined as in* (4.1) *and* $\beta \in \mathbb{R}^n$. *It holds that*

1. $\mathcal{F}^*_{A^+,\beta} \subseteq \mathcal{S}_{A^+,\beta}$,
2. $\mathcal{F}^*_{A^+,A^-} \subseteq \mathcal{S}_{A^+,A^-}$.

In particular, $\mathcal{F}^*_{A^+,\beta}$ and $\mathcal{F}^*_{A^+,A^-}$ are subcones of the nonnegativity cone. The constructed decomposition of a SAGE exponential $|A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle}$ with $v \in \mathcal{S}^*_{A^+,A^-} \cap \mathbb{R}^{\mathcal{A}}_{\neq 0}$ into AGE exponentials $\sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + v_\beta e^{\langle \beta, x\rangle}$ with $v \in \mathcal{S}^*_{A^+,\beta}$ for every $\beta \in A^-$ is a naive decomposition into AGE exponentials, and, hence, may not be the best possible decomposition providing the best possible lower bound.

Remark 4.3.2. Note that in [Dre+20] (as published in the Proceedings of ISSAC 2020), this statement was framed differently. In this paper, we claimed that the function $f(x) = \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle}$ is contained in $\mathcal{S}_{A^+,A^-}$ for every (non-scaled) vector of coefficients $v \in \mathcal{S}^*_{A^+,A^-}$. This is indeed true for $|A^-| = 1$. However, in 2021, Janin Heuer and Timo de Wolff pointed out that in general this is not true for sets A^- with cardinality bigger than 1 and proposed a partial solution. All following results, including the optimization problem in the next subsection, are adjusted accordingly and hence differ from those in [Dre+20].

Proof of Proposition 4.3.1.

1. Let $f \in \mathcal{F}^*_{A^+,\beta}$ with a corresponding vector of coefficients $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$. Per definition, we have $v \in \mathcal{S}^*_{A^+,\beta} \cap \mathbb{R}^{A^+ \cup \{\beta\}}_{\neq 0}$. Using representation (2) of Theorem 4.1.6, we have $v_\alpha > 0$ for all $\alpha \in A^+$, and for all $\lambda \in \Lambda(A^+, \beta)$, it holds that

$$\ln(|v_\beta|) \leq \sum_{\alpha \in A^+} \lambda_\alpha \ln(v_\alpha) \leq \sum_{\alpha \in A^+, \lambda_\alpha > 0} \lambda_\alpha \ln\left(\frac{v_\alpha}{\lambda_\alpha}\right).$$

The last inequality holds, as $\lambda_\alpha \in [0,1]$ for every $\alpha \in A^+$ and the logarithmic function is monotonically increasing. Thus, $-v_\beta = |v_\beta| \leq \prod_{\alpha \in A^+, \lambda_\alpha > 0} \left(\frac{v_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha}$. Applying Theorem 4.1.4, we obtain the claimed result.

2. By the Definitions 4.1.1, 4.1.5 and part (1), we obtain

$$\mathcal{F}^*_{A^+,A^-} \subseteq \sum_{\beta \in A^-} \mathcal{F}^*_{A^+,\beta} \subseteq \sum_{\beta \in A^-} \mathcal{S}_{A^+,\beta} \subseteq \mathcal{S}_{A^+,A^-}.$$

The last two sets are equivalent whenever $A^- \neq \emptyset$.

□

We remark that the reverse implication does not hold in general.

Example 4.3.3. Consider the function $f(x) := 1 - 2e^x + e^{2x}$ with $A^+ = \{0, 2\}$, $\beta = 1$ and $v_0 = v_2 = 1, v_1 = -2$. As

$$1 = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 2 \text{ and } -v_\beta = |v_\beta| = (2^{1/2})^2,$$

we have $f \in \mathcal{S}_{A^+,\beta}$. But since

$$\sum_{\alpha \in A^+} \lambda_\alpha \ln(v_\alpha) = 2\left(\frac{1}{2}\ln(1)\right) = 0 < \ln(2) = \ln(|v_1|),$$

it follows that $f \notin \mathcal{F}^*_{A^+,\beta}$.

4.3.2 Formulation of the Optimization Problem

In this subsection, we obtain a computationally fast approximation of the global optimization problem

$$\inf_{x \in \mathbb{R}^n} f(x) \tag{4.22}$$

for exponential sums $f \in \mathbb{R}^{\mathcal{A}}$, $\mathcal{A} = A^+ \cup A^-$ defined as in (4.1) via the representations of the dual SAGE-cone in Theorem 4.1.6.

First, we prove that deciding membership in the cone $\mathcal{F}^*_{A^+,A^-}$ can be done via linear programming.

Proposition 4.3.4. *Let*

$$f \;=\; |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle}$$

with $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$ *and* $\mathrm{Vert}\,(\mathrm{conv}(\mathcal{A})) \subseteq A^+$.

The following linear feasibility program in $|A^-|$ *many variables* $(\tau^{(\beta)})_{\beta \in A^-}$ *verifies containment in the cone* $\mathcal{F}^*_{A^+,A^-}$*:*

$$\ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T \tau^{(\beta)} \;\; \text{for all} \;\; \beta \in A^-,\; \alpha \in A^+ \tag{4.23}$$

The condition $\mathrm{Vert}\,(\mathrm{conv}(\mathcal{A})) \subseteq A^+$ in particular ensures $A^+ \neq \emptyset$ or $A^- = \emptyset$. The case of a coefficient vector with $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$ is chosen because this is the interesting case for optimization purposes: Whenever we have $v_\alpha = 0$ for some $\alpha \in \mathcal{A}$ we can consider the reduced support set $\mathcal{A} \setminus \{\alpha\}$.

Proof. $\mathcal{F}^*_{A^+,A^-}$ is the cone of all functions $f = |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle}$ where $v \in \mathcal{S}^*_{A^+,A^-}$. Hence, we only need to show that (4.23) verifies containment of v in the dual SAGE-cone $\mathcal{S}^*_{A^+,A^-} \cap \mathbb{R}^{\mathcal{A}}_{\neq 0}$.

The program checks the conditions of Theorem 4.1.6 (3). Due to (2.9), the assumptions $\mathrm{Vert}\,(\mathrm{conv}(\mathcal{A})) \subseteq A^+$ on f are indeed necessary. As $v \in \mathbb{R}^{\mathcal{A}}$ is fixed, the inequalities are linear and, hence, (4.23) is a linear program. Moreover, $v_\alpha > 0$ for every $\alpha \in A^+$ holds by assumption (or we know trivially that v does not belong to the dual SAGE-cone). The last inequalities in Theorem 4.1.6(3) are satisfied trivially. $\square$

In particular, fixing the non-auxiliary variables v in a lifted version of the dual cone forms a polyhedron; see Theorem 4.1.6 and Proposition 4.3.4.

To show that Proposition 4.3.4 can be used to obtain an exact linear optimization problem over the cone $\mathcal{F}^*_{A^+,A^-}$, recall from (1.1) that equivalently to (4.22), we can solve the optimization problem

$$\sup\left\{\gamma \;:\; f(x) - \gamma \geq 0 \;\; \text{for all} \;\; x \in \mathbb{R}^n\right\}.$$

Instead of using containment in the SAGE-cone as a certificate for nonnegativity, i.e., solving

$$\sup\{\gamma \ : \ f(x) - \gamma \in \mathcal{S}_{A^+,A^-}\},$$

we use the cone $\mathcal{F}^*_{A^+,A^-}$. Recall that $\mathcal{F}^*_{A^+,A^-} \subseteq \mathcal{S}_{A^+,A^-}$ by Proposition 4.3.1. In particular, we do not dualize the primal optimization problem to approximate the solution but optimize f to be a function in $\mathcal{F}^*_{A^+,A^-}$ instead of the primal cone. Hence, we compute

$$f_{\text{sage dual}} = \sup\left\{\check{\gamma} \ : \ f - \check{\gamma} \in \mathcal{F}^*_{A^+,A^-}\right\}, \tag{4.24}$$

i.e., $v_{\mathbf{0}} - \check{\gamma}$ is the coefficient corresponding to the constant term if $\mathbf{0} \in A^-$ and $|A^-|v_{\mathbf{0}} - \check{\gamma}$ is the coefficient corresponding to the constant term if $\mathbf{0} \in A^+$.

Consider v to be given via

$$\begin{aligned} f(x) - \check{\gamma} &= |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x\rangle} - \check{\gamma} \\ &= |A^-| \sum_{\alpha \in A^+ \setminus \{\mathbf{0}\}} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\beta \in A^- \setminus \{\mathbf{0}\}} v_\beta e^{\langle \beta, x\rangle} + \begin{cases} (|A^-|v_{\mathbf{0}} - \check{\gamma}) \text{ if } \mathbf{0} \in A^+, \\ (v_{\mathbf{0}} - \check{\gamma}) \text{ if } \mathbf{0} \in A^- \end{cases} \end{aligned}$$

with $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$. Note that the constant term $v_{\mathbf{0}}$ of $f(x)$ can be zero (whenever it is not contained in $\mathcal{A}$). As we need to know the cardinality of A^-, we need to decide which set to include $\mathbf{0}$ in *before* the computation. For $A^- \neq \emptyset$, by Theorem 4.1.6(3), and assuming

$$w_{\mathbf{0}} := \begin{cases} -\check{\gamma} + v_{\mathbf{0}} \text{ if } \mathbf{0} \in A^-(f - \check{\gamma}), \\ (-1/|A^-|)\check{\gamma} + v_{\mathbf{0}} \text{ if } \mathbf{0} \in A^+(f - \check{\gamma}), \end{cases}$$

and $w_\alpha := v_\alpha$ for all $\alpha \in \mathcal{A} \setminus \{\mathbf{0}\}$, solving (4.24) is equivalent to solving

$$\max\left\{\check{\gamma} : \begin{array}{l} \forall\, \alpha \in A^+, w_\alpha > 0, \forall\, \beta \in A^-, w_\beta < 0; \text{ and } \forall\, \beta \in A^-(f - \check{\gamma}) \\ \exists\, \tau \in \mathbb{R}^n \text{ s.t. } \ln\left(\frac{|w_\beta|}{w_\alpha}\right) \leq (\alpha - \beta)^T \tau \forall\, \alpha \in A^+(f - \check{\gamma}) \end{array}\right\}. \tag{4.25}$$

Note that, as we assumed $A^- \neq \emptyset$, even $w_{\mathbf{0}}$ is well defined.

Before stating the corresponding optimization program, we emphasize the fact that $\mathbf{0}$ is not necessarily contained in $\mathcal{A}$ and we do not know the sign of $w_{\mathbf{0}}$ before having solved (4.25). Hence, for the next result we need to include it either in A^+ or A^-, although we have to determine later which one of the sets it belongs to. In particular, as elements $\alpha \in A^+$ have coefficients $|A^-|v_\alpha$, with $v \in \mathcal{S}^*_{A^+,A^-}$, before including it in either of those sets, we do not know if we have to scale it with $|A^-|$.

First, we prove several statements addressing this choice.

Lemma 4.3.5. *Let $\mathcal{A} = A^+ \cup A^- \subseteq \mathbb{R}^n$ as in* (4.1) *and $f \in \mathcal{F}^*_{A^+,A^-}$ with $\mathbf{0} \in \mathcal{A}$. If f has exponents $\mathcal{A} \subseteq \mathbb{N}^n$, then $\mathbf{0} \in A^+$.*

In particular, this yields that, whenever we optimize polynomials using the dual SONC-cone, we always have $\mathbf{0} \in A^+$, simplifying this technique substantially.

Proof. As $\mathcal{A} \subseteq \mathbb{N}^n$ and $\mathbf{0} \in \mathcal{A}$, we necessarily have $\mathbf{0} \in \mathrm{Vert}(\mathrm{conv}(\mathcal{A}))$. With (2.9) and the fact that $\mathcal{F}^*_{A^+,A^-} \subseteq \mathcal{S}_{A^+,A^-}$, we obtain the statement. □

Lemma 4.3.6. *Let $\mathcal{A} = A^+ \cup A^- \subseteq \mathbb{R}^n$ as in (4.1) and $f \in \mathcal{F}^*_{A^+,A^-}$ with coefficient vector $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0} \cap \mathcal{S}^*_{A^+,A^-}$. For the optimal lower bound $f_{sage\ dual} \leq f(x)$ (as defined in (4.24)) and $\mathbf{0} \in \mathcal{A}(f - f_{sage\ dual})$, we have*

$$f_{sage\ dual} \;=\; \begin{cases} |A^-|(v_{\mathbf{0}} - e^{c^*}) & \text{if } \mathbf{0} \in A^+(f - f_{sage\ dual}) \text{ and} \\ v_{\mathbf{0}} + e^{c^*} & \text{if } \mathbf{0} \in A^-(f - f_{sage\ dual}) \end{cases} \tag{4.26}$$

with $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$ and where

$$c^* = \begin{cases} \ln(|v_{\mathbf{0}} - (1/|A^-|) f_{sage\ dual}|) & \text{if } \mathbf{0} \in A^+(f - f_{sage\ dual}) \text{ and} \\ \ln(|v_{\mathbf{0}} - f_{sage\ dual}|) & \text{if } \mathbf{0} \in A^-(f - f_{sage\ dual}). \end{cases}$$

Proof. If $\mathbf{0} \in A^+(f - f_{\text{sage dual}})$, the corresponding coefficient in $f - f_{\text{sage dual}}$ is positive, i.e., $v_{\mathbf{0}} - (1/|A^-|)f_{\text{sage dual}} > 0$ implying $e^{c^*} = v_{\mathbf{0}} - (1/|A^-|)f_{\text{sage dual}}$. Hence, $f_{\text{sage dual}} = |A^-|(v_{\mathbf{0}} - e^{c^*})$.

If $\mathbf{0} \in A^-(f - f_{\text{sage dual}})$, the corresponding coefficient is negative, i.e., we have $v_{\mathbf{0}} - f_{\text{sage dual}} < 0$ implying $-e^{c^*} = -|v_{\mathbf{0}} - f_{\text{sage dual}}| = v_{\mathbf{0}} - f_{\text{sage dual}}$. This yields the statement. □

From now on, for $\mathcal{A} = A^+ \cup A^-$ defined as in (4.1) with $A^+ \neq \emptyset, A^- \neq \emptyset$ and a fixed function

$$|A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x \rangle} + \sum_{\beta \in A^-} v_\alpha e^{\langle \beta, x \rangle}$$

where for all $\alpha \in \text{Vert}\,(\text{conv}(\mathcal{A}))$ we have $v_\alpha > 0$ (i.e., f satisfies (2.9)), with lower bound $f_{\text{sage dual}}$ and $\mathbf{0} \in \mathcal{A}(f - f_{\text{sage dual}})$, we consider the following two linear programs in $|A^-| + 1$ variables $(\tau^{(\beta)})_{\beta \in A^-}$ and $c = \ln(|w_{\mathbf{0}}|)$, where the optimization variable $w_{\mathbf{0}}$ is defined as in the representation (4.25).

$$\begin{array}{lll} & \max\ c & \\ \text{s.t.} & (1) & \text{for all } \beta \in A^-, \text{ for all } \alpha \in A^+ \setminus \{\mathbf{0}\} : \ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T \tau^{(\beta)}, \\ & (2) & \ln(|v_\beta|) - c \leq (-\beta)^T \tau^{(\beta)} \text{ for all } \beta \in A^- \end{array} \tag{LP$_{A^+}$}$$

if $\mathbf{0} \in A^+(f - \check{\gamma})$ and

$$\begin{array}{lll} & \max\ c & \\ \text{s.t.} & (1) & \text{for all } \beta \in (A^-) \setminus \{\mathbf{0}\}, \text{ for all } \alpha \in A^+ : \ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T \tau^{(\beta)}, \\ & (2) & c - \ln(v_\alpha) \leq \alpha^T \tau^{(\mathbf{0})} \text{ for all } \alpha \in A^+ \end{array} \tag{LP$_{A^-}$}$$

if $\mathbf{0} \in A^-(f - \check{\gamma})$.

Lemma 4.3.7. *Let*

$$f \;= |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x \rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x \rangle},$$

with $\mathcal{A} = A^+ \cup A^-$ defined as in (4.1), $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$ and such that for the optimal lower bound $f_{sage\ dual} \leq f$ at least one of $v_{\mathbf{0}} - f_{sage\ dual} < 0$ or $|A^-| v_{\mathbf{0}} - f_{sage\ dual} > 0$ holds. At least one of the linear programs (LP$_{A^+}$) and (LP$_{A^-}$) has a solution for its corresponding assumption

1. $\mathbf{0} \in A^+$ *or*

2. $\mathbf{0} \in A^-$,

if and only if there exists some $\check{\gamma} \in \mathbb{R}$ *such that* $f - \check{\gamma} \in \mathcal{F}^*_{A^+,A^-}$.

For either assumption, the corresponding LP is infeasible if and only if for all $\check{\gamma} \in \mathbb{R}$ *we have* $f - \check{\gamma} \notin \mathcal{F}^*_{A^+,A^-}$.

Proof. Consider $f - f_{\text{sage dual}}$ for the optimal lower bound $f_{\text{sage dual}}$ using the cone $\mathcal{F}^*_{A^+,A^-}$. By assumption, the constant coefficient of $f - f_{\text{sage dual}}$ is nonzero, yielding $\mathbf{0} \in \mathcal{A}(f - f_{\text{sage dual}})$. Hence, the inequalities are exactly the inequalities in Theorem 4.1.6, except for the fact that we use c instead of $\ln(v_{\mathbf{0}})$ due to the former substitution. □

We need to omit two cases here:

1. The case of a zero constant coefficient in the exponential sum $f - f_{\text{sage dual}}$: In this case ,we clearly have $\mathbf{0} \notin \mathcal{A}$, $f_{\text{sage dual}} = 0$ and the programs (1) and (2) in Lemma 4.3.7 are infeasible and unbounded, respectively. To still obtain a lower bound on the function f, one can verify containment in the subcone $\mathcal{F}^*_{A^+,A^-}$ of the nonnegativity cone by testing feasibility via (4.23). If f is indeed an element in $\mathcal{F}^*_{A^+,A^-}$, then 0 is always a lower bound, but not necessarily the optimal bound on $\mathcal{F}^*_{A^+,A^-}$.
2. The case of $|A^-|v_{\mathbf{0}} - f_{\text{sage dual}} \leq 0 \leq v_{\mathbf{0}} - f_{\text{sage dual}}$ with at least one of the terms on the left and right-hand side of the inequality chain nonzero: By assumption, $A^- \neq \emptyset$ and, hence, at least one of the terms on the left and right-hand side of the inequality chain can only be nonzero if $|A^-| > 1$. Clearly $v_{\mathbf{0}} \leq 0$ as well as $f_{\text{sage dual}} \leq 0$. In this case, we can assume $\mathbf{0} \in A^+$ leading to a lower bound $\tilde{f} < f_{\text{sage dual}}$ with $|A^-|v_{\mathbf{0}} - \tilde{f} \geq 0 \geq |A^-|v_{\mathbf{0}} - f_{\text{sage dual}}$, which might again not be the optimal bound on $\mathcal{F}^*_{A^+,A^-}$, but is a lower bound in any case.

From the considerations above and Proposition 4.3.4 we draw the following result.

Theorem 4.3.8. *Let*

$$f \ = |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x \rangle} + \sum_{\beta \in A^-} v_\beta e^{\langle \beta, x \rangle},$$

with $\mathcal{A} = A^+ \cup A^-$ *defined as in* (4.1)*and* $v \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$. *Moreover, let* $v_{\mathbf{0}} - f_{\textit{sage dual}} < 0$ *or* $|A^-|v_{\mathbf{0}} - f_{\textit{sage dual}} > 0$ *for the optimal lower bound* $f_{\textit{sage dual}} \leq f$. *The linear programs* (LP_{A^+}) *and* (LP_{A^-}) *solve the optimization problem* (4.25).

Proof. Let $f_{\text{sage dual}}$ be the optimal value with $f \geq f_{\text{sage dual}}$ as defined in (4.24). We set $\mathcal{A} := \mathcal{A} \cup \{\mathbf{0}\}$. First, note that we do not know the value of the constant term before computing the optimal value, and, in particular, we do not know its sign or whether it needs to be scaled with $|A^-|$. Thus, we cannot determine whether $\mathbf{0} \in A^+(f - f_{\text{sage dual}})$ or $\mathbf{0} \in A^-(f - f_{\text{sage dual}})$ before computing the optimal value.

By the assumption on the constant term in $(f - f_{\text{sage dual}})$, according to (4.3.7), at least one of the problems (LP_{A^+}) and (LP_{A^-}) is feasible if and only if there exists $\check{\gamma} \in \mathbb{R}$ such that $f - \check{\gamma} \cdot e^{\langle \mathbf{0}, x \rangle} \in \mathcal{F}^*_{A^+,A^-}$. In the case that only one linear program is feasible, $\mathbf{0}$ is contained in the corresponding set and, hence, this program yields the optimal value. If both programs are feasible, there exist $\check{\gamma}_1$ and $\check{\gamma}_2$ such that $v_{\mathbf{0}} - (1/|A^-|)\check{\gamma}_1$ is nonnegative and $f - \check{\gamma}_1 \cdot e^{\langle \mathbf{0}, x \rangle} \in \mathcal{F}^*_{A^+,A^-}$ for $\mathbf{0} \in A^+(f - \check{\gamma}_1)$, and $v_{\mathbf{0}} - \check{\gamma}_2$ is negative and $f - \check{\gamma}_2 \cdot e^{\langle \mathbf{0}, x \rangle} \in \mathcal{F}^*_{A^+,A^-}$ for $\mathbf{0} \in A^-(f - \check{\gamma}_2)$.

Thus, we select the linear program which yields the better bound.
According to Lemma 4.3.6, the lower bound on the dual SAGE-cone is

$$f_{\text{sage dual}} = \begin{cases} |A^-|(v_{\mathbf{0}} - e^{c^*}) & \text{if } \mathbf{0} \in A^+(f - f_{\text{sage dual}}) \\ v_{\mathbf{0}} + e^{c^*} & \text{if } \mathbf{0} \in A^-(f - f_{\text{sage dual}}). \end{cases} \tag{4.27}$$

□

Note that optimizing over the dual cone does not yield the actual optimal value in every case. Consider for example the exponential version of the Motzkin polynomial

$$f(x, y) = e^{2x+4y} + e^{4x+2y} - 3e^{2x+2y} + 1. \tag{4.28}$$

This is a nonnegative exponential sum on $\mathbb{R}^2$ with $\inf_{(x,y)\in\mathbb{R}^2} f(x, y) = 0$. Since in the polynomial case we always need $\mathbf{0} \in A^+$, it suffices to solve (LP$_{A^+}$), and since $|A^-| = |\{(2, 2)^T\}| = 1$, the linear program (LP$_{A^+}$) for f is the following:

$$\begin{aligned} & \max\ c \\ \text{s.t.}\quad & \ln(3) \leq 2\tau_2 \\ & \ln(3) \leq 2\tau_1 \\ & \ln(3) + 2\tau_1 + 2\tau_2 \leq c \end{aligned}$$

This LP returns the lower bound $f \geq -26$ on $\mathbb{R}^2$.

Chapter 5

The $\mathcal{S}$-Cone and a Primal-Dual View on Second-Order Representability

In this chapter, we present a cone that provides a common framework for recent nonnegativity certificates of sparse polynomials and exponential sums.

In Section 5.1, we start by introducing this cone called $\mathcal{S}$-cone. Its most basic elements are called *even and odd AG-functions*, see Definition 5.1.2. An important property of the $\mathcal{S}$-cone is that both the cone of sums of nonnegative circuit polynomials and of sums of arithmetic-geometric exponentials are special cases of it, see Remark 5.1.4.

In Section 5.2, we show that when examining the $\mathcal{S}$-cone, we can make use of the fact that fundamental properties of the cone of SAGE exponentials and SONC polynomials established in Chapter 3 — such as the projection-free characterization of the dual cone as well as the exact characterization of the extreme rays — also hold in the more general context of the $\mathcal{S}$-cone.

In Section 5.3, we focus on second-order representations of both the primal and the dual cone. In doing so, we extend results previously obtained by Averkov and Wang and Magron, who already examined the semidefinite extension degree [Ave19] and the second-order representability of the *primal* cone of *SONC polynomials* [WM20a].

For both our derivations of second-order representations of the primal and dual $\mathcal{S}$-cone, we define what we call *(dual) circuit matrices*, see Definition 5.3.2 and 5.3.11, show how to obtain second-order representations for both the dual and the primal cone of AG functions, see Theorems 5.3.5 and 5.3.12, and discuss the sizes of both resulting second-order-cone programs, see Corollaries 5.3.7 and 5.3.14.

We conclude this chapter with extending the results on second-order representations of the primal and dual cone of AG functions to the respective primal and dual of the $\mathcal{S}$-cone, see Corollaries 5.3.18 and 5.3.19. In both cases, we can take advantage of the extremality theory established in Chapter 3, namely, using *reduced circuits* and, thus, substantially reducing the size of the corresponding programs.

5.1 The $\mathcal{S}$-Cone

We start by introducing the $\mathcal{S}$-cone as a unified framework of SONC polynomials and SAGE exponentials. Our main object of study are functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ of the form

$$f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta \tag{5.1}$$

where $\mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ are finite sets of exponents, $\{c_\alpha : \alpha \in \mathcal{A}\}$ and $\{d_\beta : \beta \in \mathcal{B}\} \subseteq \mathbb{R}$. Here, we use the notations

$$|x|^\alpha = \prod_{j=1}^{n} |x_j|^{\alpha_j} \qquad \text{and} \qquad x^\beta = \prod_{j=1}^{n} x_j^{\beta_j},$$

and if one component of x is zero and the corresponding exponent is negative, then we set $|x|^\alpha = \infty$.

For two finite sets $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n, \mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$, let

$$\mathbb{R}[\mathcal{A}, \mathcal{B}] := \operatorname{span}_{\mathbb{R}}(\{|x|^\alpha \mid \alpha \in \mathcal{A}\} \cup \{x^\beta \mid \beta \in \mathcal{B}\})$$

denote the space of all functions of the form (5.1) with given sets of exponents. This is a vector space of dimension $\dim \mathbb{R}[\mathcal{A}, \mathcal{B}] = |\mathcal{A}| + |\mathcal{B}|$.

Remark 5.1.1.

1. If $\mathcal{A} \subseteq (2\mathbb{N})^n$, then $\mathbb{R}[\mathcal{A}, \mathcal{B}]$ is exactly the space of polynomials with exponent vectors in $\mathcal{A} \cup \mathcal{B}$. For this reason, we sometimes refer to elements of $\mathcal{A}$ as *even exponents* and to elements of $\mathcal{B}$ as *odd exponents*.

2. If $\mathcal{B} = \emptyset$, then $\mathbb{R}[\mathcal{A}, \mathcal{B}]$ can be identified with the space of exponential sums

$$y \mapsto \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, y \rangle}$$

via the identification $|x_i| = e^{y_i}$.

3. In fact, it is no restriction to exclude sets in $(2\mathbb{N})^n$ from $\mathcal{B}$ since for exponents $\beta \in (2\mathbb{N})^n$, we have $|x|^\beta = x^\beta$.

4. $\mathcal{A}$ and $\mathcal{B}$ are not necessarily disjoint (compare Example 5.1.9 below).

We study the nonnegativity of functions in $\mathbb{R}[\mathcal{A}, \mathcal{B}]$ using the following building blocks:

Definition 5.1.2. Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n, \mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets and let $f = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta$. We say that f is

1. an *even AG function* if at most one of the c_α is negative and all the d_β are zero; and

2. an *odd AG function* if all the c_α are nonnegative and at most one of the d_β is nonzero.

A function f is called an *AG function* (*arithmetic-geometric mean function*) if f is an even AG function or an odd AG function.

Note that nonnegative even AG functions correspond exactly to the AGE exponentials studied in previous chapters.

We arrive at the central definition of this section.

Definition 5.1.3 ($\mathcal{S}$-Cone)**.** Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets. The $\mathcal{S}$*-cone* $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ is defined as

$$C_{\mathcal{S}}(\mathcal{A}, \mathcal{B}) := \operatorname{pos}(f \in \mathbb{R}[\mathcal{A}, \mathcal{B}] \mid f \text{ is a nonnegative AG function}).$$

Remark 5.1.4.

1. If $\mathcal{B} = \emptyset$, then the $\mathcal{S}$-cone can be identified with the cone of SAGE exponentials using the substitution in Remark 5.1.1(2).

2. If $\mathcal{A} \subseteq (2\mathbb{N})^n$, then $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ is the cone of SONC polynomials supported on $\mathcal{A} \cup \mathcal{B}$ (2.11). As stated in the preliminaries, this cone was initially defined in terms of circuit polynomials. The equivalence of the definitions (see Chapter 2) was established in [Wan18a] and [MCW21a] and also follows from our more general result in Proposition 5.2.6.

3. An example where the cone $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ is different from both the cone of SAGE exponentials and the cone of SONC polynomials is given by $\mathcal{A} = \{1, 4\}$ and $\mathcal{B} = \{3\}$. Here, the $\mathcal{S}$-cone is the set of all nonnegative functions of the form $f = c_1|x| + c_4|x|^4 + c_3x^3$ (where f being nonnegative trivially implies $c_1 \geq 0$ and $c_4 \geq 0$). Functions of these forms cannot be treated as polynomials, as $c_1|x|$ is globally nonnegative but the polynomial version c_1x of this term is not. Also, we cannot treat any such f as an exponential sum, as x^3 *is not* globally nonnegative while its exponential version e^{3x} *is* globally nonnegative.

For a non-empty finite set $\mathcal{A} \subseteq \mathbb{R}^n$ and $\beta \in \mathbb{N}^n \setminus (2\mathbb{N})^n$ let

$$P^{\text{odd}}_{\mathcal{A},\beta} := \left\{ f \;\middle|\; f = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + dx^\beta, f(x) \geq 0 \;\forall\, x \in \mathbb{R}^n, c \in \mathbb{R}^{\mathcal{A}}_+, d \in \mathbb{R} \right\}$$

be the cone of nonnegative odd AG functions supported on $(\mathcal{A}, \beta)$, and similarly for $\beta \in \mathbb{R}^n \setminus \mathcal{A}$ let

$$P^{\text{even}}_{\mathcal{A},\beta} := \left\{ f \;\middle|\; f = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + d|x|^\beta, f(x) \geq 0 \;\forall\, x \in \mathbb{R}^n, c \in \mathbb{R}^{\mathcal{A}}_+, d \in \mathbb{R} \right\}$$

be the cone of nonnegative even AG functions supported on $(\mathcal{A}, \beta)$. Note that, by definition,

$$C_{\mathcal{S}}(\mathcal{A}, \mathcal{B}) = \sum_{\alpha \in \mathcal{A}} P^{\text{even}}_{\mathcal{A}\setminus\{\alpha\},\alpha} + \sum_{\beta \in \mathcal{B}} P^{\text{odd}}_{\mathcal{A},\beta}. \tag{5.2}$$

This implies the following alternative representation of the $\mathcal{S}$-cone.

Proposition 5.1.5. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite and $\delta^{(\alpha)}$ denote the unit vector in $\mathbb{R}^{\mathcal{A}\cup\mathcal{B}}$ indexed with $\alpha \in \mathcal{A} \cup \mathcal{B}$. Then, $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ equals the set*

$$\left\{ \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta \in \mathbb{R}[\mathcal{A}, \mathcal{B}] : \sum_{\alpha \in \mathcal{A}} c_\alpha \cdot \delta^{(\alpha)} - \sum_{\beta \in \mathcal{B}} |d_\beta| \cdot \delta^{(\beta)} \in C_{SAGE}(\mathcal{A} \cup \mathcal{B}) \right\}$$

$$= \Bigg\{ \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta \in \mathbb{R}[\mathcal{A}, \mathcal{B}] : \text{ there exists } t \in \mathbb{R}^{\mathcal{B}} \text{ such that}$$

$$\sum_{\alpha \in \mathcal{A}} c_\alpha \cdot \delta^{(\alpha)} + \sum_{\beta \in \mathcal{B}} t_\beta \cdot \delta^{(\alpha)} \in C_{SAGE}(\mathcal{A} \cup \mathcal{B}),\; t_\beta \leq -|d_\beta| \text{ for all } \beta \in \mathcal{B} \Bigg\}.$$

If $\mathcal{A} \cap \mathcal{B} = \emptyset$, we can shortly write $\sum_{\alpha \in \mathcal{A}} c_\alpha \cdot \delta^{(\alpha)} - \sum_{\beta \in \mathcal{B}} |d_\beta| \cdot \delta^{(\beta)} = (c, -|d|)$, where $|d|$ denotes the component-wise absolute value. If there exists some $\beta \in \mathcal{A} \cap \mathcal{B}$, then the corresponding coefficient in the SAGE-cone $c_\beta - |d_\beta|$ appears only once in the set $\mathbb{R}^{\mathcal{A}\cup\mathcal{B}}$. However, by slight abuse of notation, we also write $\sum_{\alpha\in\mathcal{A}} c_\alpha \cdot \delta^{(\alpha)} - \sum_{\beta\in\mathcal{B}} |d_\beta| \cdot \delta^{(\beta)}$ shortly as $(c, -|d|)$.

Proof. If $f = \sum_{\alpha\in\mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta\in\mathcal{B}} d_\beta x^\beta \in C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$, then (5.2) gives a decomposition

$$f = \sum_{\alpha\in\mathcal{A}} f_\alpha^{\text{even}} + \sum_{\beta\in\mathcal{B}} f_\beta^{\text{odd}}$$

with $f_\alpha^{\text{even}} \in P_{\mathcal{A}\setminus\{\alpha\},\alpha}^{\text{even}}$ for all $\alpha \in \mathcal{A}$ and $f_\beta^{\text{odd}} = \sum_{\alpha\in\mathcal{A}} c_\alpha^{(\beta)} |x|^\alpha + d_\beta x^\beta \in P_{\mathcal{A},\beta}^{\text{odd}}$ for every $\beta \in \mathcal{B}$. Defining the functions

$$\begin{aligned} \tilde{f}_\alpha^{\text{even}} &= f_\alpha^{\text{even}} \text{ for all } \alpha \in \mathcal{A} \\ \text{and } \tilde{f}_\beta^{\text{even}} &= \sum_{\alpha\in\mathcal{A}} c_\alpha^{(\beta)} |x|^\alpha - |d_\beta||x|^\beta = f_\beta^{\text{odd}} - d_\beta x^\beta - |d_\beta||x|^\beta \text{ for all } \beta \in \mathcal{B}, \end{aligned}$$

symmetry implies $\tilde{f}_\beta^{\text{even}} \in P_{\mathcal{A},\beta}^{\text{even}}$. Hence, $\tilde{f} = \sum_{\alpha\in\mathcal{A}} \tilde{f}_\alpha^{\text{even}} + \sum_{\beta\in\mathcal{B}} \tilde{f}_\beta^{\text{even}} \in C_{\mathcal{S}}(\mathcal{A}\cup\mathcal{B},\emptyset)$. Remark 5.1.4(1) then shows that $\tilde{f} \in C_{\text{SAGE}}(\mathcal{A}\cup\mathcal{B})$.

The converse direction of the first equation follows immediately with the substitution in Remark 5.1.1(2).

The second equation, which exhibits the convexity of the $\mathcal{S}$-cone, is an immediate consequence of the first one. □

In our definition of the $\mathcal{S}$-cone, we exclude sums of nonnegative AG functions with support $\mathcal{A}\cup\mathcal{B}$ for $\mathcal{A} \subseteq \mathbb{R}^n, \mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ where the corresponding AG functions have bigger support than $\mathcal{A}\cup\mathcal{B}$. This could happen, for example, if two summands cancel in the sum. For a better understanding of the problem, consider the following example.

Example 5.1.6. Let $\mathcal{A} := \{\frac{1}{3}, \frac{7}{3}\}, \mathcal{B} := \{1\}$. Consider the two nonnegative AG functions

$$\begin{aligned} f_1 &:= |x|^{\frac{1}{3}} + x + x^2, \\ f_2 &:= |x|^{\frac{1}{3}} - x^2 + |x|^{\frac{7}{3}}, \end{aligned}$$

whose support is not contained in $\mathcal{A}\cup\mathcal{B}$. But the sum

$$f := f_1 + f_2 = 2|x|^{\frac{1}{3}} + x + |x|^{\frac{7}{3}}$$

is itself a nonnegative AG function whose support is contained in $\mathcal{A}\cup\mathcal{B}$.

As observed for SONC polynomials and SAGE exponentials in Chapter 2, this restriction is not really a restriction. The following proposition states that every sum f of nonnegative AG functions whose support is bigger than the support of the sum can be decomposed into a sum of nonnegative AG functions whose supports are contained in the support of f.

For the case of SAGE exponentials, see Theorem 2.4.8, and for the case of SONC polynomials see Theorem 2.4.7.

Proposition 5.1.7. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n, \mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets and $f \in \mathbb{R}[\mathcal{A},\mathcal{B}]$. If $f \in C_{\mathcal{S}}(\mathcal{A}',\mathcal{B}')$ for some $\mathcal{A}' \supseteq \mathcal{A}$, $\mathbb{N}^n \setminus (2\mathbb{N})^n \supseteq \mathcal{B}' \supseteq \mathcal{B}$, then $f \in C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$ as* *well. Equivalently, it holds that*

$$C_{\mathcal{S}}(\mathcal{A},\mathcal{B}) = C_{\mathcal{S}}(\mathcal{A}',\mathcal{B}') \cap \mathbb{R}[\mathcal{A},\mathcal{B}].$$

Proof. By Proposition 5.1.5, the $\mathcal{S}$-cone $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$ equals the set

$$\left\{ \sum_{\alpha\in\mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta\in\mathcal{B}} d_\beta x^\beta \in \mathbb{R}[\mathcal{A},\mathcal{B}] : \sum_{\alpha\in\mathcal{A}} c_\alpha \cdot \delta^{(\alpha)} - \sum_{\beta\in\mathcal{B}} |d_\beta| \cdot \delta^{(\beta)} \in C_{\mathrm{SAGE}}(\mathcal{A}\cup\mathcal{B}) \right\},$$

with $\delta^{(\alpha)}$ denoting the unit vector with respect to α for $\alpha \in \mathcal{A}$ or $\mathcal{B}$.

Now, let $f \in C_{\mathcal{S}}(\mathcal{A}',\mathcal{B}') \cap \mathbb{R}[\mathcal{A},\mathcal{B}]$ with a vector of coefficients (c,d) and, hence, $(c,-|d|) \in C_{\mathrm{SAGE}}(\mathcal{A}'\cup\mathcal{B}') \cap \mathbb{R}^{\mathcal{A}\cup\mathcal{B}}$, where the absolute value is component-wise. The already mentioned statement for the SAGE-case [MCW21a], Theorem 2, states that $C_{\mathrm{SAGE}}(\mathcal{A}\cup\mathcal{B}) = C_{\mathrm{SAGE}}(\mathcal{A}'\cup\mathcal{B}') \cap \mathbb{R}^{\mathcal{A}\cup\mathcal{B}}$. Hence, $(c,-|d|) \in C_{\mathrm{SAGE}}(\mathcal{A}\cup\mathcal{B})$ and, again by Proposition 5.1.5, f is contained in

$$\left\{ \sum_{\alpha\in\mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta\in\mathcal{B}} d_\beta x^\beta \in \mathbb{R}[\mathcal{A},\mathcal{B}] : \sum_{\alpha\in\mathcal{A}} c_\alpha \cdot \delta^{(\alpha)} - \sum_{\beta\in\mathcal{B}} |d_\beta| \cdot \delta^{(\beta)} \in C_{\mathrm{SAGE}}(\mathcal{A}\cup\mathcal{B}) \right\},$$

which equals $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$. The other inclusion is obvious. □

Analogously to SONC polynomials and SAGE exponentials, nonnegative AG functions can be characterized as stated in the subsequent theorem, generalizing [CS17, Lemma 2.2] to the setting of AG functions. To do so, we remind the reader of the definition of the polytope $\Lambda(\mathcal{A},\beta)$ for a non-empty finite set $\mathcal{A} \subseteq \mathbb{R}^n$ and $\beta \in \mathbb{R}^n$:

$$\Lambda(\mathcal{A},\beta) := \left\{ \lambda \in \mathbb{R}_+^{\mathcal{A}} \;\middle|\; \sum_{\alpha\in\mathcal{A}} \lambda_\alpha \alpha = \beta, \; \sum_{\alpha\in\mathcal{A}} \lambda_\alpha = 1 \right\}. \tag{5.3}$$

Again we use the notation $\lambda(\mathcal{A},\beta)$ for the special case that $\mathcal{A}$ is affinely independent, where $\Lambda(\mathcal{A},\beta)$ consists of a single element.

Also, for a non-empty finite set $\mathcal{A} \subseteq \mathbb{R}^n$, we remind the reader of the definition of the *relative entropy function* from Chapter 2: $D : \mathbb{R}_{>0}^{\mathcal{A}} \times \mathbb{R}_{>0}^{\mathcal{A}} \to \mathbb{R}$,

$$D(\nu,\gamma) \;=\; \sum_{\alpha\in\mathcal{A}} \nu_\alpha \ln\left(\frac{\nu_\alpha}{\gamma_\alpha}\right), \quad \nu,\gamma \in \mathbb{R}_{>0}^{\mathcal{A}}$$

with continous extension to $\mathbb{R}_+^{\mathcal{A}} \times \mathbb{R}_+^{\mathcal{A}} \to \mathbb{R}$.

Theorem 5.1.8. *Let $\mathcal{A} \subseteq \mathbb{R}^n$ be a non-empty finite set and f be an AG function of the form*

$$f = \sum_{\alpha\in\mathcal{A}} c_\alpha |x|^\alpha + \begin{cases} d|x|^\beta \text{ with } \beta \in \mathbb{R}^n \setminus \mathcal{A} & \text{if } f \text{ is even,} \\ dx^\beta \text{ with } \beta \in \mathbb{N}^n \setminus (2\mathbb{N})^n & \text{if } f \text{ is odd} \end{cases}$$

where $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$ and $d \in \mathbb{R}$. Then, the following statements are equivalent:

1. *$f(x) \geq 0$ for all $x \in \mathbb{R}^n$.*
2. *There exists a $\nu \in \mathbb{R}_+^{\mathcal{A}}$ such that $\sum_{\alpha\in\mathcal{A}} \nu_\alpha \alpha = (\sum_{\alpha\in\mathcal{A}} \nu_\alpha)\beta$ and*

$$D(\nu, e\cdot c) \leq \begin{cases} d & \text{if } f \text{ even,} \\ -|d| & \text{if } f \text{ odd.} \end{cases}$$

3. There exists a $\lambda \in \Lambda(\mathcal{A}, \beta)$ such that

$$\prod_{\alpha \in \mathcal{A}} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} \geq \begin{cases} -d & \text{if } f \text{ even,} \\ |d| & \text{if } f \text{ odd.} \end{cases}$$

A vector $\lambda \in \Lambda(\mathcal{A}, \beta)$ as in this theorem is called an *AG witness.*

Proof. The statement can be immediately deduced from the proof of Theorem 3.1.1. □

Example 5.1.9. Let $\mathcal{A} = \mathcal{B} = \{1\} \subseteq \mathbb{N}$. A typical AG function with this support is

$$g(x) = c_1|x| + c_2 x.$$

Since the equality condition in statement (2) of Theorem 5.1.8 is trivially satisfied, we have $g(x) \geq 0$ for all $x \in \mathbb{R}$ if and only if there exists a $\nu \in \mathbb{R}_+$ with

$$\nu \ln\left(\frac{\nu}{e c_1}\right) \leq -|c_2|. \tag{5.4}$$

If $\nu \geq 0$, the latter condition can be simplified to $|c_2| \leq c_1$. For the case $\nu = 0$, this is clear from our setting $0 \cdot \ln 0 = 0$, and to see it for $\nu > 0$, rewrite (5.4) as

$$c_1 \left(\frac{\nu}{c_1}\right) \ln\left(\frac{\nu}{e c_1}\right) \leq -|c_2|.$$

Since the function $x \ln\left(\frac{x}{e}\right)$ attains its minimum at $x = 1$, we obtain the claimed result. In particular, it is the one of statement (3) in Theorem 5.1.8.

For later use, we note that our cones of interest are closed:

Proposition 5.1.10. *Let Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n, \mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets and $\beta \in \mathbb{R}^n$. The cones $P^{\text{odd}}_{\mathcal{A},\beta}, P^{\text{even}}_{\mathcal{A},\beta}$ and $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ are closed pointed convex cones.*

Proof. It is clear that all three cones are pointed since the only nonnegative function f where $-f$ is nonnegative as well is the zero function. The cones $P^{\text{odd}}_{\mathcal{A},\beta}$ and $P^{\text{even}}_{\mathcal{A},\beta}$ are defined as (infinite) intersections of closed halfspaces, and thus, they are closed. Finally, since finite sums of closed pointed convex cones are again closed, the cone $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ is closed as well. □

5.2 Circuits and the Dual of the $\mathcal{S}$-Cone

In this section, we introduce *circuit functions* as an analogon to circuit exponentials and provide several characterizations of the dual $\mathcal{S}$-cone (see Theorem 5.2.5). Notably, several important results from Chapter 3 also hold for the $\mathcal{S}$-cone.

We identify the dual space of $\mathbb{R}[\mathcal{A}, \mathcal{B}]$ with $\mathbb{R}^{(\mathcal{A},\mathcal{B})} := \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{\mathcal{B}}$. For $f \in \mathbb{R}[\mathcal{A}, \mathcal{B}]$ with coefficients $(c_\alpha)_{\alpha \in \mathcal{A}}, (d_\beta)_{\beta \in \mathcal{B}}$ and an element $(v, w) \in \mathbb{R}^{(\mathcal{A},\mathcal{B})}$, we consider the natural duality pairing

$$(v, w)(f) \ = \ \sum_{\alpha \in \mathcal{A}} v_\alpha c_\alpha + \sum_{\beta \in \mathcal{B}} w_\beta d_\beta \, . \tag{5.5}$$

This is similar to the duality theory for polynomials and exponential sums established in Chapter 2.

Using this notation, the dual cone $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})^*$ is defined as

$$C_{\mathcal{S}}(\mathcal{A},\mathcal{B})^* \;=\; \{(v,w) \in \mathbb{R}^{(\mathcal{A},\mathcal{B})} \mid (v,w)(f) \geq 0 \text{ for all } f \in C_{\mathcal{S}}(\mathcal{A},\mathcal{B})\}.$$

Now, we consider the representation of AG functions in terms of circuit functions. Recall that “relint” and “conv” denote the relative interior and the convex hull of a set and that *circuits* are defined as follows.

Definition 5.2.1. A *circuit* is a pair (A,β) where $A \subseteq \mathbb{R}^n$ is affinely independent and $\beta \in \operatorname{relint}\operatorname{conv}(A)$. For finite sets $\mathcal{A},\mathcal{B} \subseteq \mathbb{R}^n$, let

$$I(\mathcal{A},\mathcal{B}) := \{(A,\beta) \text{ circuit} \mid A \subseteq \mathcal{A}, \beta \in \mathcal{B}\}$$

denote the set of all circuits on $\mathcal{A},\mathcal{B}$. In particular, for $\mathcal{A} \subseteq \mathbb{R}^n, \mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ we call $I(\mathcal{A},\mathcal{A})$ the set of all *even* circuits and $I(\mathcal{A},\mathcal{B})$ the set of all *odd* circuits.

Note that the difference to the definition of $I(\mathcal{A})$ from Chapter 2 is the fact that we specify the set that contains the negative element — namely, it is always the second set in $I(\mathcal{A},\mathcal{B})$.

Definition 5.2.2. Let (A,β) be a circuit.

1. An *even circuit function* supported on (A,β) is an AG function of the form

$$f = \sum_{\alpha\in A} c_\alpha |x|^\alpha + d|x|^\beta.$$

2. For $\beta \in \mathbb{N}^n \setminus (2\mathbb{N})^n$, an *odd circuit function* supported on (A,β) is an AG function of the form

$$f = \sum_{\alpha\in A} c_\alpha |x|^\alpha + dx^\beta.$$

We call β the *inner* exponent of f and the other exponents are the *outer* exponents.

Remark 5.2.3. (1) As for exponential sums, the vector $\lambda \in \Lambda(\mathcal{A},\beta)$ in Theorem 5.1.8 is unique in case of a circuit, and thus, the nonnegativity of f can be expressed in terms of the *circuit number*

$$\Theta_f = \prod_{\lambda_\alpha \neq 0} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha}, \tag{5.6}$$

which was introduced in [IW16a]. For the SAGE case, compare Theorem 2.4.5.

Next, we introduce the analogous definition of *reduced* circuits — established in Chapter 3 –, which is used again in the following to determine the extreme rays of the $\mathcal{S}$-cone.

Definition 5.2.4. For a circuit (A,β) let

$$\begin{aligned} r_{\mathrm{e}}(A,\beta) &:= |\,(\operatorname{conv}(A) \setminus (A \cup \{\beta\})) \cap \mathcal{A}| \quad \text{and} \\ r_{\mathrm{o}}(A,\beta) &:= |\,(\operatorname{conv}(A) \setminus A) \cap \mathcal{A}|. \end{aligned}$$

An even circuit (A,β) is called *reduced* if $r_{\mathrm{e}}(A,\beta) = 0$ and an odd circuit (A,β) is called *reduced* if $r_{\mathrm{o}}(A,\beta) = 0$.

In contrast to the case of exponential sums, we need to distinguish between even and odd circuits: For $\beta \in \mathcal{A} \cap \mathcal{B}$, it is possible that a circuit is reduced as an even circuit, but not reduced as an odd circuit. See Example 5.2.8 below.

Analogously to the SAGE-cone, we can provide the following characterization of the dual $\mathcal{S}$-cone $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})^*$; for a proof in this setting see [KNT21]. Again, we use the convention that $0\ln(0) = 0$ and $\ln(0) = -\infty$.

Theorem 5.2.5. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets and fix an element $(v,w) \in \mathbb{R}^{(\mathcal{A},\mathcal{B})}$.*

(1) If $(v,w) \in C_{\mathcal{S}}(\mathcal{A},\mathcal{B})^*$, then $v_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$.

(2) If the condition of part (1) is satisfied, then the following are equivalent:

(a) (v,w) lies in the dual cone $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})^*$.

(b) For all $\beta \in \mathcal{A}$ (respectively $\beta \in \mathcal{B}$) and all $\lambda \in \Lambda(\mathcal{A},\beta)$, it holds that

$$\ln|v_\beta| \leq \sum_{\alpha\in\mathcal{A}} \lambda_\alpha \ln(v_\alpha) \quad \left(\textit{respectively } \ln|w_\beta| \leq \sum_{\alpha\in\mathcal{A}} \lambda_\alpha \ln(v_\alpha)\right).$$

(c) For every even circuit $(A,\beta) \in I(\mathcal{A},\mathcal{A})$ (respectively for every odd circuit $(A,\beta) \in I(\mathcal{A},\mathcal{B})$) and $\lambda = \lambda(A,\beta)$, it holds that

$$\ln|v_\beta| \leq \sum_{\alpha\in A} \lambda_\alpha \ln(v_\alpha) \quad \left(\textit{respectively } \ln|w_\beta| \leq \sum_{\alpha\in A} \lambda_\alpha \ln(v_\alpha)\right).$$

(d) For every reduced even circuit $(A,\beta) \in I(\mathcal{A},\mathcal{A})$ (respectively reduced odd circuit $(A,\beta) \in I(\mathcal{A},\mathcal{B})$) and $\lambda = \lambda(A,\beta)$, it holds that

$$\ln|v_\beta| \leq \sum_{\alpha\in A} \lambda_\alpha \ln(v_\alpha) \quad \left(\textit{respectively } \ln|w_\beta| \leq \sum_{\alpha\in A} \lambda_\alpha \ln(v_\alpha)\right).$$

The decomposition results of the SAGE-cone also hold for this specific setting; the proof in this setting can again be found in [KNT21].

Proposition 5.2.6. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets. For every $f \in C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$, the following statements hold.*

1. *f can be written as a sum of nonnegative circuit functions whose supports are contained in* $\operatorname{supp} f$.

2. *f can be written as a sum of nonnegative circuit functions supported on reduced circuits in $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$.*

Again using reduced circuits, we derive an analogous description of the extreme rays of $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$.

Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite sets and let $(A,\beta) \in I(\mathcal{A},\mathcal{A})$. Write shortly $\lambda = \lambda(A,\beta)$. Then, let

$$E_{\mathrm{e}}(A,\beta) := \left\{ \sum_{\alpha\in A} c_\alpha |x|^\alpha - \prod_{\alpha\in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} |x|^\beta \;\middle|\; c \in \mathbb{R}^A_{>0} \right\}.$$

For $(A,\beta) \in I(\mathcal{A},\mathcal{B})$ let

$$E_{\mathrm{o}}(A,\beta) := \left\{ \sum_{\alpha\in A} c_\alpha |x|^\alpha \pm \prod_{\alpha\in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} x^\beta \;\middle|\; c \in \mathbb{R}^A_{>0} \right\},$$

and for $\beta \in \mathcal{A}$ let

$$E_1(\beta) := \begin{cases} \mathbb{R}_+ \cdot |x|^\beta & \text{if } \beta \in \mathcal{A} \setminus \mathcal{B}, \\ \mathbb{R}_+ \cdot (|x|^\beta \pm x^\beta) & \text{if } \beta \in \mathcal{A} \cap \mathcal{B}. \end{cases}$$

$E_{\mathrm{e}}(A,\beta)$ and $E_{\mathrm{o}}(A,\beta)$ are the (even and odd) nonnegative circuit functions for which the inequality (5.6) on the circuit number holds with equality. $E_1(\beta)$ again provides the special case for circuits supported on a single element.

Proposition 5.2.7. *For finite sets $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$, the set $\mathcal{E}(\mathcal{A},\mathcal{B})$ of extreme rays of $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$ is*

$$\mathcal{E}(\mathcal{A},\mathcal{B}) = \left(\bigcup_{\substack{(A,\beta)\in I(\mathcal{A},\mathcal{A}),\\ r_{\mathrm{e}}(A,\beta)=0, |A|>1}} E_{\mathrm{e}}(A,\beta) \right) \cup \left(\bigcup_{\substack{(A,\beta)\in I(\mathcal{A},\mathcal{B}),\\ r_{\mathrm{o}}(A,\beta)=0, |A|>1}} E_{\mathrm{o}}(A,\beta) \right) \cup \left(\bigcup_{\beta \in \mathcal{A}} E_1(\beta) \right).$$

For a proof in this setting see again [KNT21].

The following example shows that the case distinctions in the definition of reducedness are indeed necessary.

Example 5.2.8. For $\mathcal{A} := \{0,1,2\}$ and $\mathcal{B} := \{1\}$, the sets of (even resp. odd) circuits are

$$\begin{aligned} I(\mathcal{A},\mathcal{A}) &= \{(\{0,2\},1),(\{0\},0),(\{1\},1),(\{2\},2)\} \text{ and} \\ I(\mathcal{A},\mathcal{B}) &= \{(\{1\},1),(\{0,2\},1)\}. \end{aligned}$$

We have a closer look at those elements which are both even and odd circuits.

(1) The circuit $(\{0,2\},1)$ is reduced as an even circuit and non-reduced as an odd circuit. In the context of extreme rays this is necessary. The even circuit function $a^2 - 2ab|x| + b^2x^2$ is an element of an extreme ray, but for the odd circuit function $a^2 \pm 2abx + b^2x^2$, we have

$$a^2 \pm 2abx + b^2x^2 = (a^2 - 2ab|x| + b^2x^2) + 2ab(|x| \pm x)$$

and hence, this is not an extreme ray of $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$.

(2) Further, it holds that

$$|x| = \frac{1}{2}(|x|+x) + \frac{1}{2}(|x|-x),$$

so $(\{1\},1)$ does not support an even extreme ray but in fact it does support an odd extreme ray.

5.3 A Primal-Dual View on Second-Order Representability

In this section, we examine the $\mathcal{S}$-cone and its dual for the case of rational support sets from the point of second-order representability. The results are motivated by a work of Averkov [Ave19], who has shown that the cone of SONC polynomials can be represented as a projection of a spectrahedron, using techniques from [BN01]. These techniques can be used to prove an even stronger statement, namely, second-order representability of the SONC-cone. This was already shown by Wang and Magron with an alternative proof based on binomial squares and $\mathcal{A}$-mediated sets [WM20b].

In this section, we examine the problem in terms of the rational $\mathcal{S}$-cone as well as from a primal-dual viewpoint and hence, generalize the results of Averkov and of Wang and Magron. Besides applying the techniques from [BN01], we use the extremality and duality theory established in Chapter 3 — and its generalized versions from Section 5.2 — to prevent the consideration of redundant circuits.

Beyond the specific representability result, the goal of this section is to offer further insights into the use of the framework of the $\mathcal{S}$-cone as a generalization of SONC polynomials and SAGE exponentials.

5.3.1 A Second-Order Representation for the Cone of Nonnegative AG Functions and its Dual

Here, we assume that the ground sets $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ are disjoint, finite sets. This is indeed a restriction, but for the interesting special cases of SONC polynomials and SAGE exponentials the $\mathcal{S}$-cone is built on this restriction holds anyway. Moreover, the $\mathcal{S}$-cone $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ is called *rational* if $\mathcal{A} \subseteq \mathbb{Q}^n$. Hence, we consider functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ of the form

$$f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} c_\beta x^\beta \in \mathbb{R}[\mathcal{A}, \mathcal{B}] \tag{5.7}$$

with real coefficients c_α, $\alpha \in \mathcal{A} \cup \mathcal{B}$ (i.e., we call all coefficients c_α, which can be done because the sets of exponent vectors are disjoint).

As we did in previous chapters, we can identify the dual space of $\mathbb{R}[\mathcal{A}]$ with $\mathbb{R}^{\mathcal{A}}$, and for $f \in \mathbb{R}[\mathcal{A}]$ with coefficients $c \in \mathbb{R}^{\mathcal{A}}$ and an element $v \in \mathbb{R}^{\mathcal{A}}$, we consider the natural duality pairing

$$v(f) \ = \ \sum_{\alpha \in \mathcal{A}} v_\alpha c_\alpha \, . \tag{5.8}$$

Using this notation, the dual cone $(C_{\mathcal{S}}(\mathcal{A}))^*$ is defined as

$$(C_{\mathcal{S}}(\mathcal{A}))^* \ = \ \left\{ v \in \mathbb{R}^{\mathcal{A}} \, : \, v(f) \geq 0 \text{ for all } f \in C_{\mathcal{S}}(\mathcal{A}) \right\}.$$

In contrast to the introduction of the $\mathcal{S}$-cone, we have the dual vector v instead of (v, w) here.

The following statement expresses the dual $\mathcal{S}$-cone in terms of the dual cone of SAGE exponentials.

Proposition 5.3.1. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be disjoint and finite. The dual cone of the $\mathcal{S}$-cone $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ is*

$$\begin{aligned} C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})^* &= \left\{ (v, w) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{\mathcal{B}} : (v, |w|) \in C_{SAGE}(\mathcal{A} \cup \mathcal{B})^* \right\} && (5.9) \\ &= \left\{ (v, w) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{\mathcal{B}} : \exists u \in \mathbb{R}^{\mathcal{B}} \ (v, u) \in C_{SAGE}(\mathcal{A} \cup \mathcal{B})^*, \, u \geq |w| \right\}. && (5.10) \end{aligned}$$

FIGURE 5.1: Circuit (A, β)

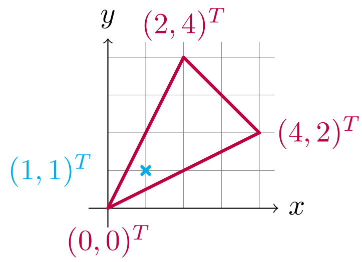

FIGURE 5.2: Circuit (A', β')

Proof. We use Proposition 5.1.5, which provides a characterization for the primal cone $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ in terms of an existential quantification. Consider its lifted cone

$$\begin{aligned}\widehat{C_{\mathcal{S}}}(\mathcal{A}, \mathcal{B}) &:= C_{\mathrm{SAGE}}(\mathcal{A} \cup \mathcal{B}) \times \mathbb{R}^{\mathcal{B}} \cap \{(c,t,d) : t_\beta \leq -|d_\beta| \text{ for all } \beta \in \mathcal{B}\} \\ &= C_{\mathrm{SAGE}}(\mathcal{A} \cup \mathcal{B}) \times \mathbb{R}^{\mathcal{B}} \cap \{(c,t,d) : t_\beta \leq d_\beta, t_\beta \leq -d_\beta \text{ for all } \beta \in \mathcal{B}\}\end{aligned} \tag{5.11}$$

in the space $\mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{\mathcal{B}} \times \mathbb{R}^{\mathcal{B}}$. The dual cone of the right-hand cone in (5.11) is the set

$$\operatorname{pos}\left\{(0, \ldots, 0, -\delta^{(\beta)}, \pm\delta^{(\beta)}) \,:\, \beta \in \mathcal{B}\right\},$$

where $\delta^{(\beta)}$ denotes the unit vector with respect to $\beta \in \mathcal{B}$. As intersection and Minkowski sum are dual operations, we obtain

$$\widehat{C_{\mathcal{S}}}(\mathcal{A}, \mathcal{B})^* = C_{\mathrm{SAGE}}(\mathcal{A} \cup \mathcal{B})^* \times \{0\} + \operatorname{pos}\left\{(0, \ldots, 0, -\delta^{(\beta)}, \pm\delta^{(\beta)}) \,:\, \beta \in \mathcal{B}\right\}.$$

Identifying the $\mathcal{S}$-cone with its coefficients, we can express $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})^*$ in terms of the lifted cone $\widehat{C_{\mathcal{S}}}(\mathcal{A}, \mathcal{B})$ by

$$C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})^* = \widehat{C_{\mathcal{S}}}(\mathcal{A}, \mathcal{B}) \cap \left\{(v,s,w) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{\mathcal{B}} \times \mathbb{R}^{\mathcal{B}} \,:\, s = 0\right\}.$$

Thus, $(v, w) \in C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ whenever $(v, |w|) \in C_{\mathrm{SAGE}}(\mathcal{A} \cup \mathcal{B})^*$. Convexity then implies the second characterization (5.10). ☐

Hence, as in the primal case, it suffices to study even AG functions in the dual situation. We will make use of a representation of the dual of the $\mathcal{S}$-cone from Theorem 5.2.5. As $\mathcal{A} \cap \mathcal{B} = \emptyset$, for a finite set $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$, the set of *circuits* supported on $\mathcal{A}$ reduces to the set

$$I(\mathcal{A}) \quad = \quad \left\{(A, \beta) \,:\, A \subseteq \mathcal{A} \text{ affinely independent}, \ \beta \in \operatorname{relint}(\operatorname{conv} A) \cap (\mathcal{A} \setminus A)\right\}.$$

Two examples of circuits are the pairs (A, β) with $A = \{0, 6\}$ and $\beta = \{2\}$ (see Figure 5.1) and (A', β') with $A' = \{(0,0)^T, (4,2)^T, (2,4)^T\}$ and $\beta' = (1,1)^T$ (see Figure 5.2).

Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite. In the previous section, we proved that a point $v \in \mathbb{R}^{\mathcal{A}}$ is contained in the dual $\mathcal{S}$-cone $C_{\mathcal{S}}(\mathcal{A})^*$ if and only if $v \geq 0$ and

$$\ln(v_\beta) \leq \sum_{\alpha \in A} \lambda_\alpha \ln(v_\alpha) \text{ for every circuit } (A, \beta) \text{ in } I(\mathcal{A}) \text{ and } \lambda = \lambda(A, \beta). \tag{5.12}$$

In order to provide a second-order representation for the $\mathcal{S}$-cone and its dual, the main task is to capture the cone of nonnegative AG functions and its dual. For a comprehensive collection of techniques for handling second-order-cones, we refer to [BN01].

Throughout the section, let (A, β) be a fixed circuit with rational barycentric coordinates $\lambda \in \mathbb{Q}_+^A$, which represent β as a convex combination of elements in A, i.e., $\beta = \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \alpha$ and $\sum_{\alpha \in \mathcal{A}} \lambda_\alpha = 1$. Let $p \in \mathbb{N}$ denote the smallest common denominator of the fractions λ_α for $\alpha \in A$, i.e., $\lambda_\alpha = \frac{p_\alpha}{p}$ with $p_\alpha \in \mathbb{N}$ for all $\alpha \in A$ and p is minimal. Recall that whenever $|A| = 1$, then $A = \{\beta\}$ by definition.

With the given circuit $(A, \beta) \in I(\mathcal{A})$, we associate a set of *dual circuit variables*

$$(y_{k,i})_{k,i}, \tag{5.13}$$

where $k \in [\lceil \log_2(p) \rceil]$ and $i \in [2^{\lceil \log_2(p) \rceil - k}]$. We denote the collection of these $\sum_{k=1}^{\lceil \log_2(p) \rceil} 2^{\lceil \log_2(p) \rceil - k} = 2^{\lceil \log_2(p) \rceil} - 1$ variables as $y^{A,\beta}$ or shortly as y. Further, denote the restriction of a vector $v \in \mathbb{R}^{\mathcal{A}}$ to the components of $A \subseteq \mathcal{A}$ by $v_{|A}$.

Definition 5.3.2. A *dual circuit matrix* $C^*_{A,\beta}(v_{|A}, v_\beta, y)$ is a block diagonal matrix consisting of the block $v_\beta \geq 0$ if $|A| = 1$ and, if $|A| > 1$, consisting of the blocks

$$\begin{pmatrix} y_{k-1,2i-1} & y_{k,i} \\ y_{k,i} & y_{k-1,2i} \end{pmatrix} \quad \text{for } k \in \{2, \ldots, \lceil \log_2(p) \rceil\} \text{ and } i \in [2^{\lceil \log_2(p) \rceil - k}], \tag{5.14}$$

the singleton blocks (v_β), and $(y_{\lceil \log_2(p) \rceil, 1} - v_\beta)$ as well as $2^{\lceil \log_2(p) \rceil - 1}$ blocks of the form

$$\begin{pmatrix} u & y_{1,l} \\ y_{1,l} & w \end{pmatrix} \quad \text{for } l \in [2^{\lceil \log_2(p) \rceil - 1}] \tag{5.15}$$

where, in each of these blocks, u and w represent a variable of $\{v_\alpha : \alpha \in A\} \cup \{v_\beta\}$ such that altogether each v_α appears p_α times and v_β appears $2^{\lceil \log_2(p) \rceil} - p$ times.

Note that in this definition, the exact order of appearances of the variables in the set $\{v_\alpha : \alpha \in A\} \cup \{v_\beta\}$ is not uniquely determined. However, since this order of appearances will not matter, we speak of *the* dual circuit matrix. The case distinction depending on the size of A ensures that also circuits supporting an atomic extremal ray of the $\mathcal{S}$-cone are captured.

Remark 5.3.3. Note that this corrects [NT21b], where the dual circuit variables as well as the dual circuit matrix were defined slightly different, causing a problem for $p = 2$.

Remark 5.3.4. Each block of the type (5.15) contains two (not necessarily identical) variables from $\{v_\alpha : \alpha \in A\} \cup \{v_\beta\}$. Since $\sum_{\alpha \in A} \lambda_\alpha = 1$, we have $\sum_{\alpha \in A} p_\alpha = p$ and, hence, the total number of occurrences of variables from the set $\{v_\alpha : \alpha \in A\} \cup \{v_\beta\}$ in the blocks of type (5.15) is

$$\sum_{\alpha \in A} p_\alpha + (2^{\lceil \log_2(p) \rceil} - p) = 2^{\lceil \log_2(p) \rceil},$$

which is twice the number of blocks of type (5.15).

Note that every $y_{k,i}$ only serves as an auxiliary variable to make the non-linear constraints $\ln(v_\beta) \leq \sum_{\alpha \in A} \lambda_\alpha \ln(v_\alpha)$ of the dual $\mathcal{S}$-cone description from (5.12) linear.

In the end, we will only multiply those constraints to obtain the original ones. In particular, factors v_β serve to cover cases where p is not a power of 2. For the purpose of the second-order descriptions, it does not matter in which order the variables appear in the blocks (5.15), because only the product of these blocks will be considered.

The goal of this subsection is to show the following characterization of the dual cone of nonnegative even AG functions $P_{A,\beta}^{\text{even}}$ supported on the circuit (A, β). Recall that positive semidefiniteness of a symmetric matrix is denoted by "$\succeq 0$".

Theorem 5.3.5. *The dual cone $(P_{A,\beta}^{\text{even}})^*$ of the cone of nonnegative even AG functions $P_{A,\beta}^{\text{even}}$ supported on the circuit $(A, \beta) \in I(\mathcal{A})$ is the projection of the spectrahedron*

$$\left\{(v, y) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{2^{\lceil \log_2(p) \rceil} - 1} \;:\; C_{A,\beta}^*(v_{|A}, v_\beta, y) \succcurlyeq 0\right\} \tag{5.16}$$

on $(v_{|A}, v_\beta)$. $(P_{A,\beta}^{\text{even}})^*$ is second-order representable.

Here, the second-order representability follows immediately from the representation (5.16) in connection with Lemma 2.1.6. Let us consider an example for the theorem.

Example 5.3.6. Let $\mathcal{A} = \{0, 6\}, \mathcal{B} = \{2\}$ and consider the circuit (A, β) with $A = \mathcal{A}$ and $\beta = 2$ (compare Figure 5.1). We have $p = 3, p_0 = 2, p_6 = 1$ and y consists of the components

$$y_{1,1},\ y_{1,2},\ y_{2,1}.$$

A vector (v_0, v_2, v_6) is contained in $(P_{A,\beta}^{\text{even}})^*$ if and only if $0 \leq v_2 \leq y_{2,1}$ and the three 2×2-matrices

$$\begin{pmatrix} v_6 & y_{1,1} \\ y_{1,1} & v_2 \end{pmatrix}, \begin{pmatrix} v_0 & y_{1,2} \\ y_{1,2} & v_0 \end{pmatrix}, \begin{pmatrix} y_{1,1} & y_{2,1} \\ y_{2,1} & y_{1,2} \end{pmatrix}$$

are positive semidefinite.

In [Ave19], Averkov considers the size of the blocks in the SDP-representation of SONC-polynomials but does not give a number or bound on the number of blocks. Here, for the $\mathcal{S}$-cone, we provide a bound on the number of inequalities of a second-order representation, which also gives a bound on the number of 2×2-blocks in a semidefinite representation. The bound depends on the smallest common denominator of the barycentric coordinates representing the inner exponent of a circuit as a convex combination of the outer ones.

Corollary 5.3.7. *The matrix $C_{A,\beta}^*(v_{|A}, v_\beta, y)$ consists of $2^{\lceil \log_2(p) \rceil} - 1$ blocks of size 2×2 and two block of size 1×1 if $|A| > 1$ and one 1×1-block if $|A| = 1$.*

Proof. Counting the 2×2-blocks, there are $\sum_{k=2}^{\lceil \log_2(p) \rceil} \left(2^{\lceil \log_2(p) \rceil - k}\right) = 2^{\lceil \log_2(p) \rceil - 1} - 1$ blocks of type (5.14) and $2^{\lceil \log_2(p) \rceil - 1}$ blocks of type (5.15). □

Remark 5.3.8. It is useful to record the set inequalities characterizing the positive semidefiniteness of the matrix $C_{A,\beta}^*(v_{|A}, v_\beta, y)$ in the case that $|A| > 1$. Besides the nonnegativity conditions for the variables,

$$v_{|A} \geq 0, \quad v_\beta \geq 0, \tag{5.17}$$

$$\text{and } y_{k,i} \geq 0 \text{ for all } k \in \{2, \dots, \lceil \log_2(p) \rceil - 1\}, i \in [2^{\lceil \log_2(p) \rceil} - k], \tag{5.18}$$

these are the determinantal conditions arising from the positive semidefiniteness of the second singleton block and the matrices in (5.14) and (5.15):

$$(0 \leq)\ v_\beta \leq y_{\lceil \log_2(p) \rceil,1}, \tag{5.19}$$
$$y_{k,i}^2 \leq y_{k-1,2i-1} y_{k-1,2i} \text{ for all } k \in \{2, \ldots, \lceil \log_2(p) \rceil\}, i \in [2^{\lceil \log_2(p) \rceil - k}] \tag{5.20}$$
$$\text{and } uw \geq (y_{1,l})^2 \text{ for } l \in [2^{\lceil \log_2(p) \rceil - 1}] \tag{5.21}$$

for $u, w \in \{v_\alpha : \alpha \in A\} \cup \{v_\beta\}$ such that v_α appears p_α times for every $\alpha \in A$ and v_β appears $2^{\lceil \log_2(p) \rceil} - p$ times.

The next lemma prepares one inclusion of Theorem 5.3.5.

Lemma 5.3.9. *Let $v \in \mathbb{R}^{A,\beta}$ such that there exists some vector $y \in \mathbb{R}^{2^{\lceil \log_2(p) \rceil} - 1}$ with $C^*_{A,\beta}(v_{|A}, v_\beta, y) \succcurlyeq 0$. Then $v_{|A}$ is nonnegative and satisfies*

$$v_\beta^p \leq \prod_{\alpha \in A} v_\alpha^{p_\alpha}.$$

Proof. If $|A| = 1$, the claim is obvious, so assume $|A| > 1$. By (5.17), we have $v_{|A} \geq 0$ and $v_\beta \geq 0$. Moreover, (5.19) and successively applying (5.20) gives

$$\begin{aligned}
v_\beta &\leq y_{\lceil \log_2(p) \rceil,1} \leq \left(y_{\lceil \log_2(p) \rceil - 1,1}\, y_{\lceil \log_2(p) \rceil - 1,2} \right)^{1/2} \\
&\leq \left(y_{\lceil \log_2(p) \rceil - 2,1}\, y_{\lceil \log_2(p) \rceil - 2,2} \right)^{1/4} \left(y_{\lceil \log_2(p) \rceil - 2,3}\, y_{\lceil \log_2(p) \rceil - 2,4} \right)^{1/4} \\
&= \left(y_{\lceil \log_2(p) \rceil - 2,1}\, y_{\lceil \log_2(p) \rceil - 2,2}\, y_{\lceil \log_2(p) \rceil - 2,3}\, y_{\lceil \log_2(p) \rceil - 2,4} \right)^{\frac{1}{2^{\lceil \log_2(p) \rceil - (\lceil \log_2(p) \rceil - 2)}}} \\
&\leq \cdots \leq \left(\left(\prod\nolimits_{\alpha \in A} v_\alpha^{p_\alpha} \right) \cdot (v_\beta)^{2^{\lceil \log_2(p) \rceil} - p} \right)^{\frac{1}{2^{\lceil \log_2(p) \rceil}}}.
\end{aligned}$$

This is equivalent to

$$(v_\beta)^{2^{\lceil \log_2(p) \rceil}} \cdot (v_\beta)^{p - 2^{\lceil \log_2(p) \rceil}} \leq \prod\nolimits_{\alpha \in A} v_\alpha^{p_\alpha},$$

which implies $v_\beta^p \leq \prod_{\alpha \in A} v_\alpha^{p_\alpha}$. □

Now we prepare the converse inclusion of Theorem 5.3.5.

Lemma 5.3.10. *For every $v \in \mathbb{R}^{A,\beta}$ with $v_{|A \cup \{\beta\}} \geq 0$ and $v_\beta^p \leq \prod_{\alpha \in A} v_\alpha^{p_\alpha}$, there exists $y \in \mathbb{R}^{2^{\lceil \log_2(p) \rceil} - 1}$ such that $C^*_{A,\beta}(v_{|A}, v_\beta, y) \succcurlyeq 0$.*

Proof. Again, if $|A| = 1$, the claim is obvious, so assume $|A| > 1$. Define y inductively by

$$y_{1,l} = \sqrt{uw} \text{ for those } u, w \text{ which occur in the block with } y_{1,l},$$
$$y_{k,i} = \sqrt{y_{k-1,2i-1} y_{k-1,2i}} \text{ for all } k \in \{2, \ldots, \lceil \log_2(p) \rceil\}, i \in [2^{\lceil \log_2(p) \rceil - k}].$$

It suffices to show that the inequalities (5.17)-(5.21) in Remark 5.3.8 are satisfied. The nonnegativity conditions (5.17) and (5.18) hold by assumption and by definition of y. The construction of y also implies that a subchain of the chain of inequalities considered in the previous proof even holds with equality, namely,

$$\begin{aligned}
y_{\lceil \log_2(p)\rceil,1} &= \left(y_{\lceil \log_2(p)\rceil-1,1}\, y_{\lceil \log_2(p)\rceil-1,2}\right)^{1/2} \\
&= \left(y_{\lceil \log_2(p)\rceil-2,1}\, y_{\lceil \log_2(p)\rceil-2,2}\right)^{1/4} \left(y_{\lceil \log_2(p)\rceil-2,3}\, y_{\lceil \log_2(p)\rceil-2,4}\right)^{1/4} \\
&= \left(y_{\lceil \log_2(p)\rceil-2,1}\, y_{\lceil \log_2(p)\rceil-2,2}\, y_{\lceil \log_2(p)\rceil-2,3}\, y_{\lceil \log_2(p)\rceil-2,4}\right)^{\frac{1}{2^{\lceil \log_2(p)\rceil-(\lceil \log_2(p)\rceil-2)}}} \\
&= \cdots = \left(\left(\prod\nolimits_{\alpha\in A} v_\alpha^{p_\alpha}\right)\cdot (v_\beta)^{2^{\lceil \log_2(p)\rceil}-p}\right)^{\frac{1}{2^{\lceil \log_2(p)\rceil}}} .
\end{aligned}$$

By the assumption $v_\beta^p \le \prod_{\alpha\in A} v_\alpha^{p_\alpha}$, we obtain $v_\beta \le y_{\lceil \log_2(p)\rceil,1}$, which shows inequality (5.19). The remaining inequalities (5.20) and (5.21) are satisfied with equality by construction. $\square$

Finally, we can conclude the proof of Theorem 5.3.5.

Proof of Theorem 5.3.5. As the claim is clear for $|A| = 1$, assume $|A| > 1$. Let p be defined as in Definition 5.3.2 and let $\lambda \in \mathbb{Q}_+^A$ with $\lambda_\alpha = \frac{p_\alpha}{p}$ and $p_\alpha \in \mathbb{N}$ for all $\alpha \in A$ denote the barycentric coordinates representing β as a convex combination of the elements in A, i.e., $\beta = \sum_{\alpha\in A} \lambda_\alpha \alpha$ and $\sum_{\alpha\in A} \lambda_\alpha = 1$. By (5.12), we have

$$\begin{aligned}
(P_{A,\beta}^{\text{even}})^* &= \left\{v \in \mathbb{R}^{A,\beta} \,:\, v_{|A\cup\{\beta\}} \ge 0,\ \ln(v_\beta) \le \sum\nolimits_{\alpha\in A} \lambda_\alpha \ln(v_\alpha)\right\} \\
&= \left\{v \in \mathbb{R}^{A,\beta} \,:\, v_{|A\cup\{\beta\}} \ge 0,\ v_\beta^p \le \prod\nolimits_{\alpha\in A} v_\alpha^{p_\alpha}\right\}.
\end{aligned}$$

Applying Lemmas 5.3.9 and 5.3.10, $C_{A,\beta}^*(x, v_\beta) \succcurlyeq 0$ if and only if $v \in P_{A,\beta}^*$. $\square$

Our derivation of the second-order representation of the dual cone $(P_{A,\beta}^{\text{even}})^*$ also suggests a simple way to derive a second-order representation of the primal cone $P_{A,\beta}^{\text{even}}$. For the dual cone, (5.12) gives — besides nonnegativity-constraints on v_α for $\alpha \in \mathcal{A}$ and on v_β — the condition $\ln(v_\beta) \le \sum_{\alpha\in A} \lambda_\alpha \ln(v_\alpha)$ for every $(A, \beta) \in I(\mathcal{A})$. These conditions can — as done in the previous proof — be stated as

$$v_\beta^p \le \prod_{\alpha\in A} v_\alpha^{p_\alpha} \text{ where } p = \frac{p_\alpha}{\lambda_\alpha} .$$

The conditions for the primal cone can be reformulated similarly. Namely, an even circuit function f with a vector of coefficients c is nonnegative if and only if $-c_\beta \le \prod_{\alpha\in A} (c_\alpha/\lambda_\alpha)^{\lambda_\alpha}$ by (5.6), which we write as

$$(-c_\beta)^p \le \prod_{\alpha\in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{p_\alpha} .$$

This motivates to carrying over the definition of the dual circuit matrix to the primal case as follows. Since c_β may be negative (in contrast to the dual case), we introduce the *primal circuit variables*, or simply *circuit variables*,

$$(x_\beta, (x_{k,i})_{k,i})$$

where $k \in [\lceil \log_2(p)\rceil]$ and $i \in [2^{\lceil \log_2(p)\rceil-k}]$. As in the dual case, we refer to these $1 + \sum_{k=1}^{\lceil \log_2(p)\rceil} 2^{\lceil \log_2(p)\rceil-k} = 2^{\lceil \log_2(p)\rceil}$ variables as $x^{A,\beta}$ or shortly as x.

Definition 5.3.11 (Circuit Matrix)**.** The *circuit matrix* $C_{A,\beta}(c_{|A\cup\{\beta\}}, x_\beta, x)$ is the block diagonal matrix consisting of the block $c_\beta \geq 0$ if $|A| = 1$ and, if $|A| > 1$, consisting of the blocks

$$\begin{pmatrix} x_{k-1,2i-1} & x_{k,i} \\ x_{k,i} & x_{k-1,2i} \end{pmatrix} \quad \text{for } k \in \{2, \ldots, \lceil \log_2(p) \rceil\},\ i \in [2^{\lceil \log_2(p)\rceil - k}],$$

the two singleton blocks

$$\left(x_{\lceil \log_2(p)\rceil,1} - \left(\prod_{\alpha\in A} (\lambda_\alpha)^{\lambda_\alpha}\right) x_\beta \right) \text{ and } \left(x_\beta + c_\beta \right), \tag{5.22}$$

as well as $2^{\lceil \log_2(p)\rceil - 1}$ blocks of the form

$$\begin{pmatrix} u & x_{1,l} \\ x_{1,l} & w \end{pmatrix} \quad \text{for } l \in [2^{\lceil \log_2(p)\rceil - 1}] \tag{5.23}$$

where $u, w \in \{c_\alpha : \alpha \in A\} \cup \{\left(\prod_{\alpha\in A}(\lambda_\alpha)^{\lambda_\alpha}\right) x_\beta\}$ such that c_α appears p_α times for every $\alpha \in A$ and $\left(\prod_{\alpha\in A}(\lambda_\alpha)^{\lambda_\alpha}\right) x_\beta$ appears $2^{\lceil \log_2(p)\rceil} - p$ times.

Note that for a circuit (A, β), the product $\left(\prod_{\alpha\in A}(\lambda_\alpha)^{\lambda_\alpha}\right)$ is always nonzero because $\beta \in \operatorname{relint} \operatorname{conv} A$ and A consists of affinely independent vectors.

Whenever $|A| > 1$, in contrast to the dual cone, there is no sign constraint on c_β in the primal cone. If p is not a power of 2, then x_β appears on the main diagonal of (5.23). In our coupling of x_β with c_β, the constraint $x_\beta + c_\beta \geq 0$ results in $-c_\beta \leq x_\beta$ and thus, reflects these sign considerations.

Note that the primal cone consists of circuit *functions*, whereas in our definition of the dual cone, the elements are coefficient *vectors*. Therefore, the projection considered in Theorem 5.3.5 only delivers the coefficients of the circuit functions rather than the cone itself.

Theorem 5.3.12. *The set of coefficients of the cone $P_{A,\beta}^{\text{even}}$ of nonnegative even circuit polynomials supported on the circuit (A, β) coincides with the projection of the spectrahedron $\widehat{P_{A,\beta}^{\text{even}}}$ defined as*

$$\left\{(c, x) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{2^{\lceil \log_2(p)\rceil}} : C_{A,\beta}(c_{|A\cup\{\beta\}}, x_\beta, x) \succcurlyeq 0,\ c_{|\mathcal{A}\setminus(A\cup\{\beta\})} = 0\right\} \tag{5.24}$$

on $(c_{|A}, c_\beta)$. The cone $P_{A,\beta}^{\text{even}}$ is second-order representable.

The last equality constraint in (5.24) is redundant and can be omitted. We include it here because this formulation is needed in Section 5.3.2 for the description of the $\mathcal{S}$-cone supported on the full set $\mathcal{A}$.

Proof. As the claim is clear for $|A| = 1$, assume $|A| > 1$.

First, let $(c, x) \in \widehat{P_{A,\beta}^{\text{even}}}$. The positive semidefiniteness of the 2×2-blocks in $C_{A,\beta}(c_{|A\cup\{\beta\}}, x_\beta, x)$ implies the inequalities

$$c_{|A} \geq 0 \text{ and } (-x_\beta)^p \cdot \left(\prod_{\alpha\in A} \lambda_\alpha{}^{\lambda_\alpha}\right)^p \leq \prod_{\alpha\in A} c_\alpha^{p_\alpha}.$$

The two 1×1-blocks from (5.22) give the inequalities $x_{\lceil \log_2(p)\rceil,1} \geq \left(\prod_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right) x_\beta$ and $x_\beta \geq -c_\beta$. They imply $-c_\beta \left(\prod_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right) \leq x_\beta \left(\prod_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right) \leq x_{\lceil \log_2(p)\rceil,1}$. Hence, similar to Lemma 5.3.9,

$$\begin{aligned}
& x_\beta \left(\prod\nolimits_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right) \leq x_{\lceil \log_2(p)\rceil,1} \ \leq \ \left(x_{\lceil \log_2(p)\rceil-1,1}\, x_{\lceil \log_2(p)\rceil-1,2}\right)^{1/2} \\
\leq \ & \left(x_{\lceil \log_2(p)\rceil-2,1}\, x_{\lceil \log_2(p)\rceil-2,2}\right)^{1/4} \left(x_{\lceil \log_2(p)\rceil-2,3}\, x_{\lceil \log_2(p)\rceil-2,4}\right)^{1/4} \\
= \ & \left(x_{\lceil \log_2(p)\rceil-2,1}\, x_{\lceil \log_2(p)\rceil-2,2}\, x_{\lceil \log_2(p)\rceil-2,3}\, x_{\lceil \log_2(p)\rceil-2,4}\right)^{\frac{1}{2^{\lceil \log_2(p)\rceil-(\lceil \log_2(p)\rceil-2)}}} \\
\leq \ & \cdots \leq \left(\left(\prod\nolimits_{\alpha\in A} c_\alpha^{p_\alpha}\right)\cdot (x_\beta)^{2^{\lceil \log_2(p)\rceil}-p} \left(\prod\nolimits_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right)^{2^{\lceil \log_2(p)\rceil}-p}\right)^{\frac{1}{2^{\lceil \log_2(p)\rceil}}}.
\end{aligned}$$

This is equivalent to

$$x_\beta^{2^{\lceil \log_2(p)\rceil}} \cdot \left(\prod\nolimits_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right)^{2^{\lceil \log_2(p)\rceil}} \cdot x_\beta^{p-2^{\lceil \log_2(p)\rceil}} \cdot \left(\prod\nolimits_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha}\right)^{p-2^{\lceil \log_2(p)\rceil}} \leq \prod\nolimits_{\alpha\in A} c_\alpha^{p_\alpha}.$$

Altogether, we obtain $(-c_\beta)^p \leq \prod_{\alpha\in A}(c_\alpha/\lambda_\alpha)^{p_\alpha}$ and further $c_{|A\cup\{\beta\}} \in P_{A,\beta}^{\text{even}}$.

For the converse inclusion, we remind the reader that $\lambda_\alpha > 0$ for all $\alpha \in A$. We set $x_\beta := x_{\lceil \log_2(p)\rceil,1}\left(\prod_{\alpha\in A}\left(\frac{1}{\lambda_\alpha}\right)^{\lambda_\alpha}\right)$ and, similar to the proof of Lemma 5.3.10, define x inductively by

$$\begin{aligned}
& x_{1,l} = \sqrt{uw} \text{ for those } u, w \text{ which occur in the block with } x_{1,l}, \\
& x_{k,i} = \sqrt{x_{k-1,2i-1}x_{k-1,2i}} \text{ for all } k \in \{2,\ldots,\lceil \log_2(p)\rceil\}, i \in [2^{\lceil \log_2(p)\rceil-k}].
\end{aligned}$$

Analogous to that proof, the construction of x gives $C_{A,\beta}(c_{A\cup\{\beta\}}, x_\beta, x) \succeq 0$.

Second-order representability is then an immediate consequence according to Lemma 2.1.6. □

Example 5.3.13. Let $\mathcal{A} = \{0,2\}$, $\mathcal{B} = \{1\}$ and consider the circuit (A,β) with $A = \mathcal{A}$ and $\beta = 1$. Since

$$1 = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 2,$$

we have $p_1 = p_2 = 1$ and the smallest common denominator $p = 2$. Hence, $\lceil \log_2(p)\rceil = \log_2(p) = 1$ and $2^{\lceil \log_2(p)\rceil} - p = 2 - p = 0$ as well as

$$\prod_{\alpha\in A} \lambda_\alpha^{\lambda_\alpha} = \frac{1}{2} \ \text{ and } \ x = \begin{pmatrix} x_1 \\ x_{1,1} \end{pmatrix}.$$

A given vector (c_0, c_1, c_2) is contained in $P_{\mathcal{A},\beta}$ if and only if

$$x_{1,1} - \frac{1}{2}x_1 \geq 0, \ x_1 + c_1 \geq 0 \ \text{ and } \ \begin{pmatrix} c_0 & x_{1,1} \\ x_{1,1} & c_2 \end{pmatrix} \succeq 0.$$

Similar to Corollary 5.3.7, we can determine the number of blocks.

Corollary 5.3.14. *The matrix $C_{A,\beta}(c_{|A\cup\{\beta\}}, x_\beta, x)$ consists of $2^{\lceil \log_2(p)\rceil} - 1$ blocks of size 2×2 and two blocks of size 1×1 if $|A| > 1$ and of a single 1×1-bock if $|A| = 1$.*

5.3.2 A Second-Order Representation of the $\mathcal{S}$-Cone and its Dual

In the previous subsection, we obtained second-order representations of the subcones of nonnegative even circuit functions and their duals, under the condition that the

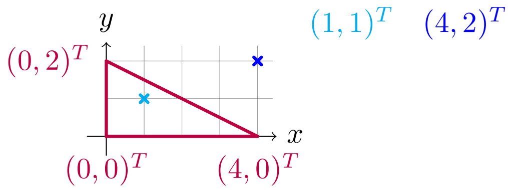

FIGURE 5.3: The circuit is reduced, as $(4,2)^T \notin \mathrm{conv}(A)$.

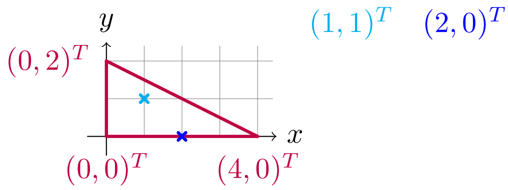

FIGURE 5.4: The circuit is not reduced, as $(2,0)^T \in \mathrm{conv}(A)$.

barycentric coordinates are rational. We now assume that $\mathcal{A}$ and $\mathcal{B}$ are rational and derive an explicit second-order representation of the rational $\mathcal{S}$-cone $C_{\mathcal{S}}(\mathcal{A},\mathcal{B})$ and its dual. In the primal case, those cones are obtained via projection and Minkowski sum, and in the dual case, they arise from projection and intersection. First we consider the lifted cones for the dual case.

Taking all circuits (A,β) into account would induce a highly redundant representation. To avoid those redundancies, we make use of the following characterization of the extreme rays of the $\mathcal{S}$-cone from the first section of this chapter, and with this, recall the following definition.

For finite and disjoint sets $\emptyset \neq \mathcal{A},\mathcal{B} \subseteq \mathbb{R}^n$, the set of *reduced* circuits contained in $\mathcal{A} \cup \mathcal{B}$ is the set

$$
\begin{aligned}
R(\mathcal{A},\mathcal{B}) \quad = \quad & \Big\{(A,\beta) \,:\, A \subseteq \mathcal{A} \text{ affinely independent}, \;\; \beta \in \mathrm{relint}(\mathrm{conv}\, A) \cap (\mathcal{B} \setminus A), \\
& \mathcal{A} \cap (\mathrm{conv}(A)) \setminus (A \cup \{\beta\}) = \emptyset\Big\}.
\end{aligned}
$$

Less formally, this is the set of all circuits (A,β) with outer exponents in $\mathcal{A}$ and inner exponents in $\mathcal{B}$ without additional support points contained in the convex hull of A. In particular, for every $\beta \in \mathcal{A}$, the circuit $(A,\beta) \in \mathcal{I}(\mathcal{A},\mathcal{A})$ with $|A| = 1$ (i.e., with $A = \{\beta\}$) is a reduced even circuit.

Note that for $\mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ disjoint and finite, the set $R(\mathcal{A},\mathcal{A})$ is exactly the set of *even reduced circuits* and the set $R(\mathcal{A},\mathcal{B})$ is exactly the set of *odd reduced circuits*. The set $R(\mathcal{A},\mathcal{A}\cup\mathcal{B})$ denotes the set of *all* reduced circuits (A,β) with $A \subseteq \mathcal{A}$ and $\beta \in \mathcal{A}\cup\mathcal{B}$. A circuit function supported on a reduced circuit in $R(\mathcal{A},\mathcal{A}\cup\mathcal{B})$ has nonnegative coefficients corresponding to exponents in $\mathcal{A}$ and a possibly negative coefficient corresponding to a single exponent in $\mathcal{A}\cup\mathcal{B}$.

Recall that the question whether a circuit is reduced or not depends on the ground set $\mathcal{A}$. For example, the circuit (A,β) with $A = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$ and $\beta = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is *reduced* for the ground set $\mathcal{A} = A \cup \{\beta\} \cup \left\{ \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right\}$ (compare Figure 5.3) but *not reduced* for $\mathcal{A} = A \cup \{\beta\} \cup \left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}$ (compare Figure 5.4).

The following proposition is a direct consequence of Theorem 5.2.5(d).

Proposition 5.3.15. *Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ and $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ be finite and disjoint sets. Then,*

$$
C_{\mathcal{S}}(\mathcal{A},\mathcal{B}) = \sum_{(A,\beta)\in R(\mathcal{A},\mathcal{A})} P^{\mathrm{even}}_{A,\beta} + \sum_{(A,\beta)\in R(\mathcal{A},\mathcal{B})} P^{\mathrm{odd}}_{A,\beta}.
$$

Using this decomposition, we can exclude many circuits from our consideration. Thus, the second-order-cone program will be much smaller than the one considering all circuits.

In the previous subsection, we only considered even circuits. To use Proposition 5.1.5 and obtain the conditions for odd circuits as well, we extend the dual circuit variables for odd circuits to

$$(y_\beta, (y_{k,i})_{k,i})$$

for $k \in [2^{\lceil \log_2(p) \rceil}]$ and $i \in [2^{\log_2(p)-k}]$. Nevertheless, we call them $y^{A,\beta}$ for a fixed circuit $(A, \beta) \in R(\mathcal{A}, \mathcal{B})$.

For the dual case, we consider the coordinates

$$y^{\mathcal{A},\mathcal{B}} = \left\{ (y^{A,\beta}) \,:\, (A, \beta) \in R(\mathcal{A}, \mathcal{A} \cup \mathcal{B}) \right\},$$

which consist of $\sum_{(A,\beta) \in R(\mathcal{A},\mathcal{A}\cup\mathcal{B})} 2^{\lceil \log_2(p_{A,\beta}) \rceil}$ components. Here, $p_{A,\beta}$ denotes the smallest common denominator of the barycentric coordinates $\lambda_{A,\beta}$ of the circuit (A, β) representing β as a convex combination of A.

For the primal case, we consider

$$x^{\mathcal{A},\mathcal{B}} = \left\{ (x^{A,\beta}) \,:\, (A, \beta) \in R(\mathcal{A}, \mathcal{A} \cup \mathcal{B}) \right\},$$

which also consist of $\sum_{(A,\beta) \in R(\mathcal{A},\mathcal{A}\cup\mathcal{B})} 2^{\lceil \log_2(p_{A,\beta}) \rceil}$ components.

Using Proposition 5.1.5, we can use our earlier characterizations of $P^{\text{even}}_{A,\beta}$ to obtain the following second-order characterization for $P^{\text{odd}}_{A,\beta}$.

Corollary 5.3.16. *Let $(A, \beta) \in R(\mathcal{A}, \mathcal{B})$ an odd reduced circuit with rational set $A \subseteq \mathcal{A} \subseteq \mathbb{Q}^n$ and rational $\beta \in \mathcal{B}$.*

(1) Let f be an odd AG function supported on (A, β) with coefficient vector c. Then, f is nonnegative if and only if there exists $x \in \mathbb{R}^{2^{\lceil \log_2(p) \rceil}}$ such that $C_{A,\beta}(c_{|A}, x_\beta, x) \succcurlyeq 0$ and

$$\begin{pmatrix} x_\beta & c_\beta \\ c_\beta & x_\beta \end{pmatrix} \succcurlyeq 0. \tag{5.25}$$

(2) A vector $v \in \mathbb{R}^{A,\beta}$ is contained in $\left(P^{\text{odd}}_{A,\beta}\right)^*$ if and only if there exist some $y \in \mathbb{R}^{2^{\lceil \log_2(p) \rceil}-1}$ and $y_\beta \in \mathbb{R}$ such that $C^*_{A,\beta}(v_{|A}, y_\beta, y) \succcurlyeq 0$ and

$$\begin{pmatrix} y_\beta & v_\beta \\ v_\beta & y_\beta \end{pmatrix} \succcurlyeq 0. \tag{5.26}$$

Note that, as for odd circuits (A, β) we don't necessarily have $v_\beta \geq 0$ for all v in the dual cone, the second argument of $C^*_{A,\beta}(v_{|A}, y_\beta, y)$ is now y_β instead of v_β, as we had in Theorem 5.3.5.

Proof. As we only consider *odd* circuits, $|A| > 1$ in every case.

1. The semidefinite condition on the matrix (5.25) is equivalent to $x_\beta \geq 0$ and $|c_\beta| \leq x_\beta$. Hence, altogether we obtain

$$f \in P^{\text{odd}}_{A,\beta} \text{ if and only if } |c_\beta| \leq \prod_{\alpha \in A} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha}$$

   for barycentric coordinates $\lambda \in \mathbb{R}^A_+$ decomposing β as a convex combination of elements in A. This is exactly Theorem 5.1.8(b).

2. If $v \in (P^{\text{odd}}_{A,\beta})^*$, then, in the notation of (5.10), there exists some u such that $(v, u) \in (P^{\text{even}}_{A,\beta})^*$ and $u \geq |v_\beta|$. In particular, $u \geq 0$ is necessary for containment in $\left(P^{\text{even}}_{A,\beta}\right)^*$. The semidefinite constraints (5.26) are equivalent to $y_\beta \geq 0$ and the latter inequality $u \geq |v_\beta|$, and the constraint $C^*_{A,\beta}(v_{|A}, y_\beta, y) \succcurlyeq 0$ is equivalent to $(v, y_\beta) \in \left(P^{\text{even}}_{A,\beta}\right)^*$ by Theorem 5.3.5.

□

Now, for every *odd* reduced circuit $(A, \beta) \in R(\mathcal{A}, \mathcal{B})$, define the block diagonal matrix $\widehat{C}^*_{A,\beta}(v_{|A\cup\{\beta\}}, y_\beta, y)$ consisting of the dual circuit matrix $C^*_{A,\beta}(v_{|A\cup\{\beta\}}, y_\beta, y)$ and (5.25) for the dual cone. Considering all the reduced circuits, these lifting matrices define the lifted cone

$$\widehat{C}^*(\mathcal{A}, \mathcal{B}) = \Big\{(v, y^{\mathcal{A},\mathcal{B}}) \ : \widehat{C}^*_{A,\beta}(v_{|A\cup\{\beta\}}, y_\beta, y) \succcurlyeq 0 \text{ for all } (A, \beta) \in R(\mathcal{A}, \mathcal{B}), \\ C^*_{A,\beta}(v_{|A}, v_\beta, y) \succcurlyeq 0 \text{ for all } (A, \beta) \in R(\mathcal{A}, \mathcal{A})\Big\},$$

where the variable vector v lives in the space $\mathbb{R}^{\mathcal{A},\mathcal{B}}$.

For a fixed odd reduced circuit $(A, \beta) \in R(\mathcal{A}, \mathcal{B})$, let

$$\widehat{P^{\text{odd}}_{A,\beta}} = \Big\{(c, x^{\mathcal{A},\mathcal{B}}) \ : \ \widehat{C}_{A,\beta}(c_{|A\cup\{\beta\}}, x_\beta, x^{A,\beta}) \succcurlyeq 0, c_{|\mathcal{A}\cup\mathcal{B}\setminus(A\cup\{\beta\})} = 0\Big\},$$

where $\widehat{C}_{A,\beta}(c_{|A\cup\{\beta\}}, x_\beta, x^{A,\beta})$ is defined analogously to the dual case. We define the lifted cone

$$\widehat{C}(\mathcal{A}, \mathcal{B}) = \sum_{(A,\beta)\in R(\mathcal{A},\mathcal{A})} \widehat{P^{\text{even}}_{A,\beta}} + \sum_{(A,\beta)\in R(\mathcal{A},\mathcal{B})} \widehat{P^{\text{odd}}_{A,\beta}}.$$

Here, for every $(A, \beta) \in R(\mathcal{A}, \mathcal{A})$, $\widehat{P^{\text{even}}_{A,\beta}}$ is the set from Theorem 5.3.12.

Corollary 5.3.17.

(1) The rational $\mathcal{S}$-cone $C_\mathcal{S}(\mathcal{A}, \mathcal{B})$ is the projection on the coordinates $v \in \mathbb{R}^{\mathcal{A},\mathcal{B}}$ of $\widehat{C}(\mathcal{A}, \mathcal{B})$.

(2) The dual of the rational $\mathcal{S}$-cone $C^*_\mathcal{S}(\mathcal{A}, \mathcal{B})$ is the projection on the coordinates $v \in \mathbb{R}^{\mathcal{A},\mathcal{B}}$ of $\widehat{C}^*(\mathcal{A}, \mathcal{B})$.

Applying this lifting to the second-order representations of Theorems 5.3.12 and 5.3.5 in standard form also gives second-order representations of the cones $C_\mathcal{S}(\mathcal{A}, \mathcal{B})$ and $C^*_\mathcal{S}(\mathcal{A}, \mathcal{B})$ in standard form.

Corollary 5.3.18 (Second-order representation of the dual rational $\mathcal{S}$-cone)**.** *A vector $v \in \mathbb{R}^{(\mathcal{A},\mathcal{B})}$ is contained in the rational $\mathcal{S}$-cone $C^*_\mathcal{S}(\mathcal{A}, \mathcal{B})$ if and only if the dual circuit vector $y^{\mathcal{A},\mathcal{B}}$ satisfies for every reduced odd circuit $(A, \beta) \in R(\mathcal{A}, \mathcal{B})$*

1. $\begin{pmatrix} y^{A,\beta}_{k-1,2i-1} & y^{A,\beta}_{k,i} \\ y^{A,\beta}_{k,i} & y^{A,\beta}_{k-1,2i} \end{pmatrix} \succeq 0, \quad 2 \leq k \leq \lceil \log_2(p_{A,\beta}) \rceil \ \forall i \in [2^{\lceil \log_2(p_{A,\beta})\rceil - k}]$,

2. $\begin{pmatrix} y^{A,\beta}_{\lceil \log_2(p_{A,\beta})\rceil,1} & y^{A,\beta}_{\beta} \\ y^{A,\beta}_{\beta} & y^{A,\beta}_{\lceil \log_2(p_{A,\beta})\rceil,2} \end{pmatrix} \succeq 0$,

3. $\begin{pmatrix} u & y^{A,\beta}_{1,l} \\ y^{A,\beta}_{1,l} & w \end{pmatrix} \succeq 0$ *for* $l \in [2^{\lceil \log_2(p_{A,\beta})\rceil - 1}]$ *and* $u, w \in \{v_\alpha : \alpha \in A\} \cup \{y^{A,\beta}_\beta\}$ *such that* $y^{A,\beta}_\beta$ *appears* $2^{\lceil \log_2(p_{A,\beta})\rceil} - p_{A,\beta}$ *times and* v_α *appears* $(p_{A,\beta})_\alpha$ *times for each* $\alpha \in A$,

4. $||v_\beta||_2 \leq y^{A,\beta}_\beta, \quad y^{A,\beta}_\beta \leq y^{A,\beta}_{\lceil \log_2(p_{A,\beta})\rceil,1}$

and for every reduced even circuit $(A, \beta) \in R(\mathcal{A}, \mathcal{A})$ *the conditions of Theorem 5.3.5.*

We need to write $y^{A,\beta}$ instead of just writing y in the previous corollary, since different $y^{A,\beta}$ may appear for every reduced circuit (A, β).

For the primal case, we have to consider every reduced circuit as well. Here, sums take the role of the intersections from the dual case.

Corollary 5.3.19 (A Second-Order Representation of the Rational $\mathcal{S}$-Cone)**.** *A function* $f \in \mathbb{R}[\mathcal{A}, \mathcal{B}]$ *with coefficient vector* c *is contained in the rational* $\mathcal{S}$*-cone* $C_{\mathcal{S}}(\mathcal{A}, \mathcal{B})$ *if and only if there exists* $c^{A,\beta}$ *for* $(A, \beta) \in R(\mathcal{A}, \mathcal{A} \cup \mathcal{B})$ *with* $c = \sum_{(A,\beta)\in R(\mathcal{A},\mathcal{A}\cup\mathcal{B})} c^{A,\beta}$ *and for the circuit vector* $x^{\mathcal{A},\mathcal{B}}$ *and for every* $(A, \beta) \in R(\mathcal{A}, \mathcal{A} \cup \mathcal{B})$, *the following inequalities hold.*

1. $\begin{pmatrix} x^{A,\beta}_{k-1,2i-1} & x^{A,\beta}_{k,i} \\ x^{A,\beta}_{k,i} & x^{A,\beta}_{k-1,2i} \end{pmatrix} \succcurlyeq 0,\ 2 \leq k \leq \lceil \log_2(p_{A,\beta}) \rceil,\ i \in [2^{\lceil \log_2(p_{A,\beta})\rceil - k}]$,

2. $x^{A,\beta}_{\lceil \log_2(p_{A,\beta})\rceil,1} - \left(\prod_{\alpha\in A} \lambda_\alpha^{(p_{A,\beta})_\alpha}\right) x^{A,\beta}_\beta \geq 0$,

3. $x^{A,\beta}_\beta + c_\beta \geq 0$,

4. $||c_\beta||_2 \leq x^{A,\beta}_\beta$ *if* (A, β) *is an odd circuit,*

5. *as well as in both the even and the odd case,*

$$\begin{pmatrix} u & x^{A,\beta}_{1,l} \\ x^{A,\beta}_{1,l} & w \end{pmatrix} \succcurlyeq 0 \qquad \text{for } l \in [2^{\lceil \log_2(\lambda_{A,\beta})\rceil - 1}]$$

for $u, w \in \{c_\alpha : \alpha \in A\} \cup \left\{\left(\prod_{\alpha\in A} \lambda_\alpha^{(\lambda_{A,\beta})_\alpha}\right) x^{A,\beta}_\beta\right\}$ *such that* c_α *appears* $(p_{A,\beta})_\alpha$ *times for every* $\alpha \in A$ *and* $\left(\prod_{\alpha\in A} \lambda_\alpha^{(\lambda_{A,\beta})_\alpha}\right) x^{A,\beta}_\beta$ *appears* $2^{\lceil \log_2(p_{A,\beta})\rceil} - p_{A,\beta}$ *times.*

Note that the cone $C_{\mathrm{SONC}}(\mathcal{A})$ of SONC *polynomials* is a closed convex cone, and it can be recognized as a special case of a rational $\mathcal{S}$-cone by observing

$$C_{\mathrm{SONC}}(\mathcal{A}) = C_{\mathcal{S}}(\mathcal{A} \cap (2\mathbb{N})^n, \mathcal{A} \cap (\mathbb{N}^n \setminus (2\mathbb{N})^n)).$$

Hence, both the cone of SONC polynomials $C_{\mathrm{SONC}}(\mathcal{A})$ and its dual are always rational $\mathcal{S}$-cones and thus, occur as a special case of Corollaries 5.3.19 and 5.3.18.

Remark 5.3.20. The specific case of the primal cone of SONC polynomials has also been studied in detail by Magron and Wang [WM20a]. Their approach is based on different methods. In particular, it relies on mediated sets and intermediately uses sums of squares representations. However, the resulting second-order programs are structurally similar. Notably, the dependence of the size of the second-order program on the parameter p in our derivation relates to the dependency on the size of the rational mediated set in [WM20a]. Also note that various amendments are integrated into the approaches (such as the handling of denominators in [WM20a] and the use of extreme rays in our approach).

Chapter 6

Sublinear Circuits and the Conditional SAGE-Cone

In this chapter, we start examining the constrained version of a global optimization problem for exponential sums, i.e., the problem

$$\min_{x \in X} f(x)$$

or some convex set $X \subseteq \mathbb{R}^n$ and some exponential sum f.

Historically, approaches examining the conditional SAGE-cone considered cones of functions

$$\sum_{\alpha \in A} c_\alpha e^{\langle \alpha, x \rangle} + c_\beta e^{\langle \beta, x \rangle}$$

(or their polynomial equivalents, see, e.g., [DIW19]), where (A, β) is a circuit in the sense of $\mathbb{R}^n$, with nonnegative coefficients corresponding to the set A of outer exponents. As already mentioned in the preliminaries, here, we build upon the circuit definition of Forsgård and de Wolff [FW19]: They define classical $\mathbb{R}^n$-circuits as vectors $\nu \in \mathbb{R}^{\mathcal{A}}$ with the following three conditions: The entries of ν sum up to 0, $\sum_{\alpha \in \mathcal{A}} \alpha \nu_\alpha = 0$ and exactly one entry of ν is negative. They call circuits with this last condition *simplicial*; we drop this term as we are *only* interested in simplicial circuits. If ν is normalized, i.e., the negative entry equals -1, then the positive entries of ν represent the barycentric coordinates of $\mathcal{A}$ with respect to β in the case of affine independence of $\mathcal{A} \setminus \{\beta\}$.

Here, instead of focusing on $\mathbb{R}^n$-circuits, we develop a new approach using so-called *sublinear circuits*, see Definition 6.1.2. Sublinear circuits are generalizations of the affine circuits from matroid theory. They arise as the convex-combinatorial core underlying constrained nonnegativity certificates of exponential sums and of polynomials based on the arithmetic-geometric inequality.

In Section 6.1, we start with the definition of these *X-circuits* of a point set $\mathcal{A}$, generalizing $\mathbb{R}^n$-circuits to a constrained setting. After revealing various elementary properties and discussing some examples, we characterize X-circuits in more geometric terms in Theorems 6.1.7 and 6.1.8. In particular, the latter theorem interprets X-circuits in terms of normal fans for the case when X is a polyhedron. It also shows that for polyhedral constraint sets the number of X-circuits is finite. Building upon this, we determine the X-circuits of a univariate support set $\mathcal{A} \subseteq \mathbb{R}$ for $X = [-1, 1]$ and $X = \mathbb{R}_+$ in Theorem 6.1.9.

In this part of the thesis, we only examine the case of *exponential sums* instead of generalized polynomial and exponential functions as introduced in the previous chapter. However, we can generalize the results of this chapter to a constrained version of the $\mathcal{S}$-cone and, with this, also to polynomials as a special case of these functions; for further explanations see Subsection 6.2.1.

In Section 6.2, we build upon the theory by Murray, Chandrasekaran, and Wierman [MCW21b] concerning X-AGE exponentials and X-SAGE exponentials as introduced in the preliminaries as the constrained equivalents to AGE exponentials and SAGE exponentials: We show that sums of X-AGE exponentials can be characterized in terms of sublinear circuits, see Theorem 6.2.2. Building upon this result, we define *λ-witnessed X-AGE-cones* and show that these cones actually construct the cone of X-AGE exponentials, see Theorem 6.2.4. Those λ-witnessed X-AGE-cones are in fact dual power-cones. Combined with the finiteness of the number of X-circuits in the case of a polyhedral constraint set, this shows that the *whole* X-SAGE-cone is power-cone representable whenever the constraint set is polyhedral. A similar representation exists for the *dual* of the X-SAGE-cone, see Proposition 6.2.7.

In Subsection 6.2.2, we give results on the relation between the sublinear circuits and their supports, and in Subsection 6.2.3, we provide necessary as well as sufficient criteria for sublinear circuits, see, e.g., Theorem 6.2.11. Building upon the results for polyhedral constraint sets and based on the former necessary or sufficient characterizations, we provide some explicit results for conic constraint sets as well as enumerations for the cube $[-1, 1]^n$ in Subsection 6.2.4.

In Section 6.3, we introduce the concept of *reduced sublinear circuits* as extremal rays of the *circuit graph*, see Definition 6.3.3. In fact, it suffices to consider these circuits to construct the X-SAGE-cone, see Theorem 6.3.5. For the case of a polyhedral constraint set, we prove that the set of reduced X-circuits is in fact the smallest set allowing such a construction, see Theorem 6.3.6.

This again motivates studying the case of a polyhedral constraint set in detail. As for X-circuits, we develop sufficient as well as necessary criteria for a vector ν to be a reduced X-circuit, see Subsection 6.3.2. We put a particular focus on the case of $X = \mathbb{R}^n_+$ and $X = [-1, 1]^n$; for the univariate special cases of these two classes, we state the set of X-circuits completely, see Theorem 6.3.12.

In the last part of this chapter, Section 6.4, those two examples, $X = [-1, 1]$ and $X = \mathbb{R}_+$, culminate in a theorem completely characterizing the extreme rays of the resulting X-SAGE-cones $C_{[-1,1]}(\mathcal{A})$ and $C_{[0,\infty)}(\mathcal{A})$.

One of the papers this chapter is based on — [MNT20] — contains various directions of research. Particularly, in various parts, we studied general convex-geometric statements, considering general convex — and non-polyhedral — sets X. This thesis focusses on the concept of sublinear circuits and statements for polyhedral constraint sets. Therefore, some technical convex-geometric lemmas will be cited without proof.

6.1 Sublinear Circuits Induced by a Point Set

Throughout this chapter, we assume that $X \subseteq \mathbb{R}^n$ is a closed, convex, and non-empty set.

Recall from Section 2.4.3 that for a finite and non-empty set of support points $\mathcal{A} \subseteq \mathbb{R}^n$ and some convex and non-empty set $X \subseteq \mathbb{R}^n$, the cone $C_X(\mathcal{A}, \beta)$ of X-AGE exponentials supported on $\mathcal{A}$ with negative term corresponding to $\beta \in \mathcal{A}$ is the set of those exponential sums

$$f = \sum_{\alpha \in \mathcal{A} \setminus \{\beta\}} c_\alpha e^{\langle \alpha, x \rangle} + c_\beta e^{\langle \beta, x \rangle}$$

where f is nonnegative on X and $c_{\setminus\beta} \in \mathbb{R}_+^{\mathcal{A}\setminus\{\beta\}}$, i.e., c contains at most one negative component which has to correspond to the index β. As in the unconstrained case, we often overload notation and also treat $C_X(\mathcal{A}, \beta)$ as a cone of coefficient vectors in $\mathbb{R}^{\mathcal{A}}$.

We say that the cone of *sums of X-AGE exponentials* with respect to exponents $\mathcal{A}$ is the Minkowski sum

$$C_X(\mathcal{A}) = \sum_{\beta\in\mathcal{A}} C_X(\mathcal{A}, \beta).$$

Note that this cone is not defined in terms of circuits.

For vectors $\beta \in \mathcal{A}$, let

$$N_\beta = \{\nu \in \mathbb{R}^{\mathcal{A}} : \nu_{\setminus\beta} \geq \mathbf{0},\ \mathbf{1}^T\nu = 0\}. \tag{6.1}$$

Recall that Murray, Chandrasekaran, and Wierman showed that it is possible to efficiently check membership in this cone using a relative entropy program.

Proposition 6.1.1 ([MCW21b, Theorem 6])**.** *An exponential $f(x) = \sum_{\alpha\in\mathcal{A}} c_\alpha e^{\langle\alpha,x\rangle}$ belongs to $C_X(\mathcal{A}, \beta)$ if and only if $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A} \setminus \{\beta\}$ and some $\nu \in \mathbb{R}^{\mathcal{A}}$ satisfies*

$$\nu \in N_\beta \quad \text{and} \quad \sigma_X(-\mathcal{A}\nu) + D(\nu_{\setminus\beta}, ec_{\setminus\beta}) \leq c_\beta. \tag{6.2}$$

Recall that here, for $\mathcal{A} \subseteq \mathbb{R}^n$ and $\nu \in \mathbb{R}^{\mathcal{A}}$, we use $\mathcal{A}$ as a linear operator $\mathcal{A} : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}, \nu \mapsto \mathcal{A}\nu = \sum_{\alpha\in\mathcal{A}} \nu_\alpha \alpha$. For $\mathcal{A} \subseteq \mathbb{R}^n$ and $X \subseteq \mathbb{R}^n$, we remind the reader of the definition of the relative entropy function $D : \mathbb{R}^{\mathcal{A}}_{>0} \times \mathbb{R}^{\mathcal{A}}_{>0} \to \mathbb{R}$ with $D(\nu, \lambda) = \sum_{\alpha\in\mathcal{A}} \nu_\alpha \ln(\nu_\alpha/\lambda_\alpha)$ and of the support function $\sigma_X : \mathbb{R}^n \to \mathbb{R}$ with $\sigma_X(\nu) = \sup_{x\in X} \nu^T x$.

Following the language introduced by Murray, Chandrasekaran, and Wierman, we call the resulting cones $C_X(\mathcal{A}, \beta)$ the *X-AGE-cone* and $C_X(\mathcal{A})$ the *X-SAGE-cone.*

The following central definition of this section helps us to examine this set N_β.

Definition 6.1.2. A vector $\nu^\star \in N_\beta$ is an *X-circuit of $\mathcal{A}$* (or simply, an *X-circuit*) if

1. it is nonzero,
2. $\sigma_X(-\mathcal{A}\nu^\star) < \infty$,
3. and it cannot be written as a convex combination of two non-proportional $\nu^{(1)}, \nu^{(2)} \in N_\beta$ for which $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ is linear on $[\nu^{(1)}, \nu^{(2)}]$.

The derived results in this section make no mention of exponential sums, and to avoid dependence on relative entropy in this section, we frame our discussion of X-circuits in terms of the cones N_β for $\beta \in \mathcal{A}$. However, the definition of X-circuits is ultimately chosen to prepare for studying X-SAGE-cones.

The third condition of Definition 6.1.2 is equivalent to strict sublinearity of the mapping $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ on any line segment in N_β that contains $\nu^\star$ except for the trivial line segments which generate a single ray. The central importance of the sublinearity condition leads us to refer to X-circuits also as *sublinear circuits.*

Remark 6.1.3. In the special case $X = \mathbb{R}^n$, condition (2) simplifies to $\mathcal{A}\nu = \mathbf{0}$. In conjunction with the definition of N_β, this shows that the special case $X = \mathbb{R}^n$ of Definition 6.1.2 exactly matches the definition of $\mathbb{R}^n$-circuits. The only difference is that we do not have circuits supported on a single element anymore.

Conceptually, Definition 6.1.2 indicates that X-circuits are essential in capturing the behavior of the augmented support function $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ on the given set N_β. While developing this concept formally, it is convenient for us to denote the set $\nu^+ := \{\alpha : \nu_\alpha > 0\}$, and to identify the index $\nu^- := \beta \in \mathcal{A}$ where $\nu_\beta < 0$. By the definition of N_β and the nonzero condition of an X-circuits, this index is unique. Note that positive homogeneity of the support function tells us that the property of being a sublinear circuit is invariant under scaling by positive constants. A sublinear circuit is *normalized* if its unique negative term ν_β has $\nu_\beta = -1$, in which case we usually denote it by the symbol λ rather than ν. We can normalize a given sublinear circuit by taking the ratio with its infinity norm $\lambda = \nu/\|\nu\|_\infty$ because $\|\nu\|_\infty = |\nu_\beta|$ for all vectors $\nu \in N_\beta$.

Example 6.1.4. *(The Conic Case.)* It is straightforward to determine which $\nu \in N_\beta$ are X-circuits of $\mathcal{A}$ when X is a cone. In such a setting, the support function of X can only take on the values zero and positive infinity. Hence, $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ is trivially linear over all of $V_\beta := \{\nu \in N_\beta : \sigma_X(-\mathcal{A}\nu) < \infty\}$. The set V_β is a cone, and reformulating $\sigma_X(-\mathcal{A}\nu) = 0$ as $\nu \in (\mathcal{A}^T X)^*$ gives

$$V_\beta = (\ker \mathcal{A} + \mathcal{A}^+ X^*) \cap N_\beta,$$

where $\mathcal{A}^+$ denotes the pseudo-inverse of $\mathcal{A} : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}^n$. Therefore, the X-circuits $\nu \in N_\beta$ are precisely the edge generators of $(\ker \mathcal{A} + \mathcal{A}^+ X^*) \cap N_\beta$.

Looking again at the special case $X = \mathbb{R}^n$ from this conic perspective, we have $X^* = \{\mathbf{0}\}$, yielding $\mathcal{A}^+ X^* = \{\mathbf{0}\}$, and $\ker \mathcal{A} + \mathcal{A}^+ X^* = \ker \mathcal{A}$, which implies that $V_\beta = \ker \mathcal{A} \cap N_\beta$. It is easy to show that edge generators of $\ker \mathcal{A} \cap N_\beta$ are precisely those $\nu \in \ker \mathcal{A} \cap N_\beta \setminus \{\mathbf{0}\}$ for which $\nu^+ = \{\alpha : \nu_\alpha > 0\}$ are affinely independent, which recovers the matroid-theoretic notion of affine-linear simplicial circuits from the point of view of subsets $A \subseteq \mathcal{A}$.

The following proposition shows that the affine-independence property is a necessary condition for all sublinear circuits. It provides insight because it shows that an X-circuit ν with $X \subseteq \mathbb{R}^n$ is restricted to $|\operatorname{supp} \nu| \leq n + 2$.

Proposition 6.1.5. *If $\nu^\star \in N_\beta$ is an X-circuit, then $(\nu^\star)^+ = \operatorname{supp} \nu^\star \setminus \{\beta\}$ is affinely independent.*

Proof. We fix $\nu^\star \in N_\beta$, set $z = -\mathcal{A}\nu^\star$ and $U = \{\nu \in N_\beta : -\mathcal{A}\nu = z, \, \nu_\beta = \nu^\star_\beta\}$. The function $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ is a constant and equal to $\sigma_X(z)$ on U, and so in order for $\nu^\star$ to be an X-circuit it must be a vertex of the polytope U. The set U is in one-to-one correspondence with $W = \{w \in \mathbb{R}_+^{\mathcal{A}\setminus\{\beta\}} : \sum_{\alpha \in \mathcal{A}\setminus\{\beta\}} (\beta - \alpha) w_\alpha = z, \, \mathbf{1}^T w = -\nu^\star_\beta\}$ by identifying $w = \nu_{\setminus\beta}$. For the matrix M with columns $\{(\beta - \alpha, 1)\}_{\alpha \in \mathcal{A}\setminus\{\beta\}}$ indexed by $\alpha \in \mathcal{A} \setminus \{\beta\}$, we can write $W = \{w \in \mathbb{R}_+^{\mathcal{A}\setminus\{\beta\}} : Mw = (z, -\nu^\star_\beta)\}$.

Basic polyhedral geometry tells us that all vertices $w^\star$ of W use an affinely independent set of columns from M. Furthermore, a given set of columns from M is affinely independent if and only if the corresponding indices of the columns (as vectors $\alpha \in \mathcal{A} \setminus \{\beta\}$) are affinely independent. Since the correspondence between $\nu \in U$ and $w \in W$ preserves extremality, the vertices of U have affinely independent positive support ν^+. $\square$

The converse of Proposition 6.1.5 is not true. This is to say: Not every vector $\nu \in N_\beta$ with affinely independent ν^+ is an X-circuit.

Example 6.1.6. Let $\mathcal{A} \subseteq \mathbb{R}^2$ contain $\alpha_1 = (0, 0)$, $\alpha_2 = (1, 0)$, and $\alpha_3 = (0, 1)$, and consider $X = \{x \in \mathbb{R}^2 : x \geq u\}$ for some fixed point $u \in \mathbb{R}^2$. Observe that the vector

$\nu^\star = (-2, 1, 1)$ has $(\nu^\star)^- = \alpha_1 = (0, 0)$, and $(\nu^\star)^+ = \{\alpha_2, \alpha_3\} = \{(1, 0), (0, 1)\}$ is affinely independent. Considering $\nu^{(1)} = (-2, 2, 0)$ and $\nu^{(2)} = (-2, 0, 2)$, we have $\nu^\star = \frac{1}{2}(\nu^{(1)} + \nu^{(2)}) \in \operatorname{relint} L$ for $L := [\nu^{(1)}, \nu^{(2)}]$. Moreover, $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ is linear on L because for any $\mu_1, \mu_2 \geq 0$ with $\mu_1 + \mu_2 = 1$, we have

$$\begin{aligned}\sigma_X(\mathcal{A}(-\mu_1\nu^{(1)} - \mu_2\nu^{(2)})) &= \sigma_X((-2\mu_1, -2\mu_2)) = -2\mu_1 u_1 - 2\mu_2 u_2 \\ &= \sigma_X((-2\mu_1, 0)) + \sigma_X((0, -2\mu_2)).\end{aligned}$$

The last equality is true since the element $(1, 1)^T$ maximizes both the objective functions $x \mapsto (-2\mu_1, 0)^T x$ and $x \mapsto (0, -2\mu_2)^T x$ on X.

With the basic exercise of Example 6.1.6 complete, we start characterizing sublinear circuits in full generality by mentioning the following theorem from [MNT20] without proof.

Theorem 6.1.7. *Fix $\beta \in \mathcal{A}$. The convex cone generated by*

$$T = \{\, (\nu, \sigma_X(-\mathcal{A}\nu)) \,:\, \nu \in N_\beta,\ \sigma_X(-\mathcal{A}\nu) < \infty\}$$

is pointed (i.e., it contains no lines) and closed. A vector $\nu^\star \in N_\beta$ is an X-circuit of $\mathcal{A}$ if and only if $(\nu^\star, \sigma_X(-\mathcal{A}\nu^\star))$ is an edge generator for $\operatorname{pos} T$.

When considering the set T in Theorem 6.1.7, it is natural to expect that for polyhedral X there are only finitely many extreme rays in $\operatorname{pos} T$, and hence, only finitely many normalized X-circuits. To prove this fact, we use the concept of *normal fans* from polyhedral geometry, see, e.g., [Zie95, Chapter 7] (for the bounded case of polytopes), [GKZ94, Section 5.4] or [Stu96, Chapter 2]. For each face F of a polyhedron P, there is an associated *outer normal cone*

$$\mathsf{N}_P(F) = \{w \,:\, z^T w = \sigma_P(w)\ \forall\, z \in F\}.$$

Clearly, the support function of a polyhedron P is linear on every outer normal cone, and, in particular, the linear representation may be given by $\sigma_P(w) = z^T w$ for any $z \in F$. We obtain the *outer normal fan* of P by collecting all outer normal cones:

$$\mathcal{O}(P) = \{\mathsf{N}_P(F) \,:\, F \trianglelefteq P\}.$$

The support of $\mathcal{O}(P)$ is the polar $\operatorname{rec}(P)^\circ$. The full-dimensional linearity domains of the support function are the outer normal cones of the vertices of P (see also [FI17, Section 1]).

Theorem 6.1.8. *Let X be polyhedral. Then, $\nu \in N_\beta \setminus \{\mathbf{0}\}$ is an X-circuit if and only if* $\operatorname{pos}\{\nu\}$ *is a ray in $\mathcal{O}(-\mathcal{A}^T X + N_\beta^\circ)$. Consequently, polyhedral X have finitely many normalized circuits.*

If X is a polyhedral cone, the situation simplifies because the support function $\sigma_X(-\mathcal{A}\nu)$ of a circuit ν can only attain the values zero and infinity. Namely, since $\mathcal{O}(-\mathcal{A}^T X + N_\beta^\circ) = (\mathcal{A}^T X)^* \cap N_\beta$ and

$$(\mathcal{A}^T X)^* = \{\nu \,:\, \nu^T y \geq 0\ \forall y \in \mathcal{A}^T X\} = \{\nu \,:\, \sigma_X(-\mathcal{A}\nu) \leq 0\},$$

the X-circuits $\nu \in N_\beta$ are precisely the edge generators of $\{\nu \in N_\beta \,:\, \sigma_X(-\mathcal{A}\nu) \leq 0\}$. See Theorem 6.1.9 for the univariate case.

Proof of Theorem 6.1.8. Let $P = -\mathcal{A}^T X + N_\beta^\circ$. Using the characterization in [Roc97, Theorem 14.2], the polar of its recession cone can be expressed as

$$(\operatorname{rec} P)^\circ = \{\nu : \sigma_X(-\mathcal{A}\nu) < \infty\} \cap N_\beta,$$

where we have also used the property $\sigma_X(-\mathcal{A}\nu) = \sup_{x\in X}(-\mathcal{A}\nu)^T x = \sigma_{-\mathcal{A}^T X}(\nu)$. In particular, this also gives $\sigma_X(-\mathcal{A}\nu) = \sigma_P(\nu)$. From P, we construct the outer normal fan $\mathcal{O} := \mathcal{O}(P)$. We claim that $\operatorname{pos}\{\nu\}$ is a ray in $\mathcal{O}$.

It is clear that if a cone $K \in \mathcal{O}$ is associated to a face $F \trianglelefteq P$, then we see that $\sigma_P(\nu) = z^T\nu$ for any $z \in F$, and so $\sigma_P(\nu) \equiv \sigma_X(-\mathcal{A}\nu)$ is linear on K. Since the support of $\mathcal{O}$ is $\operatorname{rec}(P)^\circ$, the cones $K \in \mathcal{O}$ partition $(\operatorname{rec} P)^\circ$, i.e.,

$$(\operatorname{rec} P)^\circ = \bigcup_{K\in\mathcal{O}} \operatorname{relint}(K),$$

and if K, K' are distinct elements in $\mathcal{O}$, then $\operatorname{relint} K \cap \operatorname{relint} K' = \emptyset$. Therefore, every $\nu \in N_\beta \setminus \{\mathbf{0}\}$ for which $\sigma_X(-\mathcal{A}\nu) < \infty$ is associated with a unique $K \in \mathcal{O}$ by way of $\nu \in \operatorname{relint} K$.

Fix $\nu \in (\operatorname{rec} P)^\circ$ and let K be the associated element of $\mathcal{O}$ that contains ν in its relative interior. If K is of dimension greater than 1, ν can be expressed as a convex combination of non-proportional elements $\nu^{(1)}, \nu^{(2)} \in K$, and, clearly, then $\nu \mapsto \sigma_X(-\mathcal{A}\nu) \equiv \sigma_P(\nu)$ would be linear on the interval $[\nu^{(1)}, \nu^{(2)}]$. Thus, for ν to be an X-circuit, it is necessary that K is of dimension 1. Since P is a polyhedron, $\mathcal{O}$ is induced by finitely many faces. Thus, there are finitely many $K \in \mathcal{O}$ with $\dim K = 1$ and in turn finitely many normalized X-circuits of $\mathcal{A}$.

Conversely, let $\nu^\star \in N_\beta \setminus \{\mathbf{0}\}$ and $\operatorname{pos}\{\nu^\star\}$ be a ray in $\mathcal{O}$. Since $\mathcal{O}$ is supported on $\operatorname{rec}(P)^\circ$, we have $\sigma_X(-\mathcal{A}\nu) = \sigma_P(\nu) < \infty$.

Let $\nu^{(1)} \in N_\beta$ and $\nu^{(2)} \in N_\beta$ be non-proportional and let some $\tau \in (0,1)$ satisfy $\nu^\star = \tau\nu^{(1)} + (1-\tau)\nu^{(2)}$. If $\nu^{(1)}$ or $\nu^{(2)}$ is outside of $\operatorname{rec}(P)^\circ$, say, without loss of generality $\nu^{(1)}$, then $\sigma_X(-\mathcal{A}\nu^{(1)}) = \infty$, and thus, the mapping $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ cannot be linear on $[\nu^{(1)}, \nu^{(2)}]$. Hence, we can assume that $\nu^{(1)}, \nu^{(2)} \in \operatorname{rec}(P)^\circ$.

We have to show that the mapping

$$g : [0,1] \to \mathbb{R}, \quad \theta \mapsto \sigma_P(\theta\nu^{(1)} + (1-\theta)\nu^{(2)})$$

is not linear.

Consider the restriction of the fan $\mathcal{O}$ to the cone $C := \operatorname{pos}\{\nu^{(1)}, \nu^{(2)}\}$, that is, the collection of all the cones in $\{\mathsf{N}_P(F) \cap C : F \trianglelefteq P\}$. This is a fan $\mathcal{O}'$ supported on the two-dimensional cone $S := \operatorname{rec}(P)^\circ \cap C$. On the set S, we consider the restricted mapping $(\sigma_P)|_S : S \to \mathbb{R}$, $w \mapsto \sigma_P(w)$. The linearity domains of $(\sigma_P)|_S$ are the two-dimensional cones in $\mathcal{O}'$. Since $\operatorname{pos}\{v\}$ is a ray in the fan $\mathcal{O}$ and thus, also in the fan $\mathcal{O}'$, the vectors ν_1 and ν_2 are contained in different two-dimensional cones of the fan $\mathcal{O}'$. Hence, the mapping g is not linear. Altogether, this shows that ν is an X-circuit. $\square$

We begin with a study of the univariate cases $[0,\infty)$ and $[-1,1]$. This complements the unconstrained case $X = \mathbb{R}$ from Chapter 3.

Theorem 6.1.9. *Let $\mathcal{A} = \{\alpha_1, \dots, \alpha_m\} \subseteq \mathbb{R}$ with $\alpha_1 < \dots < \alpha_m$.*

(a) Let $X = [-1, 1]$. An element $\lambda \in \bigcup_{\beta \in \mathcal{A}} N_\beta$ is a normalized X-circuit if and only if it is of the form (1) $\lambda = \delta^{(j)} - \delta^{(i)}$ *for $i \neq j$, or* (2)

$$\lambda = \frac{\alpha_k - \alpha_j}{\alpha_k - \alpha_i}\delta^{(i)} - \delta^{(j)} + \frac{\alpha_j - \alpha_i}{\alpha_k - \alpha_i}\delta^{(k)} \text{ for } i < j < k.$$

(b) Let $X = \mathbb{R}_+$. The normalized X-circuits $\lambda \in \mathbb{R}^m$ are the vectors either of the form (1) $\lambda = \delta^{(k)} - \delta^{(j)}$ for $j < k$ or of the form (2)

$$\lambda = \frac{\alpha_j - \alpha_i}{\alpha_k - \alpha_i}\delta^{(k)} - \delta^{(j)} + \frac{\alpha_k - \alpha_j}{\alpha_k - \alpha_i}\delta^{(i)} \quad \text{for} \quad i < j < k.$$

Note that vectors of type (2) are equivalent in both cases and satisfy $\mathcal{A}\lambda = 0$, and in fact are the only vectors that also satisfy $\operatorname{supp}\lambda = \{i, j, k\}$, $\lambda_j = -1$, $\lambda_i, \lambda_k > 0$, $\mathbf{1}^T\lambda = 0$.

Remark 6.1.10. By Theorem 6.1.8, the X-circuits of $\mathcal{A}$ are the outer normal vectors to facets of polyhedra $P = -\mathcal{A}^T X + N_\beta^\circ$ (for some β). As N_β is pointed, P is always full-dimensional.

Proof of Theorem 6.1.9. (a) Fix $j \in [n]$ and write $N_j := N_{(\alpha_j)}$ for short. By Theorem 6.1.8, the X-circuits are the vectors spanning the rays in the outer normal cone of the polyhedron

$$\begin{aligned} P &= -\mathcal{A}^T X + N_j^\circ \\ &= \operatorname{conv}\{(\alpha_1, \dots, \alpha_m)^T, -(\alpha_1, \dots, \alpha_m)^T\} + \mathbb{R} \cdot \mathbf{1} - \sum_{i \neq j} \operatorname{pos}\{\delta^{(i)}\} \\ &= \{\theta(\alpha_1, \dots, \alpha_m)^T + \mu\mathbf{1} : -1 \leq \theta \leq 1, \mu \in \mathbb{R}\} - \sum_{i \neq j} \operatorname{pos}\{\delta^{(i)}\}. \end{aligned}$$

Hence, a point w is contained in P if and only if

$$w_i \leq \theta\alpha_i + \mu \text{ for } i \neq j \text{ and } w_j = \theta\alpha_j + \mu \quad \text{for } \theta \in [-1, 1] \text{ and } \mu \in \mathbb{R}.$$

By eliminating μ, this is equivalent to

$$w_j - w_i + \theta(\alpha_i - \alpha_j) \geq 0 \text{ for all } i \in [m] \setminus \{j\}, \ -1 \leq \theta \leq 1.$$

Eliminating θ then gives

$$\frac{w_j - w_i}{\alpha_j - \alpha_i} \begin{cases} \leq \theta \leq 1 & \text{if } \alpha_i > \alpha_j, \\ \geq \theta \geq -1 & \text{if } \alpha_i < \alpha_j, \end{cases}$$

which yields $\frac{w_j - w_i}{\alpha_j - \alpha_i} \geq \frac{w_k - w_j}{\alpha_k - \alpha_j}$ for all $i, k \in [m]$ with $i < j < k$ as well as $w_i - w_j \leq |\alpha_i - \alpha_j|$ for all $i \in [m] \setminus \{j\}$. Hence,

$$\begin{aligned} P = \{w \in \mathbb{R}^m : \ & w_i - w_j \leq |\alpha_i - \alpha_j| \text{ for } i \in [m] \setminus \{j\} \text{ and} && (6.3) \\ & w_i(\alpha_k - \alpha_j) - w_j(\alpha_k - \alpha_i) + w_k(\alpha_j - \alpha_i) \leq 0, i, k \in [m] : i < j < k\}. && (6.4) \end{aligned}$$

We claim that none of the inequalities in the definition of P is redundant. Namely, for each inequality

$$w_i(\alpha_k - \alpha_j) - w_j(\alpha_k - \alpha_i) + w_k(\alpha_j - \alpha_i) \leq 0$$

in (6.4), the point $\delta^{(i)} + \delta^{(j)} + \delta^{(k)}$ satisfies this particular inequality with equality and satisfy all of the other inequalities strictly. Similarly, for the inequalities in (6.3), it suffices to consider the point $\alpha_j \delta^{(i)} + \alpha_i \delta^{(j)}$ in the case that $i < j$ and $\alpha_i \delta^{(i)} + \alpha_j \delta^{(j)}$ in case $i > j$. By Remark 6.1.10, the polyhedron P is full-dimensional. Hence, by Theorem 6.1.8, the normalized X-circuits in N_j are exactly the ones given in the statement of the theorem.

(b) Again, we consider, for fixed $j \in [m]$, the polyhedron $P = -\mathcal{A}^T X + N_j^\circ$ from Theorem 6.1.8. For $X = \mathbb{R}_+$, this polyhedron may be expressed as

$$P \;=\; \text{pos}\{(-\alpha_1, \ldots, -\alpha_m)\} + \mathbb{R} \cdot \mathbf{1} - \sum_{\ell \in [m] \setminus j} \text{pos}\{\delta^{(l)}\}.$$

The rays of its normal fan are the extreme rays of its polar $P^\circ = (\text{rec}\, P)^\circ$ with

$$P^\circ = \{\nu \in \mathbb{R}^m \,:\, (-\alpha_1, \ldots, -\alpha_m)^T \nu \leq 0, \;\; \mathbf{1}^T \nu = 0, \; \nu_\ell \geq 0 \text{ for } \ell \in [m] \setminus j\}. \tag{6.5}$$

By Proposition 6.1.5, each X-circuit in N_j has at most three non-vanishing components ν_i, ν_j, ν_k, and, moreover, $m-2$ of the inequalities in (6.5) are binding. If all those binding inequalities are of the form $\nu_\ell \geq 0$, then with $\sigma_X(-\mathcal{A}\nu) < \infty$, we obtain the normalized X-circuits of $\mathcal{A}$ of type (1). Now assume that the inequality $(-\alpha_1, \ldots, -\alpha_m)^T \nu \leq 0$ is binding for some normalized X-circuit ν of $\mathcal{A}$. Since the sign pattern $(-,+,+)$ for (ν_i, ν_j, ν_k) in conjunction with $\mathbf{1}^T \nu = 0$ leads to $(-\alpha_1, \ldots, -\alpha_m)^T \nu < 0$, and the sign pattern $(+,+,-)$ contradicts the X-circuit condition $\sigma_X(-\mathcal{A}\nu) < \infty$, we obtain the normalized X-circuits of $\mathcal{A}$ of type (2).

□

6.2 Sublinear Circuits in X-AGE-Cones

In this section, we show how the X-AGE-cones $C_X(\mathcal{A}, \beta)$ can be further decomposed using sublinear circuits. These decompositions lay the foundation to understand the extreme rays of the conditional SAGE-cone $C_X(\mathcal{A})$. Our first result here is a necessary criterion for an X-AGE exponential f to be extremal in $C_X(\mathcal{A}, \beta)$, which states that any underlying ν must be an X-circuit (see Theorem 6.2.2). Definition 6.2.3 introduces λ-*witnessed* X-*AGE-cones* as the subset of exponential sums in $C_X(\mathcal{A}, \beta)$ whose nonnegativity is certified by a given normalized vector λ. Theorem 6.2.4 then decomposes $C_X(\mathcal{A}, \beta)$ through the λ-witnessed X-AGE-cones where λ is a normalized X-circuit. As a consequence, for polyhedral X, the cone $C_X(\mathcal{A}, \beta)$ is power-cone representable (see Corollary 6.2.5).

In the last part of this section we consider certain necessary and sufficient conditions describing properties of X-circuits.

The following lemma provides a construction to decompose an X-AGE exponential into simpler summands, under a local linearity condition on the support function $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$.

Lemma 6.2.1. *Let $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x\rangle}$ be X-AGE with negative coefficient $c_\beta < 0$. If a vector $\nu \in \mathbb{R}^{\mathcal{A}}$ satisfying* (2.13)*, i.e., with*

$$\nu \in N_\beta \quad \text{and} \quad \sigma_X(-\mathcal{A}\nu) + D(\nu_{\setminus\beta}, ec_{\setminus\beta}) \le c_\beta,$$

can be written as a convex combination $\nu = \sum_{i=1}^k \theta_i \nu^{(i)}$ of non-proportional $\nu^{(i)} \in N_\beta$ and $\tilde{\nu} \mapsto \sigma_X(-\mathcal{A}\tilde{\nu})$ is linear on $\operatorname{conv}\{\nu^{(i)}\}_{i=1}^k$, then f is not extremal in $C_X(\mathcal{A}, \beta)$.

Proof. Construct vectors $c^{(i)}$ by

$$c_\alpha^{(i)} = \begin{cases} (c_\alpha/\nu_\alpha)\nu_\alpha^{(i)} & \text{if } \alpha \in \nu^+ \\ 0 & \text{otherwise} \end{cases} \qquad \text{for all } \alpha \in \mathcal{A} \setminus \{\beta\}, \tag{6.6}$$

and $c_\beta^{(i)} = \sigma_X(-\mathcal{A}\nu^{(i)}) + D(\nu_{\setminus\beta}^{(i)}, ec_{\setminus\beta}^{(i)})$. These $c^{(i)}$ define X-AGE exponentials by construction, and they inherit non-proportionality from the $\nu^{(i)}$. We need to show that $\sum_{i=1}^k \theta_i c^{(i)} \le c$, which will establish that f can be decomposed as a sum of these non-proportional X-AGE exponentials (possibly with added exponential monomials).

For indices $\alpha \in \nu^+$, the construction (6.6) relative to ν and $\{\nu^{(i)}\}_{i=1}^k$ ensures $\sum_{i=1}^k \theta_i c_\alpha^{(i)} = c_\alpha$. For indices $\alpha \in \operatorname{supp} c \setminus \operatorname{supp} \nu$, we have $\sum_{i=1}^k \theta_i c_\alpha^{(i)} = 0 \le c_\alpha$. The definitions of $\nu^{(i)}$ ensure

$$\sigma_X(-\mathcal{A}\nu) = \sigma_X\left(-\mathcal{A}(\textstyle\sum_{i=1}^k \theta_i \nu^{(i)})\right) = \textstyle\sum_{i=1}^k \theta_i \sigma_X(-\mathcal{A}\nu^{(i)}). \tag{6.7}$$

Meanwhile, the construction (6.6) provides $\nu_\alpha^{(i)}/c_\alpha^{(i)} = \nu_\alpha/c_\alpha$, which may be combined with $\sum_{i=1}^k \theta_i \nu_\alpha^{(i)} = \nu_\alpha$ for all $\alpha \in \mathcal{A}$ to deduce

$$\sum_{i=1}^k \theta_i D(\nu_{\setminus\beta}^{(i)}, ec_{\setminus\beta}^{(i)}) = D(\nu_{\setminus\beta}, ec_{\setminus\beta}). \tag{6.8}$$

We combine (6.7) and (6.8) to obtain the desired result

$$\sum_{i=1}^k \theta_i c_\beta^{(i)} = \sum_{i=1}^k \theta_i \left(\sigma_X(-\mathcal{A}\nu^{(i)}) + D(\nu_{\setminus\beta}^{(i)}, ec_{\setminus\beta}^{(i)})\right) = \sigma_X(-\mathcal{A}\nu) + D(\nu_{\setminus\beta}, ec_{\setminus\beta}) \le c_\beta.$$

□

Theorem 6.2.2. *Let $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x\rangle}$ be an X-AGE exponential with negative coefficient $c_\beta < 0$. If $\nu \in \mathbb{R}^{\mathcal{A}}$ satisfies* (2.13) *but is not an X-circuit, then f is not extremal in $C_X(\mathcal{A}, \beta)$.*

Proof. If f is an X-AGE exponential with $c_\beta < 0$ and ν satisfies (2.13), then we must have $\nu \ne \mathbf{0}$ and $\sigma_X(-\mathcal{A}\nu) < \infty$. By the definition of an X-circuit, ν may be written as a convex combination $\nu = \theta\nu^{(1)} + (1-\theta)\nu^{(2)}$ where $\bar{\nu} \mapsto \sigma_X(-\mathcal{A}\bar{\nu})$ is linear on $[\nu^{(1)}, \nu^{(2)}]$ and $\nu^{(1)}$ and $\nu^{(2)}$ are non-proportional. We can therefore use Lemma 6.2.1 to prove the claim. □

We turn to eliminating the degree of freedom associated with ν lying on a ray. For each $\beta \in \mathcal{A}$, following the notation in Chapter 3, we introduce the notation

$$\Lambda_X(\mathcal{A}, \beta) = \{\lambda \in N_\beta \,:\, \lambda \text{ is an } X\text{-circuit of } \mathcal{A},\ \lambda_\beta = -1\}$$

for the associated set of normalized X-circuits of $\mathcal{A}$. The set of all normalized X-circuits of $\mathcal{A}$ is denoted $\Lambda_X(\mathcal{A})$. The main reason for introducing this notation is how it interacts with the following definition.

Definition 6.2.3. Given a vector $\lambda \in N_\beta$ with $\lambda_\beta = -1$, the λ*-witnessed* X*-AGE-cone* is

$$C_X(\mathcal{A}, \lambda) = \left\{ \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle} \,:\, \prod_{\alpha \in \lambda^+} \left(\frac{c_\alpha}{\lambda_\alpha} \right)^{\lambda_\alpha} \geq -c_\beta \exp\left(\sigma_X(-\mathcal{A}\lambda)\right),\ c_{\setminus\beta} \geq \mathbf{0} \right\}. \tag{6.9}$$

We show below that every exponential sum in $C_X(\mathcal{A}, \lambda)$ is nonnegative on X. The term *witnessed* in λ*-witnessed* X*-AGE-cone* is chosen to reflect the defining role of λ in the nonnegativity certificate.

Theorem 6.2.4. *The cone* $C_X(\mathcal{A}, \beta)$ *can be written as the convex hull of* λ*-witnessed* X*-AGE-cones where* λ *runs over the normalized* X*-circuits, together with the union of all atomic nonnegative exponentials supported on* $\mathcal{A}$*, that is,*

$$C_X(\mathcal{A}, \beta) = \operatorname{conv} \bigcup_{\lambda \in \Lambda_X(\mathcal{A}, \beta)} C_X(\mathcal{A}, \lambda) \cup \bigcup_{\alpha \in \mathcal{A}} \mathbb{R}_+ \cdot \{e^{\langle \alpha, x \rangle}\}.$$

The second union particularly captures the case that $\Lambda_X(\mathcal{A}, \beta) = \emptyset$.

Proof. Theorem 6.2.2 already tells us that $C_X(\mathcal{A}, \beta)$ may be expressed as the union of the set of monomial exponential expressions with positive coefficients and the convex hull of X-AGE exponentials $f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle}$ where (c, ν) satisfies (2.13) for some X-circuit ν. Therefore, it suffices to show that (i) for any such function, the normalized X-circuit $\lambda = \nu / (\mathbf{1}^T \nu_{\setminus\beta})$ is such that (c, λ) satisfies the condition in (6.9), and (ii) if *any* (c, λ) satisfies (6.9), then, the resulting exponential sum is nonnegative on X. We will actually do both of these in one step.

Suppose $\nu \in N_\beta$ is restricted to satisfy $\nu = s\lambda$ for a variable $s \geq 0$ and a fixed $\lambda \in \Lambda_X(\mathcal{A}, \beta)$. It suffices to show that the set of $c \in \mathbb{R}^{\mathcal{A}}$ for which

$$\exists s \geq 0 \,:\, \nu = s\lambda \text{ and } \sigma_X(-\mathcal{A}\nu) + D(\nu_{\setminus\beta}, ec_{\setminus\beta}) \leq c_\beta$$

is the same as (6.9).

Let $r(\nu) = \sigma_X(-\mathcal{A}\nu) + D(\nu_{\setminus\beta}, ec_{\setminus\beta})$. We apply positive homogeneity of the support function to see that $\sigma_X(-\mathcal{A}\nu) = |\nu_\beta| \sigma_X(\mathcal{A}\nu/|\nu_\beta|)$, and use $\nu = s\lambda$ to infer $s = |\nu_\beta|$ and $\sigma_X(-\mathcal{A}\nu/|\nu_\beta|) = \sigma_X(-\mathcal{A}\lambda)$. Then, we abbreviate $d := \sigma_X(-\mathcal{A}\lambda)$ and substitute $\sum_{\alpha \in \lambda^+} \nu_\alpha = |\nu_\beta|$ to obtain

$$r(\nu) = \textstyle\sum_{\alpha \in \lambda^+} \left(\nu_\alpha \log(\nu_\alpha / c_\alpha) - \nu_\alpha + \nu_\alpha d \right).$$

The term d may be moved into the logarithm by identifying $\nu_\alpha d = \nu_\alpha \log(1/e^{-d})$. For $\alpha \in \lambda^+$, we define scaled terms $\tilde{c}_\alpha = c_\alpha e^{-d}$ so that $r(\nu) = \sum_{\alpha \in \lambda^+} \nu_\alpha \log(\nu_\alpha / \tilde{c}_\alpha) - \nu_\alpha$. Following considerations in the appendix of [MNT20], there exists a $\nu = s\lambda$ for which $r(\nu) \leq c_\beta$ if and only if

$$-c_\beta \leq \prod_{\alpha \in \lambda^+} (\tilde{c}_\alpha / \lambda_\alpha)^{\lambda_\alpha}. \tag{6.10}$$

Since $(\tilde{c}_\alpha / \lambda_\alpha)^{\lambda_\alpha} = (c_\alpha / \lambda_\alpha)^{\lambda_\alpha} \left(e^{-d}\right)^{\lambda_\alpha}$ and $\prod_{\alpha \in \lambda^+} \left(e^{-d}\right)^{\lambda_\alpha} = e^{-d}$, (6.10) can be recognized as the inequality occurring within (6.9), which completes the proof. $\square$

Theorem 6.2.4 shows how λ-witnessed X-AGE-cones provide a window to the structure of *full* X-AGE-cones $C_X(\mathcal{A}, \beta)$. To appreciate the benefit of this perspective, it is necessary to consider the more elementary *power-cone*. In our context, the primal power-cone associated with a normalized X-circuit $\lambda \in \mathbb{R}^{\mathcal{A}}$ is

$$\text{Pow}(\lambda) = \{z \in \mathbb{R}^{\operatorname{supp}\lambda} : \textstyle\prod_{\alpha\in\lambda^+} z_\alpha^{\lambda_\alpha} \geq |z_\beta|,\ z_{\setminus\beta} \geq \mathbf{0},\ \beta := \lambda^-\};$$

the corresponding dual cone is given by

$$\text{Pow}(\lambda)^* = \{w \in \mathbb{R}^{\operatorname{supp}\lambda} : \textstyle\prod_{\alpha\in\lambda^+} (w_\alpha/\lambda_\alpha)^{\lambda_\alpha} \geq |w_\beta|,\ w_{\setminus\beta} \geq \mathbf{0},\ \beta := \lambda^-\}.$$

It should be evident that $C_X(\mathcal{A}, \lambda)$ can be formulated in terms of a dual λ-weighted power-cone; a precise formula is provided momentarily. For now, we give a corollary concerning power-cone representability and second-order representability of $C_X(\mathcal{A})$ when X is a polyhedron, see, e.g., [Ave19; BN01] for formal definitions.

Corollary 6.2.5. *If X is a polyhedron, then $C_X(\mathcal{A})$ is power-cone representable. If in addition $\mathcal{A}^T X$ is rational, then $C_X(\mathcal{A})$ is second-order representable and thus, has semidefinite extension degree 2.*

We provide a concrete description of the second-order-cone program in Chapter 7.

Proof. By Theorem 6.1.8, polyhedral X have finitely many X-circuits, up to multiples. Apply Theorem 6.2.2 and finiteness of the normalized circuits $\Lambda_X(\mathcal{A})$ to write

$$C_X(\mathcal{A}) = \sum_{\lambda\in\Lambda_X(\mathcal{A})} C_X(\mathcal{A}, \lambda) + \sum_{\alpha\in\mathcal{A}} \mathbb{R}_+ \cdot e^{\langle\alpha,x\rangle}.$$

The first claim follows as each of the finitely many sets $C_X(\mathcal{A}, \lambda)$ appearing in the above sum are (dual) power-cone representable. Elements in the sum $\sum_{\alpha\in\mathcal{A}} \mathbb{R}_+ \cdot e^{\langle\alpha,x\rangle}$ trivially have semidefinite extension degree one. For every other element, observe that under the rationality assumptions, we have $\Lambda_X(\mathcal{A}) \subseteq \mathbb{Q}^{\mathcal{A}}$. Using $\beta := \lambda^-$ and $m := |\operatorname{supp}\lambda|$, it is well-known that the m-dimensional λ-weighted power-cone as well as its dual are second-order representable when $\lambda_{\setminus\beta}$ is a rational vector in the $(m-1)$-dimensional probability simplex [BN01, Section 3.4]. The last claim follows as the semidefinite extension degree of the second-order-cone is 2 [BN01, Section 2.3]. ∎

The first part of Corollary 6.2.5 generalizes the case $X = \mathbb{R}^n$ considered by Papp for polynomials [Pap19]. That aspect of the corollary has uses in computational optimization when applied judiciously. The second part of Corollary 6.2.5 generalizes results by Averkov [Ave19] and Wang and Magron [WM20b] for ordinary SONC polynomials and the results on the $\mathcal{S}$-cone in Chapter 5.

We now work towards finding a simple representation of dual λ-witnessed X-AGE-cones $C_X(\mathcal{A}, \lambda)^*$. We begin this process by treating the primal as a cone of coefficients contained in $\mathbb{R}^{\mathcal{A}}$, and finding an explicit representation of the primal in terms of the elementary dual power-cone $\text{Pow}(\lambda)^*$.

Proposition 6.2.6. *For $\lambda \in N_\beta$ with $\lambda_\beta = -1$ and $\sigma_X(-\mathcal{A}\lambda) < \infty$, the λ-witnessed X-AGE-cone admits the representation*

$$\begin{aligned} C_X(\mathcal{A}, \lambda) = \{c \in \mathbb{R}^{\mathcal{A}} : \beta := \lambda^-,\ c_{\setminus\beta} \geq \mathbf{0},\ w \in \text{Pow}(\lambda)^*, \\ w_{|\lambda^+} = c_{|\lambda^+}, w_\beta = c_\beta e^{\sigma_X(-\mathcal{A}\lambda)} - r,\ r \geq 0\}. \end{aligned} \tag{6.11}$$

Proof. First, we note that some inequality constraints $c_{\setminus\beta} \geq \mathbf{0}$ are already implied by $w \in \mathrm{Pow}(\lambda)^*$. It is necessary to include the inequality constraints explicitly to account for the case when $\operatorname{supp}\lambda \subsetneq \mathcal{A}$. The condition $w \in \mathrm{Pow}(\lambda)^*$ can be rewritten as

$$\prod_{\alpha \in \lambda^+} (c_\alpha/\lambda_\alpha)^{\lambda_\alpha} \geq |c_\beta e^{\sigma_X(-\mathcal{A}\lambda)} - r|. \tag{6.12}$$

Meanwhile, the minimum of $|c_\beta e^{\sigma_X(-\mathcal{A}\lambda)} - r|$ over $r \geq 0$ is attained at $r = 0$ when $c_\beta < 0$, and at $r = c_\beta e^{\sigma_X(-\mathcal{A}\lambda)}$ when $c_\beta \geq 0$. In the $c_\beta < 0$ case, the constraint (6.12) becomes

$$\prod_{\alpha \in \lambda^+} (c_\alpha/\lambda_\alpha)^{\lambda_\alpha} \geq -c_\beta e^{\sigma_X(-\mathcal{A}\lambda)}.$$

In the $c_\beta = 0$ case, the constraint (6.12) is meaningless, since $\prod_{\alpha \in \lambda^+}(c_\alpha/\lambda_\alpha)^{\lambda_\alpha} \geq 0$ is implied by $c_{\setminus\beta} \geq \mathbf{0}$. As the constraint in the preceding display is similarly meaningless when $c_\beta > 0$, we see that it can be used instead of (6.12) without loss of generality. □

We can appeal to Proposition 6.2.6 to find a representation for $C_X(\mathcal{A},\lambda)^*$ which is analogous to Equation (6.9). Again, the dual is computed by treating the primal as a cone of *coefficients*.

Proposition 6.2.7. *For $\lambda \in N_\beta$ with $\lambda_\beta = -1$ and $\sigma_X(-\mathcal{A}\lambda) < \infty$, the dual λ-witnessed X-AGE-cone is given by*

$$C_X(\mathcal{A},\lambda)^* = \left\{ v \in \mathbb{R}^{\mathcal{A}}_+ \,:\, \beta := \lambda^-,\ e^{\sigma_X(-\mathcal{A}\lambda)} \prod_{\alpha \in \lambda^+} v_\alpha^{\lambda_\alpha} \geq v_\beta \right\}. \tag{6.13}$$

Proof. Let $\beta = \lambda^-$ as is usual. To $v \in \mathbb{R}^{\mathcal{A}}$ associate

$$\inf\{v^T c \,:\, c \in C_X(\mathcal{A},\lambda)\}. \tag{6.14}$$

A vector v belongs to $C_X(\mathcal{A},\lambda)^*$ if and only if $\inf\{v^T c \,:\, c \in C_X(\mathcal{A},\lambda)\} = 0$. We will find constraints on v so that the dual feasible set for (6.14) is non-empty, which in turn implies the claim.

We begin by noting that for any element $\alpha \in \mathcal{A} \setminus \operatorname{supp}\lambda$, the only constraints on c_α, v_α for $c \in C_X(\mathcal{A},\lambda), v \in C_X(\mathcal{A},\lambda)^*$ are $c_\alpha \geq 0, v_\alpha \geq 0$; therefore we assume $\mathcal{A} = \operatorname{supp}\lambda$ for the remainder of the proof. When considering the given expression for (6.14) as a primal problem, we compute a dual using (6.11) from Proposition 6.2.6. Under the assumption $\mathcal{A} = \operatorname{supp}\lambda$, the constraint $c_{\setminus\beta} \geq \mathbf{0}$ is implied by $w \in \mathrm{Pow}(\lambda)^*$ for $w_{|\lambda^+} = c_{|\lambda^+}, w_\beta = c_\beta e^{\sigma_X(-\mathcal{A}\lambda)} - r,\ r \geq 0$. Therefore, when forming a Lagrangian for (6.14) using (6.11), the dual variable to $c_{\setminus\beta} \geq \mathbf{0}$ may be omitted.

For the remaining constraints $w \in \mathrm{Pow}(\lambda)^*$ and $r \geq 0$, we use dual variables $\mu \in \mathrm{Pow}(\lambda)$ and $t \in \mathbb{R}_+$, respectively; the Lagrangian is

$$\begin{aligned}
\mathcal{L}(c,r,\mu,t) &= v^T c - \sum_{\alpha \in \lambda^+} \mu_\alpha c_\alpha - \mu_\beta(c_\beta e^{\sigma_X(-\mathcal{A}\lambda)} - r) - tr \\
&= \sum_{\alpha \in \lambda^+} c_\alpha(v_\alpha - \mu_\alpha) + c_\beta(\nu_\beta - \mu_\beta e^{\sigma_X(-\mathcal{A}\lambda)}) - r(t - \mu_\beta).
\end{aligned}$$

Since we have assumed $\operatorname{supp}\lambda = \mathcal{A}$ and $\sigma_X(-\mathcal{A}\lambda) < \infty$, for the Lagrangian to be bounded below over $c \in \mathbb{R}^{\mathcal{A}}$ and $r \in \mathbb{R}$, it is necessary and sufficient that $v_\alpha = \mu_\alpha$

for all $\alpha \in \lambda^+, \nu_\beta e^{-\sigma_X(-\mathcal{A}\lambda)} = \mu_\beta$ and $\mu_\beta = t$. Hence,

$$
\begin{aligned}
&\inf\{v^T c \,:\, c \in C_X(\mathcal{A}, \lambda)\} \\
=&\inf\left\{ \sup\{\mathcal{L}(c, r, \mu, t) \,:\, (\mu, t) \in \mathrm{Pow}(\lambda) \times \mathbb{R}_+\} \,:\, (c, r) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R} \right\} \\
=&\sup\left\{ \inf\{\mathcal{L}(c, r, \mu, t) \,:\, (c, r) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}\} \,:\, (\mu, t) \in \mathrm{Pow}(\lambda) \times \mathbb{R}_+ \right\} \\
=&0.
\end{aligned}
$$

The proposition follows by applying the definition of $\mathrm{Pow}(\lambda)$. □

6.2.1 Excursus: The Conditional $\mathcal{S}$-Cone

Nonnegativity certificates of a given exponential sum f on a convex set $X \subseteq \mathbb{R}^n$ were originally studied and introduced by Murray, Chandrasekaran, and Wierman (not only for exponential sums but also for polynomials). Even though we study the setting of exponential sums in this chapter, there exist similar X-nonnegativity certificates for elements in the $\mathcal{S}$-cone, i.e., generalized polynomial and exponential functions of the form $f = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta$.

Let $X \subseteq \mathbb{R}^n$ be convex and closed. We consider the set $C_{\mathcal{S}}(X, \mathcal{A}, \mathcal{B})$ of functions of the form (5.1) supported on $\mathcal{A}$ and $\mathcal{B}$ that are nonnegative on X. In contrast to the situation studied in Chapter 5, in the constrained case, observe that even if the term $d_\beta x^\beta$ for every element $\beta \in \mathbb{N}^n \setminus (2\mathbb{N})^n$ is *never globally* nonnegative, there might exist some convex and closed $X \subseteq \mathbb{R}^n$ such that $d_\beta x^\beta$ is nonnegative on X.

Let $\mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$, and $f = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta$ with $c_\alpha, d_\beta \in \mathbb{R}$ for all $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$.

Theorem 6.2.8. *Let $X \subseteq \mathbb{R}^n$ be closed and convex and suppose there exists exactly one nonpositive term, i.e., either*

1. *there exists a designated $\beta' \in \mathcal{B}$ such that $d_{\beta'} x^{\beta'} < 0$ for some $x \in X$, $d_\beta x^\beta \geq 0$ for all $\beta \in \mathcal{B} \setminus \{\beta'\}, x \in X$ and $c_\alpha \geq 0$ for all $\alpha \in \mathcal{A}$ or*
2. *there exists a designated $\alpha' \in \mathcal{A}$ such that $c_{\alpha'} < 0, c_\alpha \geq 0$ for all $\alpha \in \mathcal{A} \setminus \{\alpha'\}$ and $d_\beta x^\beta \geq 0$ for all $\beta \in \mathcal{B}$ and $x \in X$.*

Then, f is nonnegative if and only if

1. *for $X_{\beta'} = X \cap \{x \in \mathbb{R}^n : d_{\beta'} x^{\beta'} < 0\}$, there exists $\nu \in \mathbb{R}_+^{\mathcal{A}, \mathcal{B} \setminus \{\beta'\}}$ such that*
$$\sigma_{X_{\beta'}}\left(\sum_{\alpha \in \mathcal{A} \cup (\mathcal{B} \setminus \{\beta\})} \nu_\alpha(\beta' - \alpha)) + D(\nu_{|\mathcal{A}}, ec) + D(\nu_{|\mathcal{B}}, ed \right) \leq -|d_\beta|$$
for case (1), and
2. *for $X_{\alpha'} = X \cap \{x \in \mathbb{R}^n : c_{\alpha'} < 0\}$, there exists $\nu \in \mathbb{R}_+^{\mathcal{A} \setminus \{\alpha'\}, \mathcal{B}}$ such that*
$$\sigma_{X_{\alpha'}}\left(\sum_{\alpha \in (\mathcal{A} \setminus \{\alpha'\}) \cup \mathcal{B}} \nu_\alpha(\beta - \alpha)) + D(\nu_{|\mathcal{A}}, ec) + D(\nu_{|\mathcal{B}}, ed \right) \leq c_\beta,$$
for case (2).

Proof. We only prove the first statement. The second one is an immediate consequence of the statement on exponential sums by Murray, Chandrasekaran, and Wierman [MCW21b].

Let $X_{\beta'}$ be defined as declared above and let

$$\delta_{X_{\beta'}}(x) = \begin{cases} 0 & \text{if } x \in X_{\beta'}, \\ -\infty & \text{otherwise} \end{cases}.$$

Following the proof of Murray, Chandrasekaran, and Wierman for conditional SAGE certificates, f is nonnegative if and only if

$$p^* = \inf\{\delta_{X_\beta}(x) + \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B} \setminus \{\beta'\}} d_\beta x^\beta + d_{\beta'} x^{\beta'}, x \in \mathbb{R}^n\} \geq 0.$$

With the definition of $\delta_{X_{\beta'}}$, we assume, without loss of generality, $d_{\beta'} x^{\beta'} = -|d_{\beta'}||x^{\beta'}|$ and $f = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B} \setminus \{\beta'\}} |d_\beta||x|^\beta - |d_{\beta'}||x|^{\beta'}$ because we have $d_\beta x^\beta \geq 0$ for all $\beta \in \mathcal{B} \setminus \{\beta'\}$.

Hence, we can equivalently compute

$$\begin{aligned} f^* &= \inf\{\delta_{X_{\beta'}}(x) + \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^{\alpha - \beta'} + \sum_{\beta \in \mathcal{B} \setminus \{\beta'\}} d_\beta |x|^{\beta - \beta'}, x \in \mathbb{R}^n\} \\ &= \inf\{\delta_{X_{\beta'}}(x) + \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle x, (\alpha - \beta')\rangle} + \sum_{\beta \in \mathcal{B} \setminus \{\beta'\}} d_\beta e^{\langle x, \beta - \beta'\rangle)}, x \in \mathbb{R}^n\} \geq |d_{\beta'}|. \end{aligned}$$

Following [MCW21b], this is equivalent to

$$\sup_{v \in \mathbb{R}_+^{\mathcal{A}}} \sigma_{X_{\beta'}} \left(\sum_{\alpha \in \mathcal{A} \cup (\mathcal{B} \setminus \{\beta'\})} \nu_\alpha (\alpha - \beta')) + D(\nu_{|\mathcal{A}}, ec) + D(\nu_{|\mathcal{B}}, ed \right) \leq -|c_{\beta'}|,$$

Hence, f is X-nonnegative if and only if there exists $\nu \in \mathbb{R}_+^{\mathcal{A}, \mathcal{B} \setminus \{\beta'\}}$ such that

$$\sigma_{X_{\beta'}} \left(\sum_{\alpha \in \mathcal{A} \cup (\mathcal{B} \setminus \{\beta'\})} \nu_\alpha (\alpha - \beta')) + D(\nu_{|\mathcal{A}}, ec) + D(\nu_{|\mathcal{B}}, ed \right) \leq -|c_{\beta'}|,$$

□

Of course, having $\mathcal{B} = \emptyset$, this should match the case of X-AGE exponentials, and it indeed does because for $c_{\alpha'} < 0$ for some $\alpha' \in \mathcal{A}$, we naturally have $c_{\alpha'} = -|c_{\alpha'}|$.

Note moreover that in the constrained case, in general we do not have (as we have in the unconstrained case) that a polynomial is X-nonnegative if and only if the corresponding exponential sum is X-nonnegative, because $-X \subseteq X$ does not necessarily hold in every case.

6.2.2 X-Circuits and Their Supports

In this subsection, we study the relationship between X-circuits and their supports.

As stated in the introduction, in the classical case of affine circuits, the normalized circuits are uniquely determined by their supports. Moreover, as a consequence of Theorem 6.1.9, in the case $X = [-1, 1]$, the normalized X-circuits are uniquely determined by their signed supports. As explained in the following, this phenomenon does not extend to sublinear circuits for arbitrary sets. In the special case of sublinear circuits supported on two elements, the two nonzero entries are additive inverses of each other, so that, for a given β and a given support, this signed support indeed uniquely determines the circuit up to a positive factor. In order to exhibit the mentioned phenomenon, we present a counterexample with support size 3.

Example 6.2.9. Let $\mathcal{A} = \{\alpha_1, \alpha_2, \alpha_3\} = \left\{(0,0)^T, (1,0)^T, (0,1)^T\right\} \subseteq \mathbb{R}^2$. We show that for $\beta := \alpha_1$, there are two non-proportional circuits which are supported on all three elements of $\mathcal{A}$. Specifically, we construct an example in which

$$\nu^{(1)} := (-2,1,1)^T \quad \text{and} \quad \nu^{(2)} := (-3,1,2)^T$$

are sublinear circuits. Note that both of them have the same signed support but they are not multiples of each other. Observe that

$$-\mathcal{A}\nu^{(1)} = (-1,-1)^T, \qquad -\mathcal{A}\nu^{(2)} = (-1,-2)^T.$$

We set up X in such a way that $(-1,-1)^T$ and $(-1,-2)^T$ are normal vectors of X. For example, choose X as the cone in $\mathbb{R}^2$ spanned by $(-1,1)^T$ and $(2,-1)^T$. We obtain

$$\begin{aligned} -\mathcal{A}^T X &= \operatorname{pos}\left\{ \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\} \\ &= \operatorname{pos}\left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

Since $N_\beta^\circ = N_{(0,0)}^\circ = \mathbb{R} \cdot (1,1,1)^T + \mathbb{R} \times \mathbb{R}_{\le 0} \times \mathbb{R}_{\le 0}$, it can be verified (for example, using a computer calculation) that $\nu^{(1)}$ and $\nu^{(2)}$ are indeed sublinear circuits, and that they are the only ones having a negative component ν_β up to scaling by a positive factor.

In the example, the two distinct sublinear circuits $\nu^{(i)}$, $1 \le i \le 2$, with identical signed supports, have different expressions $\mathcal{A}\nu^{(i)}$, that is, $\mathcal{A}\nu^{(1)} \neq \mathcal{A}\nu^{(2)}$. By the following statement, it is not possible to have two distinct sublinear circuits with the same signed support and identical nonzero values of $\mathcal{A}\nu^{(i)}$.

Lemma 6.2.10. *Let $\nu^{(1)}$ and $\nu^{(2)}$ be sublinear circuits with the same signed support and such that $\mathcal{A}\nu^{(1)} = \mathcal{A}\nu^{(2)}$. Then, $\nu^{(1)}$ and $\nu^{(2)}$ are proportional, and in case $\mathcal{A}\nu^{(1)} = \mathcal{A}\nu^{(2)} \neq 0$, the equality $\nu^{(1)} = \nu^{(2)}$ holds.*

Proof. Let $\nu^{(1)}$ and $\nu^{(2)}$ have the same signed support with $\mathcal{A}\nu^{(1)} = \mathcal{A}\nu^{(2)}$. Set β as the index of the negative component of $\nu^{(1)}$ and $\nu^{(2)}$. Assuming $\nu^{(1)} \neq \nu^{(2)}$, the precondition $\operatorname{supp} \nu^{(1)} = \operatorname{supp} \nu^{(2)}$ implies that for sufficiently small $\varepsilon > 0$, the vectors

$$\nu' := \nu^{(1)} - \varepsilon\nu^{(2)} \quad \text{and } \nu'' := \nu^{(1)} + \varepsilon\nu^{(2)}$$

are both contained in $N_\beta \setminus \{\mathbf{0}\}$ as well. Observe that

$$\begin{aligned} &\sigma_X(-\mathcal{A}\nu') = \sigma_X(-\mathcal{A}\nu^{(1)}) - \varepsilon\sigma_X(-\mathcal{A}\nu^{(2)}) < \infty \text{ and} \\ &\sigma_X(-\mathcal{A}\nu'') = \sigma_X(-\mathcal{A}\nu^{(1)}) + \varepsilon\sigma_X(-\mathcal{A}\nu^{(2)}) < \infty. \end{aligned}$$

Moreover, $\nu^{(1)}$ is a convex combination $\nu^{(1)} = \frac{1}{2}\nu' + \frac{1}{2}\nu''$ for which $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ is linear on $[\nu', \nu'']$. Since $\mathcal{A}\nu^{(1)} = \mathcal{A}\nu^{(2)}$, the vectors $\nu^{(1)}$ and $\nu^{(2)}$ are non-proportional or we have $\mathcal{A}\nu^{(1)} = \mathcal{A}\nu^{(2)} = \mathbf{0}$. In both cases, if $\nu^{(1)}$ and $\nu^{(2)}$ are non-proportional, then this contradicts that $\nu^{(1)}$ is a sublinear circuit. $\square$

6.2.3 Necessary and Sufficient Conditions

In this subsection, we obtain some criteria for elements $\nu \in \bigcup_{\beta\in\mathcal{A}} N_\beta$ to be X-circuits of some fixed set X. These criteria only involve the supports rather than the exact values of the coefficients. Recall that for an X-circuit ν, we denote by ν^- the single index β with $\nu_\beta < 0$ and $\nu^+ := \{\alpha : \nu_\alpha \geq 0\}$. Moreover, recall that in the classical case of affine matroids, any simplicial circuit ν supported on at least three elements has no other support point except ν^- contained in the relative interior of the convex hull of all its support points, and the coefficients of ν^+ are positive multiples of the barycentric coordinates of β, i.e., $\operatorname{relint}\operatorname{conv}(\operatorname{supp}(\nu)) \cap \nu^+ = \emptyset$ and $\mathcal{A}\nu = \mathbf{0}$ (compare Chapter 3 or see, e.g., [FW19]). In the following theorem, we generalize this property to the case of X-circuits.

Theorem 6.2.11. *If $\lambda \in \Lambda_X(\mathcal{A}, \beta)$ for $\beta \in \mathcal{A}$, then $\operatorname{relint}\operatorname{conv}(\operatorname{supp}(\lambda)) \cap \lambda^+ = \emptyset$. Moreover, if $\beta \in \operatorname{conv}(\lambda^+)$, then $\mathcal{A}\lambda = \mathbf{0}$.*

Proof. Fix $\lambda \in \Lambda_X(\mathcal{A}, \beta)$. For the first statement, suppose there exists $\bar{\alpha} \in \lambda^+$ such that $\bar{\alpha} \in \operatorname{relint}\operatorname{conv}(\operatorname{supp}(\lambda))$. Hence, there exist $\theta_\alpha \in [0, 1)$ for $\alpha \in (\lambda^+ \setminus \{\bar{\alpha}\}) \cup \{\beta\}$ such that

$$\sum_{\alpha\in\lambda^+\setminus\{\bar{\alpha}\}} \theta_\alpha + \theta_\beta = 1 \quad \text{and} \quad \sum_{\alpha\in\lambda^+\setminus\{\bar{\alpha}\}} \theta_\alpha \alpha + \theta_\beta \beta = \bar{\alpha}.$$

Let $\tau \in (0, 1]$ be maximal such that $\tau\theta_\alpha\lambda_{\bar{\alpha}} \leq \lambda_\alpha$ for $\alpha \in (\lambda^+ \setminus \{\bar{\alpha}\}) \cup \{\beta\}$ and $(1 + \tau)\lambda_{\bar{\alpha}} < 1$. As $\lambda_{\bar{\alpha}} < 1$, this does indeed exist. The two vectors $\nu^{(1)}$ and $\nu^{(2)}$ defined by

$$\nu^{(1)}_\alpha = \begin{cases} \lambda_\alpha + \tau\theta_\alpha\lambda_{\bar{\alpha}} & \text{for } \alpha \in (\lambda^+ \setminus \{\bar{\alpha}\}) \cup \{\beta\}, \\ (1-\tau)\lambda_{\bar{\alpha}} & \text{for } \alpha = \bar{\alpha} \end{cases}$$
$$\text{and} \quad \nu^{(2)}_\alpha = \begin{cases} \lambda_\alpha - \tau\theta_\alpha\lambda_{\bar{\alpha}} & \text{for } \alpha \in (\lambda^+ \setminus \{\bar{\alpha}\}) \cup \{\beta\}, \\ (1+\tau)\lambda_{\bar{\alpha}} & \text{for } \alpha = \bar{\alpha} \end{cases}$$

(and 0 outside of $\lambda^+ \cup \{\beta\}$) are non-proportional elements of N_β with $(\nu^{(i)})^+ \subseteq \lambda^+$ for $i = 1, 2$. Moreover, $\mathcal{A}\nu^{(i)} = \mathcal{A}\lambda$ for $i = 1, 2$ and $\lambda \in \operatorname{relint}[\nu^{(1)}, \nu^{(2)}]$, which contradicts the X-circuit property of λ. For the second statement, suppose that $\beta \in \operatorname{conv}(\lambda^+)$ and $\mathcal{A}\lambda \neq \mathbf{0}$. Then, there exists a normalized $\lambda' \in N_\beta$ with $\lambda^+ = (\lambda')^+$ and $\mathcal{A}\lambda' = \mathbf{0}$. Let τ be the maximal real number with $\nu^{(1)} := \lambda - \tau\lambda' \in N_\beta$. That maximum clearly exists, and, since $(\lambda')^+ = \lambda^+$, the number τ is positive. Moreover, since λ and λ' are normalized, we have $\tau \leq 1$. The sublinear circuit $\nu^{(2)} := \lambda + \tau\lambda'$ is also clearly contained in N_β. Since λ, λ' are non-proportional and $\tau > 0$, the sublinear circuits $\nu^{(1)}$ and $\nu^{(2)}$ are non-proportional. Further, since $\nu^{(1)} + \nu^{(2)} = 2\lambda$, we see that λ can be written as a convex combination of the two non-proportional elements $\nu^{(1)} \in N_\beta$ and $\nu^{(2)} \in N_\beta$. Due to $\mathcal{A}\lambda' = \mathbf{0}$, we obtain $\sigma_X(-\mathcal{A}\nu^{(1)}) = \sigma_X(-\mathcal{A}\nu^{(2)}) = \sigma_X(-\mathcal{A}\lambda)$, and thus

$$\sigma_X(-\mathcal{A}\lambda) = \frac{1}{2}(\sigma_X(-\mathcal{A}\nu^{(1)}) + \sigma_X(-\mathcal{A}\nu^{(2)})).$$

Hence, $\lambda \notin \Lambda_X(\mathcal{A}, \beta)$. $\square$

We can provide the following two cases of the converse of Theorem 6.2.11. In particular, both cases will be applicable for $X = [-1,1]^n$. We can assume that $\beta \in \operatorname{conv}(\lambda^+) - \operatorname{rec}(X)^*$ since otherwise any $\lambda \in N_\beta \setminus \{\mathbf{0}\}$ will have $\sigma_X(-\mathcal{A}\lambda) = \infty$ and hence, violate condition (1) in the definition of an X-circuit.

Lemma 6.2.12. *Given $\beta \in \mathcal{A}$, let $\lambda \in N_\beta \setminus \{\mathbf{0}\}$ be a normalized element such that λ^+ consists of affinely independent vectors and with $\beta \in \operatorname{conv}(\lambda^+) - \operatorname{rec}(X)^*$.*

1. *If $|\operatorname{supp}(\lambda)| = 2$ or*
2. *if X is full-dimensional, $\beta \in \operatorname{conv}(\lambda^+)$, $\mathcal{A}\lambda = \mathbf{0}$,*

then $\lambda \in \Lambda_X(\mathcal{A}, \beta)$.

Note that, since in the previous theorem, λ^+ consists of affinely independent vectors, we have $\operatorname{relint}\operatorname{conv}(\lambda^+) \cap \lambda^+ = \emptyset$.

Remark 6.2.13. If the property of full-dimensionality is omitted in the second condition, the statement is not true anymore. As a counterexample, let X be the singleton set $X = \{1\}$ and let $\mathcal{A} = \{1,2,3\}$. Then $\lambda = \frac{1}{2}(1,-2,1)^T$ is not an X-circuit because $\lambda = \frac{1}{2}\lambda^{(1)} + \frac{1}{2}\lambda^{(2)}$ with $\lambda^{(1)} = (1,-1,0)^T$ and $\lambda^{(2)} = (0,-1,1)^T$ and $\nu \to \sigma_X(-\mathcal{A}\nu)$ is linear on $[\lambda^{(1)}, \lambda^{(2)}]$. Note that the functions $x \mapsto e^{\langle x, \alpha \rangle}$, $\alpha \in \mathcal{A}$ are not linearly independent on X.

Proof of Lemma 6.2.12. For the first statement, suppose there exist $\nu^{(1)}, \nu^{(2)} \in N_\beta$ decomposing λ. Then, $\operatorname{supp}(\nu^{(i)}) \subseteq \operatorname{supp}(\lambda)$ for $i \in \{1,2\}$ because the cancellation of terms not contained in $\operatorname{supp}(\lambda)$ is not possible, as the negative term always corresponds to β. Since $\nu_\beta^{(1)} < 0$ and $\nu_\beta^{(2)} < 0$ and $|\operatorname{supp}(\lambda)| = 2$, both $\nu^{(1)}$ and $\nu^{(2)}$ are proportional to λ.

Now, consider the second condition. Since the property of being an X-circuit is invariant under translation of X, we can assume, without loss of generality, that $\mathbf{0} \in \operatorname{int} X$. Suppose that there exist non-proportional, normalized $\lambda^{(1)}, \lambda^{(2)} \in N_\beta$ and $\theta_1, \theta_2 \in (0,1)$ with $\theta_1 + \theta_2 = 1$ such that

$$\sum_{i=1}^{2} \theta_i (\lambda^{(i)}, \sigma_X(-\mathcal{A}\lambda^{(i)})) = (\lambda, \sigma_X(-\mathcal{A}\lambda)).$$

We distinguish two cases. If $\mathcal{A}\lambda^{(1)} = \mathbf{0}$, then $\mathcal{A}\lambda^{(2)} = -\frac{\theta_1}{\theta_2}\mathcal{A}\lambda^{(1)} = \mathbf{0}$. Hence, the uniqueness of the barycentric coordinates with respect to a given affinely independent ground set implies $\lambda^{(1)} = \lambda^{(2)}$, which is a contradiction to their non-proportionality.

By the argument above $\mathcal{A}\lambda^{(2)} = \mathbf{0}$ implies $\mathcal{A}\lambda^{(1)} = \mathbf{0}$. Hence, if $\mathcal{A}\lambda^{(1)} \neq \mathbf{0}$, we have $\mathcal{A}\lambda^{(2)} = -\frac{\theta_1}{\theta_2}\mathcal{A}\lambda^{(1)} \neq \mathbf{0}$ as well. Then $\mathbf{0} \in \operatorname{int} X$ implies $\sigma_X(-\mathcal{A}\lambda^{(1)}) > 0$ and $\sigma_X(-\mathcal{A}\lambda^{(2)}) > 0$. Since $\sigma_X(-\mathcal{A}\lambda) = -\sigma_X(\mathbf{0}) = 0$, the mapping $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ cannot be linear on $[\lambda^{(1)}, \lambda^{(2)}]$. □

X-Circuits of Polyhedral Cones X

As discussed after Theorem 6.1.8, in the case of polyhedral cones X, we have that $\sigma_X(-\mathcal{A}\lambda) = 0$ whenever this value is finite. Since we will reduce the determination of the sublinear circuits $\Lambda_X(\mathcal{A})$ for a cone X in some prominent cases to the classical affine circuits $\Lambda_{\mathbb{R}^n}(\mathcal{A})$ (which of course is also a case of a polyhedral cone), we first look at an example for the latter case. In the following, we examine sublinear circuits for various sets $X \subseteq \mathbb{R}^n$ (for some $n \in \mathbb{N}$) and support sets, which have the form

$\mathcal{A} = \{(i,j) \ : \ 1 \leq i,j \leq k\}$, $k \in \mathbb{N}$. In these situations, we can write a sublinear circuit ν as a matrix $M^{(\nu)} \in \mathbb{R}^{k \times k}$ such that $M^{(\nu)}_{i,j} = \nu_{(i,j)}$ for all $(i,j) \in \mathcal{A}$.

Example 6.2.14. For $X = \mathbb{R}^2$ and support $\mathcal{A} = \{(i,j) \ : \ 1 \leq i,j \leq 3\}$, there are 16 sublinear circuits (up to multiples). Namely, there are 8 sublinear circuits with support size 3 (all of them have nonzero entries 1,-2,1; they appear in the three rows, the three columns and the two diagonals of the 3×3-matrix). Moreover, there are the following 8 sublinear circuits of support size 4. Here, the upper left entry of the matrices refers to the support point $(1,1)$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 0 & 2 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{6.15}$$

as well as the 90-degree, 180-degree and 270-degree rotations about the $(2,2)$-element of these matrices. As $\mathbf{0} \in \operatorname{int} \mathbb{R}^2$ and $\operatorname{rec}(\mathbb{R}^2)^* = \{\mathbf{0}\}$, in particular this reflects the statements of Theorem 6.2.11 and Lemma 6.2.12.

Next, we consider the sublinear circuits of the nonnegative orthant $\mathbb{R}^n_+$. For a non-empty subset $S \subseteq [n]$ and a support point $\alpha \in \mathcal{A} \subseteq \mathbb{R}^n$, we write α_S for the projection of α onto the components of S, i.e., $\alpha_S := (\alpha_s)_{s \in S}$. We also set $\mathcal{A}_S := \{\alpha_S \ : \ \alpha \in \mathcal{A}\}$ and for a matrix M with n rows, we set M_S as the submatrix of M defined by the rows with indices in S, which, in particular, yields $M_S \lambda = (M\lambda)_S$.

Theorem 6.2.15. *Let $X = \mathbb{R}^n_+$ and $\beta \in \mathcal{A}$. A normalized element $\lambda \in N_\beta$ is contained in $\Lambda_X(\mathcal{A}, \beta)$ if and only if there exists a non-empty subset $S \subseteq [n]$ such that $|\{\alpha_S \ : \ \alpha \in \operatorname{supp} \lambda\}| = |\operatorname{supp}(\lambda)|$ and where λ is an $\mathbb{R}^{|S|}$-circuit for the support set $\mathcal{A}_S$ and $(\mathcal{A}\lambda)_{[n] \setminus S} > \mathbf{0}$.*

Remark 6.2.16. The latter condition in Theorem 6.2.15 implies $\beta_S = (\mathcal{A}\lambda^+)_S$ and, hence, $\beta_S \in \operatorname{relint} \operatorname{conv}((\lambda^+)_S)$ and $\beta_{[n]\setminus S} \in \operatorname{conv}((\lambda^+)_{[n]\setminus S}) - \mathbb{R}^{[n]\setminus S}_+$.

Proof of Theorem 6.2.15. Let $\lambda \in \Lambda_X(\mathcal{A}, \beta)$. Hence, $\mathcal{A}\lambda \geq \mathbf{0}$. For every $s \in [n]$ with $(\mathcal{A}\lambda)_{\{s\}} > 0$, we observe that λ is also an $\mathbb{R}^{n-1}_+$-circuit for $\mathcal{A}_{n \setminus \{s\}}$. Similarly, the X-circuit property of λ necessarily implies that there exists at least one $s \in [n]$ with $(\mathcal{A}\lambda)_{\{s\}} = 0$. Let S be the inclusion-maximal subset $S \subseteq [n]$ with $(\mathcal{A}\lambda)_S = \mathbf{0}$. By the initial considerations, $S \neq \emptyset$ and λ is an $\mathbb{R}^{|S|}$-circuit of $\mathcal{A}_S$. This implies the cardinality statement $|\{\alpha_S \ : \ \alpha \in \operatorname{supp} \lambda\}| = |\operatorname{supp}(\lambda)|$. By definition of S, we have $(\mathcal{A}\lambda)_{[n]\setminus S} > \mathbf{0}$. Conversely, let $S \subseteq [n]$ with $|\{\alpha_S \ : \ \alpha \in \operatorname{supp} \lambda\}| = |\operatorname{supp}(\lambda)|$ such that λ is an $\mathbb{R}^{|S|}$-circuit of $\mathcal{A}_S$ and $(\mathcal{A}\lambda)_{[n]\setminus S} > \mathbf{0}$. Then λ is an $\mathbb{R}^{|S|}_+$-circuit for $\mathcal{A}_S$ and, further, an X-circuit for $\mathcal{A}$. □

Theorem 6.2.15 can be used in the reduction of the enumeration of all X-circuits to the enumeration of all classical affine circuits.

Example 6.2.17. For $X = \mathbb{R}^2_+$ and the support set $\mathcal{A} = \{(i,j) \ : \ 1 \leq i,j \leq 3\}$, there are 65 normalized sublinear circuits. Namely, by Theorem 6.2.15, there are

1. 27 normalized sublinear circuits of cardinality 2, namely, $\lambda = -\delta^{(i_1,j_1)} + \delta^{(i_2,j_2)}$ for $1 \leq i_1 \leq i_2 \leq 3$, $1 \leq j_1 \leq j_2 \leq 3$; that is, the entry 1 appears in the *lower right* quadrant of the entry -1.
2. 16 normalized sublinear circuits in which the entries $\frac{1}{2}, -1, \frac{1}{2}$ appear in columns 1,2, and 3, respectively, such that the entry -1 appears above the line through the two entries $\frac{1}{2}$.

3. 16 normalized sublinear circuits in which the entries $\frac{1}{2}$ appear in rows 1,2, and 3, respectively, such that the -1 appears to the left of the line containing the two entries $\frac{1}{2}$.

4. 8 $\mathbb{R}^2$-circuits of cardinality 4, which are the normalized versions of the ones from Example 6.2.14.

Since the diagonal and the anti-diagonal are counted both in cases (2) and (3), we have to subtract 2, which gives $27 + 16 + 16 + 8 - 2 = 65$. The following table shows the number of sublinear circuits with $\nu^- = \{(i, j)\}$ in row i and column j .

	1	2	3
1	8	14	2
2	14	21	2
3	2	2	0

Exemplarily, for the case $\nu^- = \{(2, 1)\}$, there are five circuits of type (1) as well as the following nine (in the subsequent list not normalized) sublinear circuits ν with $\nu^- = \{(2, 1)\}$, i.e., the component with index $(1, 2)$ is the negative component. As before, the upper left entry of the matrices refers to the support point $(1, 1)$:

$$\begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

The following theorem characterizes the connection between the X-circuits and the $\mathbb{R}^n$-circuits for more general polyhedral cones X.

Theorem 6.2.18. *Let $X = \operatorname{pos}\{v^{(1)}, \ldots, v^{(k)}\}$ be an n-dimensional polyhedral cone spanned by the vectors $v^{(1)}, \ldots, v^{(k)}$ where $k \geq n$. Then,*

$$\{\lambda \in \Lambda_X(\mathcal{A}) \ : \ \mathcal{A}\lambda = \mathbf{0}\} \ = \ \Lambda_{\mathbb{R}^n}(\mathcal{A}). \tag{6.16}$$

In the definition of $\Lambda_{\mathbb{R}^n}(\mathcal{A})$ we use in this chapter, there are no elements in $\Lambda_{\mathbb{R}^n}(\mathcal{A})$ which describe circuits supported on a single element as we used them in Chapters 3, 4 and 5.

Proof. Fix $\beta \in \mathcal{A}$ and denote by W the $k \times n$-matrix whose rows are the transposed vectors $(v^{(1)})^T, \ldots, (v^{(k)})^T$. Hence, $X^* = \{x \in \mathbb{R}^n \ : \ Wx \geq \mathbf{0}\}$. The set $\Lambda_X(\mathcal{A}, \beta)$ is the set of normalized vectors spanning the extreme rays of the cone

$$\begin{aligned} K_X \ &= \ \{\nu \in N_\beta \ : \ \sigma_X(-\mathcal{A}\nu) \leq \mathbf{0}\} \ = \ \{\nu \in N_\beta \ : \ \mathcal{A}\nu \in X^*\} \\ &= \ \{\nu \in N_\beta : W\mathcal{A}\nu \geq \mathbf{0}\} \end{aligned}$$

and the set $\Lambda^n_{\mathbb{R}}(\mathcal{A}, \beta)$ is the set of normalized vectors spanning the extreme rays of the cone

$$K_{\mathbb{R}^n} \ = \ \{\nu \in N_\beta \ : \ \sigma_{\mathbb{R}^n}(-\mathcal{A}\nu) \leq \mathbf{0}\} \ = \ \{\nu \in N_\beta \ : \ \mathcal{A}\nu = \mathbf{0}\}.$$

Since the matrix W has rank n, the linear mapping $x \mapsto Wx$ is injective, and thus, its kernel is $\{\mathbf{0}\}$. Hence, $K_{\mathbb{R}^n} = \{\nu \in N_\beta \ : \ W\mathcal{A}\nu = \mathbf{0}\}$. The cone $K_{\mathbb{R}^n}$ is contained in K_X. Consequently, if $\lambda \in N_\beta$ is not contained in the right hand side of (6.16), it is not contained in the left hand side. Conversely, let $\lambda \in N_\beta$ be contained in the right hand side of (6.16). Then, $\mathcal{A}\lambda = \mathbf{0}$ and $W\mathcal{A}\lambda = \mathbf{0}$. Assume there exists a decomposition into a convex combination $\lambda = \theta_1\lambda^{(1)} + \theta_2\lambda^{(2)}$ with $W\mathcal{A}\lambda^{(1)} \neq \mathbf{0}$. At least one component of $W\mathcal{A}(\lambda - \theta_1\lambda^{(1)}) = W\mathcal{A}\theta_2\lambda^{(2)}$ is smaller than zero since $W\mathcal{A}\lambda = \mathbf{0}$ as well as $W\mathcal{A}\lambda^{(1)} \geq \mathbf{0}$ and nonzero. This is a contradiction. Hence, λ is contained in the left hand side of (6.16). $\square$

6.2.4 The n-Dimensional Cube $X = [-1, 1]^n$

We discuss the sublinear circuits of the n-dimensional cube $[-1, 1]^n$, which is a prominent case of a compact polyhedron. Throughout this subsection, we assume $X = [-1, 1]^n$ for some fixed $n \in \mathbb{N}$ and $\mathcal{A} \subseteq \mathbb{R}^n$ non-empty and finite. We can already apply some of the former statements to gain knowledge about the structure of X-circuits. For example, as $\operatorname{rec}(X)^* = \mathbb{R}^n = -\operatorname{rec}(X)^*$, Lemma 6.2.12 implies that every element supported on exactly two points is an X-circuit. Hence, we examine the structure of those X-circuits $\lambda \in \Lambda_X(\mathcal{A})$ that have more than two support points. We begin with a necessary criterion.

Lemma 6.2.19. *Let $\lambda \in N_\beta$ with $\lambda_\beta = -1$ for some $\beta \in \mathcal{A}$ and $|\operatorname{supp}(\lambda)| \geq 3$. If for all $j \in [n]$*

$$\left(\alpha_j \leq \beta_j \text{ for all } \alpha \in \lambda^+\right) \quad \text{or} \quad \left(\alpha_j \geq \beta_j \text{ for all } \alpha \in \lambda^+\right), \tag{6.17}$$

then $\lambda \notin \Lambda_X(\mathcal{A})$.

Note that the precondition expresses that there exists a vertex v of $[-1, 1]^n$ such that for all $\alpha \in \lambda^+$, the maximal face of the function $x \mapsto (\beta - \alpha)^T x$ contains v.

Proof. We can assume $\beta \notin \operatorname{relint}(\operatorname{conv}(\lambda^+))$ since otherwise the preconditions imply $\beta = \alpha$ for all $\alpha \in \lambda^+$, violating $|\operatorname{supp}(\lambda)| \geq 3$. Hence, we have $\mathcal{A}\lambda \neq \mathbf{0}$ and the supremum of $x \mapsto (-\mathcal{A}\nu)^T x$ is attained at some vertex of $[-1, 1]^n$.

Now assume $\lambda \in \Lambda_X(\mathcal{A})$. In order to come up with a contradiction, we construct a decomposition of $\lambda = \sum_{\alpha \in \lambda^+} \nu^{(\alpha)}$ with supports $\operatorname{supp}\{\nu^{(\alpha)}\} = \{\alpha, \beta\}$ of cardinality 2 by setting

$$\theta_\alpha \nu_\alpha^{(\alpha)} := \lambda_\alpha \ \text{ and } \ \theta_\alpha \nu_\beta^{(\alpha)} := -\theta_\alpha \nu_\alpha^{(\alpha)} = -\lambda_\alpha \text{ for all } \alpha \in \lambda^+.$$

We observe that $\nu^{(\alpha)} \in N_\beta$ for all $\alpha \in \lambda^+$ and $(\theta_\alpha)_{\alpha \in \lambda^+}$ can be chosen with the property $\sum_{\alpha \in \lambda^+} \theta_\alpha = 1$. Moreover, $\sum\limits_{\alpha \in \lambda^+} \theta_\alpha \nu^{(\alpha)} = \lambda$ and

$$\begin{aligned}
\sum_{\alpha \in \lambda^+} \theta_\alpha \sigma_X(-\mathcal{A}\nu^{(\alpha)}) \ &= \ \sum_{\alpha \in \lambda^+} \theta_\alpha \sum_{j=1}^n \left| \nu_\alpha^{(\alpha)}(\alpha_j - \beta_j) \right| = \sum_{\alpha \in \lambda^+} \sum_{j=1}^n |\lambda_\alpha(\alpha_j - \beta_j)| \\
&\overset{(6.17)}{=} \sum_{j=1}^n \left| \sum_{\alpha \in \lambda^+} \lambda_\alpha(\alpha_j - \beta_j) \right| = \sigma_X(-\mathcal{A}\lambda).
\end{aligned}$$

By distinguishing the cases $\alpha_j = \beta_j$ and $\alpha_j \neq \beta_j$, it is straightforward to see that this expression in terms of a convex combination is locally linear. Hence, λ cannot be an X-circuit, which is the contradiction. $\square$

We provide a slightly more general version of Lemma 6.2.19, whose proof is analogous.

Lemma 6.2.20. *Let $\lambda \in N_\beta$ with $\lambda_\beta = -1$ for some $\beta \in \mathcal{A}$ and $|\operatorname{supp}(\lambda)| \geq 3$. Further, suppose that for $J(\lambda) := \{j : \beta_j = \sum_{\alpha \in \mathcal{A}} \lambda_\alpha \alpha_j\}$, the support can be disjointly decomposed into the two sets*

$$\mathcal{A}^{(1)} = \{\alpha : \alpha_j = \beta_j \ \forall\, j \notin J(\lambda)\} \ \text{ and } \ \mathcal{A}^{(2)} = \{\alpha : \alpha_j = \beta_j \ \forall\, j \in J(\lambda)\} \neq \emptyset$$

such that for all $j \in [n] \setminus J(\lambda)$, we have

$$\left(\alpha_j \leq \beta_j \text{ for all } \alpha \in \mathcal{A}^{(2)}\right) \ \text{ or } \ \left(\alpha_j \geq \beta_j \text{ for all } \alpha \in \mathcal{A}^{(2)}\right).$$

Then, λ is not an X-circuit of $\mathcal{A}$.

Example 6.2.21. *The planar case* $[-1,1]^2$. For the planar square $X = [-1,1]^2$, we provide some explicit descriptions of the sublinear circuits for support sets located on a grid $\{(i,j) : 1 \leq i,j \leq k\}$ for some $k \in \mathbb{N}$. If λ is a normalized $[-1,1]^2$-circuit, then, due to $\operatorname{rec}([-1,1]^2)^* = -\operatorname{rec}([-1,1]^2)^* = \mathbb{R}^2$, there is no restriction on the location of the negative coordinate. However, using Theorem 6.2.11, we can exclude potential sublinear circuits $\lambda \in N_\beta$ for some $\beta \in \mathcal{A}$ where $\operatorname{relint}\operatorname{conv}(\operatorname{supp}(\lambda)) \cap \lambda^+ \neq \emptyset$ and those where $\beta \in \operatorname{conv}(\lambda^+)$ but $\mathcal{A}\lambda \neq \mathbf{0}$; in particular, the latter situation excludes the case $\beta \in \operatorname{conv}(\lambda^+) \setminus \operatorname{relint}\operatorname{conv}(\lambda^+)$. Moreover, using Lemma 6.2.19, we can exclude all those potential $[-1,1]^2$-circuits where $|\operatorname{supp}(\lambda)| \geq 3$ and $[(\alpha_j \leq \beta_j$ for all $\alpha \in \lambda^+)$ or $(\alpha_j \geq \beta_j$ for all $\alpha \in \lambda^+)]$. For $\mathcal{A} = \{(i,j) : 1 \leq i,j \leq 3\}$, i.e., for the case $k = 3$, the structural statements allow us to obtain the exact set of sublinear circuits. Up to multiples, there are 132 X-circuits:

1. 72 sublinear circuits supported on two elements, namely, $\delta^{(i_1,j_1)} - \delta^{(i_2,j_2)}$ for $1 \leq i_1, i_2, j_1, j_2 \leq 3$ with $(i_1,j_1) \neq (i_2,j_2)$.
2. 27 sublinear circuits in which the entries $\frac{1}{2}, -1, \frac{1}{2}$ appear in columns 1,2, and 3, respectively.
3. 27 sublinear circuits in which the entries $\frac{1}{2}, -1, \frac{1}{2}$ appear in rows 1,2, and 3, respectively.
4. 8 sublinear circuits supported on 4 elements.

Since the diagonal and the anti-diagonal are counted both in cases (2) and (3), this gives 72+27+27+8-2 = 132 sublinear circuits. The following table shows the number of normalized sublinear circuits λ with $\lambda^- = \{(i,j)\}$ in row i and column j.

	1	2	3
1	8	17	8
2	17	32	17
3	8	17	8

The subsequent list contains all 17 – not necessarily normalized – X-circuits ν with $\nu^- = \{(1,2)\}$, i.e., where the component with index $(1,2)$ is the negative component.

As before, the upper left entry of each matrix refers to the support point $(1,1)$:

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$\begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

The case $k = 4$. In the case $\mathcal{A} = \{(i,j) : 1 \le i,j \le 4\}$, a computer calculation shows that there are 980 normalized X-circuits, which come in the following classes with regard to λ^-:

	1	2	3	4
1	15	47	47	15
2	47	136	136	47
3	47	136	136	47
4	15	47	47	15

Note that in this case, the criteria of this and the previous section are not sufficient to determine the set of sublinear circuits.

6.3 Reduced Sublinear Circuits in X-SAGE-Cones

The previous section showed that an X-SAGE-cone is generated using only X-circuits and nonnegative exponentials supported on a single point. In this section, we seek to understand whether all X-circuits are necessary in this representation.

The answer to this question depends on whether one means to reconstruct an individual X-AGE-cone, or the larger X-SAGE-cone. For example, by reinterpreting results from [MCW21a], we may infer that every simplicial $\mathbb{R}^n$-circuit $\lambda \in \Lambda_{\mathbb{R}^n}(\mathcal{A},\beta)$ generates a λ-witnessed AGE-cone containing an extreme ray of $C_{\mathbb{R}^n}(\mathcal{A},\beta)$. In this way, every $\mathbb{R}^n$-circuit is needed if one requires complete reconstruction of individual AGE-cones. However, in Chapter 3, we showed that many extreme rays of AGE-cones are *not* extreme when considered in the sum $C_{\mathbb{R}^n}(\mathcal{A}) = \sum_{\beta\in\mathcal{A}} C_{\mathbb{R}^n}(\mathcal{A},\beta)$. Specifically, an $\mathbb{R}^n$-circuit $\lambda \in \Lambda_{\mathbb{R}^n}(\mathcal{A})$ is only needed in $C_{\mathbb{R}^n}(\mathcal{A})$ if exactly one element of $\mathcal{A}$ hits the relative interior of $\operatorname{conv}(\operatorname{supp}\lambda)$, see Chapter 3, Theorem 3.2.1. Circuits satisfying this property were called *reduced*.

The goal of this section is to develop a reducedness criterion for X-circuits that yields the most efficient construction of $C_X(\mathcal{A})$ by λ-witnessed X-AGE-cones, see Theorems 6.3.5 and 6.3.6. We also seek to providing sufficient and necessary criteria for extremality in certain polyhedric special cases.

6.3.1 Definitions, Results, and Discussion

The definition of a reduced $\mathbb{R}^n$-circuit is of a purely combinatorial nature, involving the circuit's support. This is appropriate because when speaking of affine-linear simplicial circuits, the normalized vector representation λ is completely determined by its support. As explained in the previous section, in the context of X-circuits, we no longer have this property.

Therefore, when developing reduced *X-circuits*, it is useful to have a different characterization of reduced *$\mathbb{R}^n$-circuits*. Here, we can consider how Forsgård and de Wolff defined the *Reznick cone of $\mathcal{A}$* as the conic hull $R_{\mathbb{R}^n}(\mathcal{A}) := \operatorname{pos} \Lambda_{\mathbb{R}^n}(\mathcal{A})$ and — in the language of Chapter 3 — subsequently proved that an $\mathbb{R}^n$-circuit λ is an edge generator of $R_{\mathbb{R}^n}(\mathcal{A})$ if and only if it is reduced [FW19].

Our definition of reduced X-circuits involves edge generators of a certain cone in a dimension one higher than the Reznick cone. To describe the cone and facilitate later analysis, we need the following definition.

Definition 6.3.1. The *functional form* of an X-circuit $\nu \in \mathbb{R}^{\mathcal{A}}$ is $\phi_\nu : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}$ defined by

$$\phi_\nu(y) = \sum_{\alpha \in \mathcal{A}} y_\alpha \nu_\alpha + \sigma_X(-\mathcal{A}\nu).$$

We routinely overload notation and use $\phi_\nu = (\nu, \sigma_X(-\mathcal{A}\nu)) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}$ to denote the functional form of a given X-circuit. When representing the functional form of an X-circuit by a vector in $\mathbb{R}^{\mathcal{A}} \times \mathbb{R}$, the scalar $\phi_\nu(y)$ can be expressed as an inner product $\phi_\nu(y) = (y, 1)^T \phi_\nu$.

Definition 6.3.2. The *circuit graph* of $(\mathcal{A}, X)$ is

$$G_X(\mathcal{A}) = \operatorname{pos}\left(\{\phi_\lambda \,:\, \lambda \in \Lambda_X(\mathcal{A})\} \cup \{(\mathbf{0}, 1)\}\right)$$

where $(\mathbf{0}, 1) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}$.

While the cones from Theorem 6.1.7 are considered for one $\beta \in \mathcal{A}$ at a time, the circuit graph accounts for all X-circuits at once. We note that

$$G_X(\mathcal{A}) = \operatorname{pos}\left(\{\phi_\lambda \,:\, \lambda \in \Lambda_X^\star(\mathcal{A})\} \cup \{(\mathbf{0}, 1)\}\right). \tag{6.18}$$

For the proof, we refer to [MNT20], Section 5.

The circuit graph also includes an extra generator that ultimately serves to make the following definition more stringent.

Definition 6.3.3. The *reduced X-circuits of $\mathcal{A}$* are vectors ν where $\nu / \|\nu\|_\infty \in \Lambda_X(\mathcal{A})$ and the corresponding functional form ϕ_ν generates an extreme ray of $G_X(\mathcal{A})$. The set of *normalized* reduced X-circuits is henceforth denoted by $\Lambda_X^\star(\mathcal{A})$.

There is a subtle issue here that, in order for reduced X-circuits to be of any use to us, the circuit graph must be pointed. This is ensured by our assumption of linear independence of the functions $\{e^{\langle \alpha, x \rangle}\}_{\alpha \in \mathcal{A}}$ on X (stated in Chapter 2). Regardless of whether or not the circuit graph is pointed, the following theorem holds; for a proof, compare [MNT20].

Theorem 6.3.4. *$C_X(\mathcal{A})^* = \operatorname{cl}\{\exp y \,:\, (y, 1) \in G_X(\mathcal{A})^*\}$.*

Theorem 6.3.4 is a tool that we combine with convex duality to obtain the following results.

Theorem 6.3.5. *If $\Lambda_X(\mathcal{A})$ is empty, then $C_X(\mathcal{A}) = \mathbb{R}^{\mathcal{A}}_+$. Otherwise,*

$$C_X(\mathcal{A}) = \mathrm{cl}\left(\mathrm{conv} \bigcup \left\{ C_X(\mathcal{A}, \lambda) \;:\; \lambda \in \Lambda^\star_X(\mathcal{A}) \right\}\right). \tag{6.19}$$

Proof. Using the representation $G_X(\mathcal{A}) = \mathrm{pos}\left(\{\phi_\lambda : \lambda \in \Lambda^\star_X(\mathcal{A})\} \cup \{(\mathbf{0}, 1)\}\right)$ provided by (6.18), we can express

$$(y, 1) \in G_X(\mathcal{A})^* \Leftrightarrow (y, 1)^T (\lambda, \sigma_X(-\mathcal{A}\lambda)) \geq 0 \;\forall\; \lambda \in \Lambda^\star_X(\mathcal{A}). \tag{6.20}$$

We obtain the following refined definition of $C_X(\mathcal{A})^*$ by combining (6.20) with Theorem 6.3.4:

$$C_X(\mathcal{A})^* = \left\{ v \in \mathbb{R}^{\mathcal{A}}_+ \;:\; \forall\, \lambda \in \Lambda^\star_X(\mathcal{A}),\; \beta := \lambda^-,\; e^{\sigma_X(-\mathcal{A}\lambda)} \prod_{\alpha \in \lambda^+} v_\alpha^{\lambda_\alpha} \geq v_\beta \right\}. \tag{6.21}$$

Whenever $\Lambda_X(\mathcal{A}) \neq \emptyset$, we can write Equation 6.21 as $C_X(\mathcal{A})^* = \bigcap_{\lambda \in \Lambda^\star_X(\mathcal{A})} C_X(\mathcal{A}, \lambda)^*$. We appeal to conic duality principles (see [Roc97, Corollary 16.5.2]) to obtain the claim of the theorem. □

We point out how Theorem 6.3.5 involves a closure around the union over λ-witnessed X-AGE-cones, while Theorem 6.2.4 has no such closure. The need for the closure here stems from an application of an infinite version of conic duality in the course of the theorem's proof, while our proof of Theorem 6.2.4 required no duality at all. The requisite use of conic duality is simpler when X is a polyhedron, as the following theorem suggests.

Theorem 6.3.6. *If X is a polyhedron and $\Lambda_X(\mathcal{A})$ is non-empty, then the associated conditional SAGE-cone is given by the finite Minkowski sum*

$$C_X(\mathcal{A}) = \sum_{\lambda \in \Lambda^\star_X(\mathcal{A})} C_X(\mathcal{A}, \lambda). \tag{6.22}$$

Moreover, there is no proper subset $\Lambda \subsetneq \Lambda^\star_X(\mathcal{A})$ for which $C_X(\mathcal{A}) = \sum_{\lambda \in \Lambda} C_X(\mathcal{A}, \lambda)$.

The first part of Theorem 6.3.6 follows easily from the arguments we use to prove Theorem 6.3.5. The second part of the theorem is much more delicate, and, in fact, is the reason why $G_X(\mathcal{A})$ is defined in the manner of 6.3.2 rather than merely $\mathrm{pos}\{\phi_\lambda : \lambda \in \Lambda_X(\mathcal{A})\}$. We will use another statement without proof from Section 5 of [MNT20]:

Lemma 6.3.7. *If X is polyhedral and $\Lambda \subsetneq \Lambda^\star_X(\mathcal{A})$, then there must exist a $\tilde{y} \in \mathbb{R}^{\mathcal{A}}$ satisfying $\phi_{\lambda'}(\tilde{y}) \geq 0$ for all $\lambda' \in \Lambda$, yet for some $\lambda \in \Lambda^\star_X(\mathcal{A}) \setminus \Lambda$, we have $\phi_\lambda(\tilde{y}) < 0$.*

Moreover, if $\tilde{y} \in \mathbb{R}^{\mathcal{A}}$ satisfies $\phi_\lambda(\tilde{y}) < 0$ for some $\lambda \in \Lambda_X(\mathcal{A})$, then we have $\exp \tilde{y} \notin C_X(\mathcal{A})^*$.

Proof of Theorem 6.3.6. Using Theorem 6.3.4, we can work with the dual description $C_X(\mathcal{A})^* = \mathrm{cl}\{\exp y \;:\; (y, 1) \in G_X(\mathcal{A})^*\}$. Applying (6.18) then gives

$$C_X(\mathcal{A})^* = \mathrm{cl}\{\exp y \;:\; \phi_\lambda(y) \geq 0 \;\;\forall\, \lambda \in \Lambda^\star_X(\mathcal{A})\},$$

as $\Lambda_X(\mathcal{A}) \neq \emptyset$. We rewrite the condition on $\phi_\lambda(y)$ as a condition on $v = \exp y$ using the power-cone formulation in Proposition 6.2.7. Since X is polyhedral, Theorem 6.1.8 tells us that there are finitely many normalized X-circuits $\Lambda_X(\mathcal{A})$. We may therefore express $C_X(\mathcal{A})^*$ as a finite intersection of dual λ-witnessed X-AGE-cones,

$$C_X(\mathcal{A})^* = \bigcap_{\lambda \in \Lambda_X^\star(\mathcal{A})} C_X(\mathcal{A}, \lambda)^*.$$

Moreover, each dual λ-witnessed X-AGE-cone $C_X(\mathcal{A}, \lambda)^*$ is an outer approximation of the full-dimensional moment cone $\operatorname{pos}\{e^{\mathcal{A}^T x} : x \in X\}$, hence, there exists a point v_0 in the interior of the moment cone where $v_0 \in \operatorname{int} C_X(\mathcal{A}, \lambda)^*$ for all $\lambda \in \Lambda_X^\star(\mathcal{A})$. Therefore, by [Roc97, Corollary 16.4.2] we have

$$C_X(\mathcal{A}) = (C_X(\mathcal{A})^*)^* = \sum_{\lambda \in \Lambda_X^\star(\mathcal{A})} (C_X(\mathcal{A}, \lambda)^*)^* = \sum_{\lambda \in \Lambda_X^\star(\mathcal{A})} C_X(\mathcal{A}, \lambda),$$

which establishes the first part of the theorem.

For the second part of the theorem, suppose Λ is a proper subset of $\Lambda_X^\star(\mathcal{A})$. Consider the set $C = \sum_{\lambda \in \Lambda} C_X(\mathcal{A}, \lambda)$ and its dual $C^* = \bigcap\{C_X(\mathcal{A}, \lambda)^* : \lambda \in \Lambda\}$. Clearly, since $C \subseteq C_X(\mathcal{A})$, we have $C^* \supset C_X(\mathcal{A})^*$ — we will show that this containment is *strict*, i.e., $C^* \supsetneq C_X(\mathcal{A})^*$. Once this is done, duality will tell us that $C \subsetneq C_X(\mathcal{A})$.

Since C is contained within the exponential sum nonnegativity cone, we have that C^* contains the moment cone but is still contained in the nonnegative orthant. As we have assumed X is non-empty, $C_X(\mathcal{A})^*$ contains a point $\exp(\mathcal{A}^T x) \in \mathbb{R}_{++}^{\mathcal{A}}$, so $C_X(\mathcal{A})^* \cap \operatorname{relint} \mathbb{R}_+^{\mathcal{A}} \neq \emptyset$. Rockafellar's [Roc97, Theorem 18.2] states that every relatively open set in $\mathbb{R}_+^{\mathcal{A}}$ is contained in the relative interior of some face of $\mathbb{R}_+^{\mathcal{A}}$. By our assumption $C_X(\mathcal{A})^* \cap \mathbb{R}_{++}^{\mathcal{A}} \neq \emptyset$, the only face of $\mathbb{R}_+^{\mathcal{A}}$ which contains $C_X(\mathcal{A})^*$ is $\mathbb{R}_+^{\mathcal{A}}$ itself. Since $\operatorname{relint} C_X(\mathcal{A})^*$ is relatively open, we have $\operatorname{relint} C_X(\mathcal{A})^* \subset \mathbb{R}_{++}^{\mathcal{A}}$, and the claim follows by the identity $C_X(\mathcal{A})^* = \operatorname{cl} \operatorname{relint} C_X(\mathcal{A})^*$. So, $C^* = \operatorname{cl}(C^* \cap \mathbb{R}_{++}^{\mathcal{A}})$.

Work with C^* over the positive orthant using Proposition 6.2.7 to express it as $C^* = \operatorname{cl}\{\exp y : \phi_\lambda(y) \geq 0 \ \forall \lambda \in \Lambda\}$. By the first statement of Lemma 6.3.7 there exists an element $\tilde{y} \in Y$ for which some $\lambda \in \Lambda_X^\star(\mathcal{A}) \setminus \Lambda$ satisfies $\phi_\lambda(\tilde{y}) < 0$. Apply the second statement of Lemma 6.3.7 to this pair (ϕ_λ, y) to see that $\exp \tilde{y}$ can be separated from the closed convex set $C_X(\mathcal{A})^*$. We have therefore found a point $\tilde{y}$ where $\exp \tilde{y} \in C^*$ and yet $\exp \tilde{y}$ can be separated from $C_X(\mathcal{A})^*$, so we conclude $C^* \supsetneq C_X(\mathcal{A})^*$. □

The task of actually finding the reduced X-circuits of $\mathcal{A}$ is difficult. When X is a polyhedron, there are finitely many such X-circuits, but the naive method for finding them involves Fourier-Motzkin elimination on a set of potentially very high dimension. There is more hope for this problem when X is a cone. In that case, X-circuits are the extreme rays of $(\ker \mathcal{A} + \mathcal{A}^+ X^*) \cap N_\beta$ for $\beta \in \mathcal{A}$, and no lifting is needed to find these extreme rays with a computer. The reduced X-circuits could then be computed by finding the extreme rays of the convex cone generated by the X-circuits.

6.3.2 Reducibility and Extremality

By Theorem 6.3.6, the reduced sublinear circuits provide an irredundant decomposition of conditional SAGE-cones. In this section, we discuss some criteria and key examples for reduced sublinear circuits. As an application of the criteria, we will determine the extreme rays of the $\mathbb{R}_+$-SAGE-cone in Theorem 6.4.1 and of the $[-1, 1]$-SAGE-cone in Theorem 6.4.2. Remember from Chapter 3, Proposition 3.2.5, that for the classical case of affine circuits supported on a finite set $\mathcal{A}$, the following exact characterization in terms of the support is known. For the course of this section, we fix $\mathcal{A} \subseteq \mathbb{R}^n$ non-empty and finite.

Proposition 6.3.8. (also compare [KNT21, Corollary 4.7], [FW19, Theorem 3.2]) *A vector ν is a reduced $\mathbb{R}^n$-circuit if and only if*

$$\mathcal{A} \cap \operatorname{relint} \operatorname{conv} \nu^+ = \{\nu^-\}.$$

For example, with regard to the two matrices in (6.15) of Example 6.2.14, the left one is not reduced, but the right one is. The following theorem gives a generalization to the constrained situation for the necessary direction of Proposition 6.3.8 when X is a non-empty, convex set in $\mathbb{R}^n$.

Theorem 6.3.9. *Let $\lambda \in \Lambda_X(\mathcal{A}, \beta)$. If there exists $\beta' \in \mathcal{A} \setminus \operatorname{supp} \lambda$ and some normalized $\lambda' \in N_{\beta'}$ where $(\lambda')^+ \subseteq \operatorname{supp}(\lambda)$ and $\mathcal{A}\lambda' = \gamma \mathcal{A}\lambda$ for some $\gamma \geq 0$, then λ is not reduced.*

Before providing the proof within this section, we discuss its consequences.

Corollary 6.3.10. *Let $\lambda \in \Lambda_X(\mathcal{A}, \beta)$. If $(\operatorname{conv}(\operatorname{supp}(\lambda)) \cap \mathcal{A}) \setminus \operatorname{supp}(\lambda) \neq \emptyset$, then λ is not reduced. Consequently,*

$$\{\lambda \in \Lambda_X^\star(\mathcal{A}) : \mathcal{A}\lambda = \mathbf{0}\} \subseteq \Lambda_{\mathbb{R}^n}^\star(\mathcal{A}).$$

Proof. Let $(\lambda')^+$ denote the vertices of $\operatorname{supp}(\lambda)$ and set $\gamma = 0$. Then, the first statement is a consequence of Theorem 6.3.9 for $\beta' \in \operatorname{conv}(\operatorname{supp}(\lambda)) \setminus \operatorname{supp}(\lambda)$. The second one is a direct consequence of Proposition 6.3.8 and the fact that for $X = \mathbb{R}^n$ all X-circuits λ have the property $\mathcal{A}\lambda = \mathbf{0}$. □

Using this corollary, we can provide an analogue to Theorem 6.2.18.

Theorem 6.3.11. *Let $X = \operatorname{pos}\{v^{(1)}, \ldots, v^{(k)}\}$ be an n-dimensional polyhedral cone spanned by the vectors $v^{(1)}, \ldots, v^{(k)}$ where $k \geq n$. Then,*

$$\{\lambda \in \Lambda_X^\star(\mathcal{A}) \ : \ \mathcal{A}\lambda = 0\} \ = \ \Lambda_{\mathbb{R}^n}^\star(\mathcal{A}). \tag{6.23}$$

As in Theorem 6.2.18, there are no elements in $\Lambda_{\mathbb{R}^n}^\star(\mathcal{A})$ which describe circuits supported on a single element, as we used them in Chapters 3, 4 and 5.

Proof. By Corollary 6.3.10, every $\lambda \in \Lambda_X^\star(\mathcal{A})$ is contained in $\Lambda_{\mathbb{R}^n}^\star(\mathcal{A})$. Suppose there exists some $\lambda \in \Lambda_{\mathbb{R}^n}^\star(\mathcal{A})$ that is not contained in $\Lambda_X^\star(\mathcal{A})$. By Theorem 6.2.18, we have $\lambda \in \Lambda_X(\mathcal{A})$. As $\lambda \notin \Lambda_X^\star(\mathcal{A})$, there exist $m \in \mathbb{N}$ and X-circuits $\nu^{(1)}, \ldots, \nu^{(m)}$, non-proportional to λ, which satisfy $\sum_{i \leq m}(\nu^{(i)}, \sigma_X(-\mathcal{A}\nu^{(i)})) = (\lambda, \sigma_X(-\mathcal{A}\lambda))$. Since $\sigma_X(-\mathcal{A}\lambda) = 0$ and $\sigma_X(y) \in \{0, \infty\}$ for all $y \in \mathbb{R}^n$, we have $\sigma_X(-\mathcal{A}\nu^{(i)}) = 0$ for all $i \in [m]$.

As in Theorem 6.2.18, denote by W the $k \times n$-matrix whose rows are the transposed vectors $(v^{(1)})^T, \ldots, (v^{(k)})^T$. Again,

$$\sigma_X(-y) < \infty \text{ if and only if } Wy \geq \mathbf{0}.$$

Since $\mathcal{A}\lambda = W\mathcal{A}\lambda = \mathbf{0}$, we obtain $W\mathcal{A}\nu^{(i)} = \mathbf{0}$ and, as the kernel of W is $\{\mathbf{0}\}$, further, $\mathcal{A}\nu^{(i)} = \mathbf{0}$ for all $i \in [m]$. Hence, $\nu^{(i)} \in \Lambda_{\mathbb{R}^n}^\star(\mathcal{A})$ and therefore $\lambda \notin \Lambda_{\mathbb{R}^n}^\star(\mathcal{A})$, which is a contradiction. □

We illustrate the applicability of Theorem 6.3.9 in determining the reduced sublinear circuits by returning to the univariate examples $X = [-1, 1]$ and $X = \mathbb{R}_+$, which were started in Theorem 6.1.9.

Theorem 6.3.12. *Let $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$ be sorted ascendingly where $m \geq 3$.*

(a) Let $X = [-1, 1]$. Then, $\Lambda_X^\star(\mathcal{A})$ consists of those X-circuits which have the form (1) $\lambda = \delta^{(2)} - \delta^{(1)}$ or $\lambda = \delta^{(m-1)} - \delta^{(m)}$ or (2)

$$\lambda = \frac{\alpha_{i-1} - \alpha_i}{\alpha_{i-1} - \alpha_{i+1}} \delta^{(i-1)} - \delta^{(i)} + \frac{\alpha_{i-1} - \alpha_i}{\alpha_{i-1} - \alpha_{i+1}} \delta^{(i+1)} \text{ for some } i \in \{2, \ldots, m-1\}.$$

(b) Now, let $X = \mathbb{R}_+$. Then, $\Lambda_X^\star(\mathcal{A})$ consists of those X-circuits which have the form (1) $\lambda = \delta^{(2)} - \delta^{(1)}$ or (2)

$$\lambda = \frac{\alpha_{i-1} - \alpha_i}{\alpha_{i-1} - \alpha_{i+1}} \delta^{(i-1)} - \delta^{(i)} + \frac{\alpha_{i-1} - \alpha_i}{\alpha_{i-1} - \alpha_{i+1}} \delta^{(i+1)} \text{ for some } i \in \{2, \ldots, m-1\}.$$

Note that, in particular, this gives

$$\Lambda_X^\star(\mathcal{A}) \cap \left\{ \lambda \in \bigcup_{\beta \in \mathcal{A}} N_\beta : |\operatorname{supp}(\lambda)| = 3 \right\} \;=\; \Lambda_{\mathbb{R}}^\star(\mathcal{A})$$

for both cases. Remember from Chapter 3 that

$$\Lambda_{\mathbb{R}}^\star(\mathcal{A}) = \left\{ \left(\frac{\alpha_{i+1} - \alpha_i}{\alpha_{i+1} - \alpha_{i-1}} \right) e^{i-1} + \left(\frac{\alpha_i - \alpha_{i-1}}{\alpha_{i+1} - \alpha_{i-1}} \right) e^{i+1} - e^i \;:\; 1 < i < m \right\}.$$

Proof of Theorem 6.3.12.

(a) By Theorem 6.1.9 and Corollary 6.3.10, the only candidates for normalized reduced X-circuits are

  (i) $\lambda = \delta^{(i)} - \delta^{(i\pm1)}$ or

  (ii) $\lambda = \dfrac{\alpha_i - \alpha_{i+1}}{\alpha_{i-1} - \alpha_{i+1}} \delta^{(i-1)} - \delta^{(i)} + \dfrac{\alpha_{i-1} - \alpha_i}{\alpha_{i-1} - \alpha_{i+1}} \delta^{(i+1)}$ for $i \in \{2, \ldots, m-1\}$.

For every X-circuit $\delta^{(i+1)} - \delta^{(i)}$ with $i > 1$, the X-circuit $\delta^{(i+1)} - \delta^{(1)}$ satisfies the precondition of Theorem 6.3.9 and for every X-circuit $\delta^{(i-1)} - \delta^{(i)}$ with $i < m$, the X-circuit $\delta^{(i-1)} - \delta^{(m)}$ satisfies the precondition of Theorem 6.3.9. Hence, all those X-circuits are not reduced.

We see that for all $i \in [m]$, there is precisely one normalized X-circuit λ that appears in the listed set of *possible* reduced X-circuits. As $\operatorname{rec}(X)^* = \mathbb{R}$, there exists at least one X-AGE exponential where the i-th coefficient is negative, hence, $C_X(\mathcal{A}, \alpha_i) \neq \emptyset$ for all $i \in [m]$. As $C_X(\mathcal{A}, \alpha_i)$ is the union of several λ-witnessed X-AGE-cones and — by results earlier in this chapter — those cones can be solely represented by reduced X-circuits, for every $i \in [m]$, there exists at least one reduced X-circuit in $C_X(\mathcal{A}, \alpha_i)$. With this, the statement follows.

(b) As a first step towards seeing this, observe that since $X = [0, \infty)$ is a cone, the functional form of a $[0, \infty)$-circuit ν is simply $\phi_\nu(y) = \sum_{i=1}^m y_i \nu_i$. Hence, the reduced $[0, \infty)$-circuits are exactly the edge generators of the cone $\operatorname{pos} \Lambda_{[0,\infty)}$ generated by all the $[0, \infty)$-circuits of types (1) and (2) listed in Theorem 6.1.9.

Therefore, we have to show that $\{\delta^{(2)} - \delta^{(1)}\} \cup \Lambda^{\star}_{\mathbb{R}}(\mathcal{A})$ are exactly the normalized edge generators of $\operatorname{pos} \Lambda_{[0,\infty)}$.

For the X-circuits $\delta^{(j)} - \delta^{(i)}$ $(j > i)$ of type (1) in Theorem 6.1.9, we show that they decompose if $j > i+1$ or $i > 1$. For $j > i+1$, this is apparent from the decomposition

$$\delta^{(j)} - \delta^{(i)} \;=\; (\delta^{(j)} - \delta^{(j-1)}) + (\delta^{(j-1)} - \delta^{(i)}).$$

For $j = i+1$ and $i > 1$, $\delta^{(i+1)} - \delta^{(i)}$ decomposes into

$$\begin{aligned} &\left(-\frac{\alpha_{i+1} - \alpha_i}{\alpha_i - \alpha_{i-1}}\delta^{(i-1)} + \frac{\alpha_{i+1} - \alpha_i}{\alpha_i - \alpha_{i-1}}\delta^{(i)}\right) \\ +&\left(\frac{\alpha_{i+1} - \alpha_i}{\alpha_i - \alpha_{i-1}}\delta^{(i-1)} - \frac{\alpha_{i+1} - \alpha_{i-1}}{\alpha_i - \alpha_{i-1}}\delta^{(i)} + \delta^{(i+1)}\right), \end{aligned}$$

i.e., into X-circuits with three non-vanishing components. As final consideration for type (1), the X-circuit $\delta^{(2)} - \delta^{(1)}$ cannot be written as a conic combination of X-circuits with three nonzero entries because any conic combination of those X-circuits has a positive entry in its non-vanishing component with maximal index. For X-circuits of type (2) from Theorem 6.1.9, simply note that these are also $\mathbb{R}$-circuits. Therefore a necessary condition for a type (2) X-circuit λ to be extremal in $\operatorname{pos} \Lambda_{[0,\infty)}$ is that λ belongs to $\Lambda^{\star}_{\mathbb{R}}(\mathcal{A})$.

It remains to be shown that none of the remaining X-circuits can be written as a convex combination of the others. First note that an X-circuit $\nu \in \Lambda^{\star}_{\mathbb{R}}(\mathcal{A})$ cannot be decomposed into a sum which involves an X-circuit $\tilde{\nu}$ with two vanishing components. Namely, since $\mathcal{A}\nu = 0$ and $\mathcal{A}\tilde{\nu} > 0$, we would obtain $\nu - \tilde{\nu}$ with the property $\mathcal{A}(\nu - \tilde{\nu}) < 0$, and thus, $\sigma_{[0,\infty)}(-\mathcal{A}(\nu - \tilde{\nu})) = \infty$, a contradiction. And of course it is trivially true that no element $\lambda \in \Lambda^{\star}_{\mathbb{R}}(\mathcal{A})$ can be written as a convex combination of other such elements. Since $\operatorname{pos} \Lambda_{[0,\infty)}$ is finitely generated and there is no $S \subsetneq \{\delta^{(2)} - \delta^{(1)}\} \cup \Lambda^{\star}_{\mathbb{R}}(\mathcal{A})$ for which $\operatorname{pos} \Lambda_{[0,\infty)} = \operatorname{pos} S$, we conclude that $\{\delta^{(2)} - \delta^{(1)}\} \cup \Lambda^{\star}_{\mathbb{R}}(\mathcal{A})$ are the reduced X-circuits of $\mathcal{A}$.

□

Proof of Theorem 6.3.9. Since λ and λ' are normalized elements in N_β and $N_{\beta'}$, we have

$$\begin{aligned} \textstyle\sum_{\alpha \in \lambda^+} \lambda_\alpha &= 1 \text{ and } \quad \lambda_\beta = -1, \lambda_\alpha \geq 0 \text{ for } \alpha \in \mathcal{A} \setminus \{\beta\}, \\ \textstyle\sum_{\alpha \in (\lambda')^+} \lambda'_\alpha &= 1 \text{ and } \quad \lambda'_{\beta'} = -1, \lambda'_\alpha \geq 0 \text{ for } \alpha \in \mathcal{A} \setminus \{\beta'\}. \end{aligned}$$

Let τ be the maximal real number in $[0, 1/\gamma]$ (with the convention $1/\gamma := \infty$ if $\gamma = 0$) such that $\nu^{(1)} := \lambda - \tau\lambda' \in N_\beta$. That maximum clearly exists, and since $(\lambda')^+ \subseteq \operatorname{supp}(\lambda)$, the number τ is positive. Moreover, since λ and λ' are normalized and distinct, we have $\tau < 1$.

Similarly, let τ' be the maximal real number in $[0, \gamma]$ with $\nu^{(2)} := \lambda' - \tau'\lambda \in N_{\beta'}$. Here, we have $0 \leq \tau' \leq 1$ (and, in particular, $\tau' = 0$ if $\gamma = 0$ or $(\lambda')^+ \subsetneq \lambda^+$). Hence, $\nu^{(1)} \in N_\beta, \nu^{(2)} \in N_{\beta'}$ and $1 - \tau\tau' \in (0, 1]$.

Since $\nu^{(1)} + \tau\nu^{(2)} = \lambda - \tau\lambda' + \tau\lambda' - \tau\tau'\lambda = (1 - \tau\tau')\lambda$, we see that λ can be written as a conic combination of the two non-proportional (not necessarily normalized) elements $\nu^{(1)} \in N_\beta$ and $\nu^{(2)} \in N_{\beta'}$. Due to $\mathcal{A}\lambda' = \gamma\mathcal{A}\lambda$ and as both $1 - \tau\gamma \geq 0$ and

$\gamma - \tau' \geq 0$, we obtain

$$\begin{aligned}
\sigma_X(-\mathcal{A}\nu^{(1)}) &= \sigma_X(-\mathcal{A}\lambda + \tau\mathcal{A}\lambda') = \sigma_X(-\mathcal{A}\lambda + \tau\gamma\mathcal{A}\lambda) \\
&= (1-\tau\gamma)\sigma_X(-\mathcal{A}\lambda) = \sigma_X(-\mathcal{A}\lambda) - \tau\sigma_X(-\mathcal{A}\lambda'), \\
\sigma_X(-\mathcal{A}\nu^{(2)}) &= \sigma_X(-\mathcal{A}\lambda' + \tau'\mathcal{A}\lambda) = \sigma_X(-\gamma\mathcal{A}\lambda + \tau'\mathcal{A}\lambda) \\
&= (\gamma - \tau')\sigma_X(-\mathcal{A}\lambda) = \sigma_X(-\mathcal{A}\lambda') - \tau'\sigma_X(-\mathcal{A}\lambda)
\end{aligned}$$

and further

$$\begin{aligned}
\sigma_X(-\mathcal{A}\lambda) &= \frac{1}{1-\tau\tau'}\left(\sigma_X\left(-\mathcal{A}\lambda\right) - \tau\sigma_X\left(-\mathcal{A}\lambda'\right) + \tau\sigma_X\left(-\mathcal{A}\lambda'\right) - \tau\tau'\sigma_X\left(-\mathcal{A}\lambda\right)\right) \\
&= \frac{1}{1-\tau\tau'}\left(\sigma_X\left(-\mathcal{A}\nu^{(1)}\right) + \tau\sigma_X\left(-\mathcal{A}\nu^{(2)}\right)\right),
\end{aligned}$$

which shows that $(\lambda, \sigma_X(-\mathcal{A}\lambda))$ does not generate an extreme ray in $G_X(\mathcal{A})$. By definition of a reduced sublinear circuit, $\lambda \in \Lambda_X^{\star}(\mathcal{A})$. □

Example 6.3.13. *The reduced sublinear circuits for the cube* $[-1,1]^2$*.* We consider again the support $\mathcal{A} = \{(i,j) : 1 \leq i,j \leq k\}$ for some $k \in \mathbb{N}$. In the case $k = 3$, there are 24 normalized reduced X-circuits, which come in the following classes:

1. 12 sublinear circuits with entries $1, -1$, namely,
   (a) 8 with entry -1 in a corner and entry $+1$ beside or below the corner,
   (b) 4 with entry -1 in a non-corner boundary entry and entry $+1$ in the central, interior entry,
2. 8 sublinear circuits, where the sequence $\frac{1}{2}, -1, \frac{1}{2}$ appears in a row (3 possibilities), in a column (3 possibilities) or on the diagonal or antidiagonal,
3. 4 sublinear circuits supported on 4 elements, namely

$$\begin{pmatrix} 0 & 1/3 & 0 \\ 1/3 & -1 & 0 \\ 0 & 0 & 1/3 \end{pmatrix},$$

   as well as the 90-degree, 180-degree and 270-degree rotation of this matrix. Note that when starting from the set of all sublinear circuits λ for $[-1,1]^2$, Theorem 6.3.9 is applicable to rule out that λ is reduced in a number of cases. For example, the matrices

$$\begin{pmatrix} 0 & 0 & 1/2 \\ 1/2 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 & 1/2 \\ 1/2 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

   represent sublinear circuits λ and λ' with the property $\mathcal{A}\lambda = (-1/2, 0)^T$ and $\mathcal{A}\lambda' = (-3/2, 0)^T$, to which Theorem 6.3.9 can be applied in order to show that λ is not reduced. Also note that all reduced $\mathbb{R}^2$-circuits for the support set $\mathcal{A}$ are also reduced $[-1,1]^2$-circuits. Namely, since for all other $[-1,1]^2$-circuits λ, we have $\sigma_X(-\mathcal{A}\lambda) \neq \mathbf{0}$, those circuits cannot be used to decompose an $\mathbb{R}^2$-circuit (which has $\sigma_X(-\mathcal{A}\lambda) = \mathbf{0}$).

In the case $k = 4$ with 16 support points, a computer calculation shows that there are 72 reduced sublinear circuits.

6.4 Extreme Rays of Conditional SAGE-Cones in Dimension 1

In the previous section, we showed that by appropriate appeals to convex duality, one may derive representations of $C_X(\mathcal{A})$ with little to no redundancy. Here we build upon those results to completely characterize the extreme rays of the X-SAGE-cone for the univariate cases $X = [0, \infty)$ and $X = [-1, 1]$.

Theorem 6.4.1. *For $\alpha_1 < \cdots < \alpha_m$, the extreme rays of $C_{[0,\infty)}(\{\alpha_1, \ldots, \alpha_m\})$ are:*

(1) $\mathbb{R}_+ \cdot e^{\alpha_1 x}$

(2) $\mathbb{R}_+ \cdot \{e^{\alpha_2 x} - e^{\alpha_1 x}\}$

(3) $\mathbb{R}_+ \cdot \{c_{i+1}e^{\alpha_{i+1}x} + c_i e^{\alpha_i x} + c_{i-1}e^{\alpha_{i-1}x} \,:\, 2 \le i \le m-1\}$ *with*

$$c_{i+1} > 0, \qquad c_{i-1} > 0, \quad \textit{and} \quad c_i = -\left(\frac{c_{i-1}}{\lambda_{i-1}}\right)^{\lambda_{i-1}} \left(\frac{c_{i+1}}{\lambda_{i+1}}\right)^{\lambda_{i+1}}$$

where

$$\lambda_{i+1} = \frac{\alpha_i - \alpha_{i-1}}{\alpha_{i+1} - \alpha_{i-1}}, \quad \lambda_{i-1} = \frac{\alpha_{i+1} - \alpha_i}{\alpha_{i+1} - \alpha_{i-1}}, \quad \textit{and} \quad \frac{c_{i-1}}{c_{i+1}} \ge \frac{\lambda_{i-1}}{\lambda_{i+1}}$$

Theorem 6.4.2. *Let $X = [-1, 1]$ and $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$ be sorted ascendingly where $m \ge 3$. The extremal rays of $C_X(\mathcal{A})$ are the following:*

1. $\mathbb{R}_+ \cdot \left(e^{\alpha_2 x} - e^{\alpha_1 - \alpha_2} e^{\alpha_1 x}\right)$

2. $\mathbb{R}_+ \cdot \left(e^{\alpha_{m-1} x} - e^{\alpha_{m-1} - \alpha_m} e^{\alpha_m x}\right)$

3. $\mathbb{R}_+ \cdot \{c_{i-1}e^{\alpha_{i-1}x} + c_i e^{\alpha_i x} + c_{i+1}e^{\alpha_{i+1}x}\}$*, with*

$$c_{i-1} > 0,\ c_{i+1} > 0 \ \textit{and}\ c_i = -\left(\frac{c_{i-1}}{\lambda_{i-1}}\right)^{\lambda_{i-1}} \left(\frac{c_{i+1}}{\lambda_{i+1}}\right)^{\lambda_{i+1}} \quad \textit{where}$$

$$\lambda_{i-1} = \frac{\alpha_{i+1} - \alpha_i}{\alpha_{i+1} - \alpha_{i-1}},\ \lambda_{i+1} = \frac{\alpha_i - \alpha_{i-1}}{\alpha_{i+1} - \alpha_{i-1}} \ \textit{and}$$

$$\alpha_{i-1} - \alpha_{i+1} \le \ln \frac{c_{i-1}\lambda_{i+1}}{c_{i+1}\lambda_{i-1}} \le \alpha_{i+1} - \alpha_{i-1}$$

The rest of this section focuses on the proofs of the previous two theorems.

Proof of Theorem 6.4.1. Let $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$. By Theorem 6.3.12, all edge generators of $C_{[0,\infty)}(\mathcal{A})$ are either monomials or λ-witnessed X-AGE exponentials where λ is a reduced $[0, \infty)$-circuit. By Theorem 6.1.9, $\Lambda^\star_{[0,\infty)}(\mathcal{A}) = \{e^2 - e^1\} \cup \Lambda^\star_{\mathbb{R}}(\mathcal{A})$. Since $n = 1$, by Proposition 6.1.5 all X-circuits λ have $|\operatorname{supp} \lambda| \le 3$. We therefore divide the proof into considering cases of monomials and cases of X-AGE exponentials with two or three terms.

First we address the monomials. Given $f(x) = e^{\alpha_i x}$ with $i > 1$, we can write $f = f_1 + f_2$ with $f_1(x) = e^{\alpha_i x} - e^{\alpha_{i-1} x}$ and $f_2(x) = e^{\alpha_{i-1} x}$ — the summand f_1 is nonnegative on $[0, \infty)$ because $\alpha_i > \alpha_{i-1}$, and f_2 is globally nonnegative. Therefore, the only possible extremal monomial in $C_{[0,\infty)}(\mathcal{A})$ is $f(x) = e^{\alpha_1 x}$. Since $X = [0, \infty)$, the leading term of any $g \in C_X(\mathcal{A})$ must have a positive coefficient. Moreover, if g is non-proportional to f, the leading term of g must have an exponent greater than

α_1. Therefore, any convex combination of X-AGE exponentials $g \in C_{[0,\infty)}(\mathcal{A})$ which are non-proportional to f must disagree with $f(x)$ in the limit as x tends to infinity. We conclude f is extremal in $C_{[0,\infty)}(\mathcal{A})$.

Now we consider the 2-term case, where, by Theorem 6.3.12, we have to consider exponential sums of the form $f(x) = c_2 e^{\alpha_2 x} - c_1 e^{\alpha_1 x}$. We observe that f is nonnegative on $[0,\infty)$ if and only if $c_2 \geq c_1 \geq 0$, and, furthermore that such exponential sums are nonextremal unless $c_1 = c_2$. To see that $f(x) = e^{\alpha_2 x} - e^{\alpha_1 x}$ is indeed extremal, note that f cannot be written as a convex combination involving any 3-term X-AGE exponentials because any conic combination of 3-term X-AGE exponentials has a leading term with positive coefficient on $e^{\alpha_i x}$ for some $i \geq 3$.

We have already proven cases (1) and (2) of the proposition. By Theorem 6.3.12, we know that any extremal 3-term X-AGE exponential belongs to a λ-witnessed X-AGE-cone where λ is a reduced $\mathbb{R}$-circuit. These reduced $\mathbb{R}$-circuits have the property $\operatorname{supp}\lambda = \{i-1, i, i+1\}$ $\alpha_{i-1}\lambda_{i-1} + \alpha_{i+1}\lambda_{i+1} = \alpha_i$, $\lambda_i = -1$. Any X-AGE exponential with such a witness is nonnegative on all of $\mathbb{R}$. Therefore, any 3-term X-AGE exponential f that is *extremal* in $C_{[0,\infty)}(\mathcal{A})$ is also extremal in $C_{\mathbb{R}}(\mathcal{A}) \subseteq C_{[0,\infty)}(\mathcal{A})$, thus Theorem 3.2.1 implies

$$f(x) = c_{i+1}e^{\alpha_{i+1}x} - \left(\left(\frac{c_{i+1}}{\lambda_{i+1}}\right)^{\lambda_{i+1}}\left(\frac{c_{i-1}}{\lambda_{i-1}}\right)^{\lambda_{i-1}}\right)e^{\alpha_i x} + c_{i-1}e^{\alpha_{i-1}x}. \tag{6.24}$$

We have arrived at the final phase of proving part (3) of this proposition. By the equality case in the AM/GM inequality and using $e^{\alpha_i x} = \left(e^{\lambda_{i+1}\alpha_{i+1}x}\right)\left(e^{\lambda_{i+1}\alpha_{i-1}x}\right)$, one finds that the unique minimizer $x^\star$ for functions (6.24) satisfies

$$\left(\frac{c_{i+1}e^{\langle x^\star, \alpha_{i+1}\rangle}}{\lambda_{i+1}}\right) = \left(\frac{c_{i-1}e^{\langle x^\star, \alpha_{i-1}\rangle}}{\lambda_{i-1}}\right) \quad \Leftrightarrow \quad x^\star = \ln\left(\frac{c_{i-1}}{c_{i+1}}\frac{\lambda_{i+1}}{\lambda_{i-1}}\right)/(\alpha_{i+1} - \alpha_{i-1}).$$

If $V_i(\lambda, c) := (c_{i-1}\lambda_{i+1})/(c_{i+1}\lambda_{i-1})$ satisfies $V_i(\lambda, c) < 1$, then $x^\star < 0$ and, by continuity, we have $\inf\{f(x) : x \geq 0\} > 0$ — hence, the condition $V_i(\lambda, c) \geq 1$ is necessary for extremality. Furthermore, if $V_i(\lambda, c) > 1$, then the unique minimizer of f given by (6.24) occurs at $x^\star > 0$. Such f cannot be decomposed as a convex combination which involves 1-term or 2-term X-AGE exponentials (which have $f(x) > 0$ for $x > 0$), and cannot be written as a convex combination consisting solely of 3-term X-AGE exponentials, see Theorem 3.2.1, therefore any f given by (6.24) with $V_i(\lambda, c) > 1$ is extremal in $C_{[0,\infty)}(\mathcal{A})$. All that remains is to show extremality of functions (6.24) with $V_i(\lambda, c) = 1$. This follows from the same argument as $V_i(\lambda, c) > 1$, but we must use the stationarity condition $f'(0) = 0$ to preclude using 2-term extremal X-AGE exponentials in a decomposition of f. □

To prove the case $X = [-1, 1]$, we first deal with the atomic extreme rays, that is, extreme rays which are supported on a single element. These extreme rays are not captured by the X-circuit view.

Lemma 6.4.3 (Atomic Extreme Rays of $C_X(\mathcal{A})$ for Compact Sets X)**.** *Let $X \subseteq \mathbb{R}^n$ be a compact set and $\mathcal{A} \subseteq \mathbb{R}^n$ be finite with $|\mathcal{A}| \geq 2$. Then, there are no atomic extreme rays of $C_X(\mathcal{A})$.*

Proof. As in Lemma 6.2.12, we use the invariance of the X-circuits under translation of X and can, without loss of generality, assume $\mathbf{0} \in X$. Let $\alpha \neq \beta \in \mathcal{A}$ be arbitrary. Assume that $f = c_\alpha e^{\langle x, \alpha\rangle}$ with $c_\alpha > 0$ is extremal. We observe that $\lambda \in N_\beta$ with

$\lambda_\alpha = 1 = -\lambda_\beta$ is an X-circuit inducing the ray

$$\mathbb{R}_+ \cdot \left(e^{\langle x,\alpha \rangle} - \frac{1}{e^s} e^{\langle x,\beta \rangle} \right)$$

where $s \geq 0$ is finite and such that $\sigma_X(-\mathcal{A}\lambda)) = s$. Hence, the X-AGE exponentials

$$f^{(1)} = c_\alpha e^{\langle x,\alpha \rangle} - \frac{c_\alpha}{e^s} e^{\langle x,\beta \rangle}, f^{(2)} = \frac{c_\alpha}{s} e^{\langle x,\beta \rangle}$$

sum to f, contradicting the extremality of f. □

Proof of Theorem 6.4.2. Let $\mathcal{A} = \{\alpha_1, \ldots, \alpha_m\}$ be sorted ascendingly. By Lemma 6.4.3, there are no atomic extreme rays, and by Theorems 6.3.12 and 6.3.6, all the extreme rays are supported on two or three elements.

We start by considering the 2-term case. By Theorem 6.3.12, the only candidates for extreme rays are the ones given in the cases (1) and (2) in this statement. Since these cases are symmetric, it suffices to consider the case (1), i.e., some function $f(x) = e^{\alpha_2 x} - e\alpha_1 - \alpha_2 e \alpha_1 x$. As a conic combination of 3-term AGE exponentials and of functions of case (2) has a lowest-exponent term with positive coefficient, f cannot be written as a convex combination involving any 3-term X-AGE functions and of functions of case (2). Thus, f is indeed extremal.

Now, consider the 3-term case. By Theorem 6.3.12, the only candidates for extreme rays are of the form $f(x) = c_{i-1}e^{\alpha_{i-1}x} + c_i e^{\alpha_i x} + c_{i+1}e^{\alpha_{i+1}x}$ with $c_{i-1} > 0$, $c_{i+1} > 0$ and $c_i < 0$. The proof of Theorem 6.4.1 shows that f must have a zero in $[-1, 1]$ and that the location x^* of the zero is

$$x^* \;=\; \ln \left(\frac{c_{i-1}\lambda_{i+1}}{c_{i+1}\lambda_{i-1}} \right) / (\alpha_{i+1} - \alpha_{i-1})$$

where λ_{i-1} and λ_{i+1} are defined as in case (3) of this theorem. This gives the defining condition for c_i as well as the inequality conditions in case (3).

Any decomposition of f cannot involve a 2-term X-AGE exponential. For $x^* \in (-1, 1)$, this follows from the strict positivity of the 2-term X-AGE functions of type (1) and (2). For the boundary situations $x^* \in \{-1, 1\}$, we can additionally use the derivative condition $f'(x^*) = 0$ to exclude the 2-term X-AGE exponentials.

It remains to show that the 3-term X-AGE exponential f cannot be decomposed in terms of 3-term AGE exponentials. However, since f has a zero in $[-1, 1]$, and thus, in $\mathbb{R}$, by Theorem 3.2.1 it induces an extremal ray of the cone $C_{\mathbb{R}}(\mathcal{A})$ and cannot be decomposed using only 3-term X-AGE exponentials. □

Chapter 7

Constrained AM/GM-Based Optimization

As an application of the concepts introduced in the previous chapter, we examine some constrained optimization approaches based on the results for the unconstrained case studied in Chapters 4 and 5. Here, we are no longer interested in the optimal value of a given exponential sum on $\mathbb{R}^n$ but in the optimal value on a given set X.

As in the unconstrained case, we use sign constraints and the signed X-SAGE-cone, i.e., $\mathcal{A} = A^+ \cup A^-$ where $c_\alpha > 0$ for all $\alpha \in A^+$ and where $c_\alpha < 0$ for each $\alpha \in A^-$, and consider functions of the form

$$f \ = \sum_{\alpha \in A^+} c_\alpha e^{\langle x, \alpha \rangle} + \sum_{\alpha \in A^-} c_\alpha e^{\langle x, \alpha \rangle}. \tag{7.1}$$

Then, for a closed and convex set $X \subseteq \mathbb{R}^n$ and a finite set $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ with decomposition $\mathcal{A} = A^+ \cup A^-$ in the sense of (7.1), the *conditional signed SAGE-cone* $S_X(A^+, A^-)$ is the cone of all functions that can be written either as a sum of X-AGE exponentials of the form (7.1) or as elements $c_\alpha e^{\langle x, \alpha \rangle}$ with $\alpha \in A^+$ and $c_\alpha > 0$. Again, we denote the special case $A^- = \{\beta\}$ as $S_X(A^+, \beta)$.

Then, we examine the problem

$$\inf_{x \in X} f(x) \tag{7.2}$$

for functions $f \in \mathbb{R}^{\mathcal{A}}$ and $\mathcal{A} = A^+ \cup A^-$ as defined in (7.1).

In Section 7.1, we start by stating a symmetry-adapted decomposition theorem that is analogous to the unconstrained case, namely, Theorem 7.1.1. It is followed in Corollary 7.1.2 by a reduced relative entropy program certifying nonnegativity for G-symmetric constraint sets.

In Section 7.2, we take the dual point of view again. For a certain subset of the X-SAGE-cone that can be described by elements in the dual cone and a polyhedral conic constraint set, we present a linear program yielding a certificate for nonnegativity, which is based on two linear programs $(\mathrm{LP}^X_{A^+})$ and $(\mathrm{LP}^X_{A^-})$, see Theorem 7.2.3.

In the last part of this chapter, Section 7.3, we do not have the optimization viewpoint anymore — in particular, we consider the original X-SAGE-cone instead of the signed X-SAGE-cone. Here, we derive a second-order representation of the X-SAGE-cone and its dual for polyhedral constraint sets X and under the assumption on the support set $\mathcal{A}$ that $\mathcal{A}^T X$ is rational, see Corollaries 7.3.9 and 7.3.10. The sizes of the second-order-cone programs only depend on the number of sublinear circuits, which, by Chapter 6, is finite, their structure, and the number of constraints in the matrix representation of the polyhedral constraint set. Similar to the unconstrained case, the techniques used in this subsection feature a *(dual) circuit matrix.*

7.1 Exploiting Symmetries for Conditional SAGE

Both the symmetry-adapted decomposition theorem as well as the symmetrized relative entropy program in Chapter 4 allow a generalization to the constrained setting as soon as we have a symmetric constraint set. Exploiting this structure can be used to reduce the computation time of the resulting implementations.

First, we provide the constrained version of Theorem 4.2.3 for symmetric constraint sets X.

Theorem 7.1.1. *Let $X \subseteq \mathbb{R}^n$ be convex and G-invariant, let f be a G-invariant exponential sum of the form* (7.1) *and let $\hat{A}^-$ be a set of orbit representatives for A^-. Then, $f \in S_X(A^+, A^-)$ if and only if for every $\hat{\beta} \in \hat{A}^-$, there exists an X-AGE exponential $h_{\hat{\beta}} \in S_X(A^+, \hat{\beta})$ such that*

$$f = \sum_{\hat{\beta} \in \hat{A}^-} \sum_{\rho \in G/\operatorname{Stab}(\hat{\beta})} \rho h_{\hat{\beta}}. \tag{7.3}$$

The functions $h_{\hat{\beta}}$ can be chosen to be invariant under the action of $\operatorname{Stab}(\hat{\beta})$.

Similarly, we obtain a symmetrized version of the corresponding relative entropy program, analogous to Corollary 4.2.6.

Corollary 7.1.2. *Let $X \subseteq \mathbb{R}^n$ be convex and G-invariant. A G-invariant exponential sum f of the form* (7.1) *is contained in $S_X(A^+.A^-)$ if and only if for every $\hat{\beta} \in \hat{A}^-$, there exist $c^{(\hat{\beta})} \in \mathbb{R}_+^{\mathcal{A}}$ and $\nu^{(\hat{\beta})} \in \mathbb{R}_+^{\mathcal{A}}$ such that*

$$\begin{aligned} D(\nu^{(\hat{\beta})}, e \cdot c^{(\hat{\beta})}) + \sigma_X\Big(- \sum_{\alpha \in A^+} \nu_\alpha^{(\hat{\beta})}(\alpha - \hat{\beta})\Big) \quad &\leq \quad c_{\hat{\beta}} \quad \textit{for every } \hat{\beta} \in \hat{A}^- \textit{ and} \\ \sum_{\hat{\beta} \in \hat{A}^-} \sum_{\sigma \in \operatorname{Stab}\hat{\beta} \backslash G} c_{\sigma(\alpha)}^{(\hat{\beta})} \quad &\leq \quad c_\alpha \quad \textit{for every } \alpha \in A^+. \end{aligned}$$

Both statements can be proven following the respective proofs for the unconstrained setting from Chapter 4.

7.2 Constrained Optimization via the Dual of the X-SAGE-Cone and Linear Programming

Following the unconstrained case, in this subsection, we obtain a computationally fast approximation of constrained optimization problems with polyhedral conic constraint sets $X \subseteq \mathbb{R}^n$ via the dual X-SAGE-cone

$$S_X^*(A^+, A^-) \;=\; \left\{ \boldsymbol{v} \in \mathbb{R}^{\mathcal{A}} : \boldsymbol{v}(f) \geq 0 \text{ for all } f \in S_X(A^+, A^-) \right\}.$$

As we did in the unconstrained case, we associate a function

$$f = |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x \rangle} + \sum_{\alpha \in A^-} v_\alpha e^{\langle \alpha, x \rangle}$$

with every dual variable $v \in S_X^*(A^+, A^-) \cap \mathbb{R}_{\neq 0}^{\mathcal{A}}$, and call the constructed cone $\mathcal{F}_X^*(A^+, A^-)$; for the special case $A^- := \{\beta\}$, we define the cone $\mathcal{F}_X^*(A^+, \beta)$ analogously.

Proposition 7.2.1. *Let $X \subseteq \mathbb{R}^n$ be a conic, closed, convex set. It holds that*

1. $\mathcal{F}_X^*(A^+, \beta) \subseteq S_X(A^+, \beta)$ *for any* $A^+ \subseteq \mathbb{R}^n$ *and* $\beta \in \mathbb{R}^n$ *and*
2. $\mathcal{F}_X^*(A^+, A^-) \subseteq S_X(\mathcal{A})$ *for any* $A^+ \subseteq \mathbb{R}^n$ *and* $A^- \subseteq \mathbb{R}^n$.

As in the unconstrained case, in particular, every function in $\mathcal{F}_X^*(A^+, \beta)$ is non-negative on X.

We emphasize the fact that constraints describing the primal X-SAGE-cone involve some term "$-\sigma_X(-\mathcal{A}\lambda)$" on the greater side of the inequality whereas constraints describing the dual X-SAGE-cone involve some term "$\sigma_X(-\mathcal{A}\lambda)$" on the greater side of the inequality. Indeed, this fact makes it necessary to use the assumption that X is conic to prove the statement.

Proof.

1. Let $X \subseteq \mathbb{R}^n$ be a polyhedral, non-empty, and finite set, and $\mathcal{A} = A^+ \cup \{\beta\}$ be defined as in (7.1). We first recall that for polyhedral sets $X \subseteq \mathbb{R}^n$, the conditional signed SAGE-cone is the set

$$S_X(A^+, \beta) = \left\{ \sum_{\alpha \in A^+} c_\alpha e^{\langle x, \alpha \rangle} + c_\beta e^{\langle x, \beta \rangle} \ : \ c_\alpha \geq 0 \text{ for all } \alpha \in A^+, c_\beta < 0 \text{ and} \right. \\ \left. \exists\, \lambda \in \Lambda_X(\mathcal{A}, \beta) \text{ with } \prod_{\alpha \in \lambda^+} \left(\frac{c_\alpha}{\lambda_\alpha}\right)^{\lambda_\alpha} e^{\sigma_X(-\mathcal{A}\lambda)} \geq -c_\beta \right\}.$$

Moreover, for polyhedral X, we can use the representaion of the proof of Corollary 6.2.5. Then, applying Proposition 6.2.7, the dual of the signed X-AGE-cone can be described as

$$\left\{ \boldsymbol{v} \in \mathbb{R}^{\mathcal{A}} : \begin{array}{l} \text{for all } \ \alpha \in A^+, v_\alpha \geq 0; \text{ and for all } \ \beta \in A^-, \\ \text{for all } \ \lambda \in \Lambda_X(\mathcal{A}, \beta) \ , \ |v_\beta| \leq \prod_{\alpha \in \lambda^+} v_\alpha^{\lambda_\alpha} e^{\sigma_X(-\mathcal{A}\lambda)} \end{array} \right\}.$$

As X is a cone, we have $\sigma_X(-\mathcal{A}\lambda) = 0$ for all normalized X-circuits λ, and with this, in particular, $\sigma_X(-\mathcal{A}\lambda) = 0 = -\sigma_X(-\mathcal{A}\lambda)$. Let $f \in \mathcal{F}_X^*(A^+, \beta)$ with a corresponding vector of coefficients $\boldsymbol{v} \in \mathcal{S}_X(A^+, \beta) \cap \mathbb{R}_{\neq 0}^{A^+ \cup \{\beta\}}$. By the representation of the dual cone above, we have $v_\alpha > 0$ for all $\alpha \in A^+$, $v_\beta < 0$ by the definition of $\mathcal{F}_X^*(A^+, \beta)$, and for all $\lambda \in \Lambda_X(\mathcal{A}, \beta)$, it holds that

$$-v_\beta \leq |v_\beta| \leq \prod_{\alpha \in \lambda^+} v_\alpha^{\lambda_\alpha} e^{\sigma_X(-\mathcal{A}\lambda)} \leq \prod_{\alpha \in \lambda^+} \frac{v_\alpha}{\lambda_\alpha}^{\lambda_\alpha} e^{-\sigma_X(-\mathcal{A}\lambda)}.$$

The last inequality holds — besides the previous discussion on $\sigma_X(-\mathcal{A}\lambda)$ — for the same reason as in Proposition 4.3.1.

2. Applying the definitions of the primal and dual signed X-SAGE-cone and part (1), we obtain

$$\mathcal{F}_X^*(A^+, A^-) \subseteq \sum_{\beta \in A^-} S_X^*(A^+, \beta) \subseteq \sum_{\beta \in A^-} S_X(A^+, \beta) \subseteq S_X(A^+, A^-).$$

$\square$

As $X = \mathbb{R}^n$ is indeed a polyhedral cone, we know from examinations concerning the $\mathbb{R}^n$-SAGE-cone that the reverse implication does not hold in general.

As in the unconstrained case, an approximation of the constrained optimization problem using the dual X-SAGE-cone can be computed via linear programming.

Corollary 7.2.2. *Let $X \subseteq \mathbb{R}^n$ be a polyhedral cone and*

$$f \;=\; |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle x,\alpha \rangle} + \sum_{\alpha \in A^-} v_\alpha e^{\langle x,\alpha \rangle}$$

with $\boldsymbol{v} \in \mathbb{R}^{\mathcal{A}}_{\neq 0}$ and $A^+ \neq \emptyset$.

The following linear feasibility program in $|A^-|$ many variables $(x^{(\beta)})_{\beta \in A^-}$ verifies nonnegativity of f by certifying that $f \in \mathcal{F}^*_X(A^+, A^-)$ and, equivalently, containment of v in the dual X-SAGE-cone with respect to X.

$$\ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T x^{(\beta)} \;\; \textit{for all} \;\; \beta \in A^-,\; \alpha \in A^+,\; x^\beta \in X \tag{7.4}$$

Proof. Applying Theorem 2.4.14 together with Theorem 2.4.13, the dual of the signed X-SAGE-cone can be expressed as

$$\mathrm{cl}\left\{ \boldsymbol{v} \in \mathbb{R}^{\mathcal{A}} : \begin{array}{l} \text{for all} \;\; \alpha \in A^+, v_\alpha > 0; \text{ and for all } \;\beta \in A^- \\ \text{there exists } x \in X, \ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T x \;\; \text{for all} \;\; \alpha \in A^+ \end{array} \right\}.$$

Then, the proof of Corollary 7.2.2 can be deduced from the proof of Proposition 4.3.4. The only difference is the replacement of "$\tau \in \mathbb{R}^n$" by "$x^\beta \in X$". Note that here, we use the polyhedral condition on X: A polyhedron can be described using a set of linear inequalities, hence, the condition "$x^\beta \in X$" ensures the linearity of the program. □

Note that the previous proofs reveal that this is not the only reason we need the polyhedral condition here: It also ensures that the set of X-circuits is finite, and thus, we can describe the dual cone $\mathcal{S}^*_X(A^+, A^-)$ in terms of a power-cone, which allowed us earlier to prove containment of the cone $\mathcal{F}_X(A^+, A-)$ in $\mathcal{S}_X(A^+, A^-)$.

We proceed as we did in the unconstrained case.

For $\mathcal{A} = A^+ \cup A^-$ defined as in (7.1) and a fixed function

$$|A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x \rangle} + \sum_{\alpha \in A^-} v_\alpha e^{\langle \alpha, x \rangle},$$

we define a variable $w_{\mathbf{0}} := \begin{cases} -\check{\gamma} + v_{\mathbf{0}} \text{ if } \mathbf{0} \in A^-(f - \check{\gamma}), \text{ and} \\ (-1/|A^-|)\check{\gamma} + v_{\mathbf{0}} \text{ if } \mathbf{0} \in A^+(f - \check{\gamma}) \end{cases}.$

Then, we consider the following two linear programs in $|A^-|$ variables $(x^{(\beta)})_{\beta \in A^-}$ and $c = \ln(|w_{\mathbf{0}}|)$.

$$\begin{array}{ll} & \max \;\; c \\ & (1) \;\; \text{for all} \;\; \beta \in A^-, \text{for all } \alpha \in A^+ \setminus \{\mathbf{0}\} : \ln\left(\frac{|v_\beta|}{v_\alpha}\right) \leq (\alpha - \beta)^T x^{(\beta)}, \\ \text{s.t.} & (2) \;\; \ln\left(|v_\beta|\right) - c \leq (-\beta)^T x^{(\beta)} \text{ for all } \;\beta \in A^-, \\ & (3) \;\; x^{(\beta)} \in X \end{array} , \tag{$\mathrm{LP}^X_{A^+}$}$$

if $\mathbf{0} \in A^+(f - \check{\gamma})$ and

$$\begin{aligned}
&\max\ c && (\mathrm{LP}^X_{A^-})\\
\text{s.t.}\quad &(1)\quad \text{for all}\ \ \beta \in A^- \setminus \{\mathbf{0}\}, \text{for all}\ \alpha \in A^+ : \ln\left(\tfrac{|v_\beta|}{v_\alpha}\right) \leq (\alpha-\beta)^T x^{(\beta)},\\
&(2)\quad c - \ln(v_\alpha) \leq \alpha^T \tau^{(\mathbf{0})} \text{for all}\ \ \alpha \in A^+,\\
&(3)\quad x^{(\beta)} \in X
\end{aligned}$$

if $\mathbf{0} \in A^-(f - \check{\gamma})$.

As in the unconstrained case, we can draw the following result.

Theorem 7.2.3. *Let $X \subseteq \mathbb{R}^n$ be a polyhedral cone and*

$$f\ = |A^-| \sum_{\alpha \in A^+} v_\alpha e^{\langle \alpha, x\rangle} + \sum_{\alpha \in A^-} v_\alpha e^{\langle \alpha, x\rangle},$$

with $\emptyset \neq \mathcal{A} = A^+ \cup A^-$ defined as in (7.1), and let $f_X^{sage\ dual}$ be the optimal value with $f \geq f_X^{sage\ dual}$ on X. Assume that $f_X^{sage\ dual}$ satisfies $-f_X^{sage\ dual} + v_{\mathbf{0}} < 0$ or $-f_X^{sage\ dual} + |A^-|v_{\mathbf{0}} > 0$. The linear programs $(\mathrm{LP}^X_{A^+})$ and $(\mathrm{LP}^X_{A^-})$ solve the constrained optimization problem.

In the statement above, we omitted the same cases as in Chapter 4, namely,

1. $-f_X^{\text{sage dual}} + v_{\mathbf{0}} = 0 = -f_X^{\text{sage dual}} + |A^-|v_{\mathbf{0}}$ and
2. $-f_X^{\text{sage dual}} + v_{\mathbf{0}} \geq 0 \geq -f_X^{\text{sage dual}} + |A^-|v_{\mathbf{0}}$.

Both, however, can be treated precisely as in the unconstrained case.

7.3 Second-Order Representations for the Cone of X-AGE Exponentials and its Dual

Building upon the duality theory for constraint sets $X \subsetneq \mathbb{R}^n$ as examined in Chapter 6, we follow the approach from Chapter 5. We show that for polyhedral constraint sets X and the assumption on both the constraint and the support set $\mathcal{A} \subseteq \mathbb{R}^n$ that $\mathcal{A}^T X$ is rational, the conditional SAGE-cone as well as its dual can be represented by a second-order-cone program. This completes the observations in Corollary 6.2.5. The restriction on the constraint set stems from the result on power-cone representability of the X-SAGE-cone for polyhedral sets X, which we exploit in the following. The assumption on $\mathcal{A}^T X$ ensures that the set of normalized X-circuit $\Lambda_X(\mathcal{A})$ is rational.

Let $\emptyset \neq \mathcal{A} \subseteq \mathbb{R}^n$ be finite. Proposition 6.2.7 in conjunction with Corollary 6.2.5 and the representation in its proof tells us that a point $v \in \mathbb{R}^{\mathcal{A}}$ is contained in the dual X-SAGE-cone $C_X^*(\mathcal{A})$ if and only if $v \geq 0$ and

$$v_{\lambda^-} \leq \prod_{\alpha \in \lambda^+} v_\alpha^{\lambda_\alpha} e^{\sigma_X(-\mathcal{A}\lambda)} \text{ for every } \lambda \in \Lambda_X(\mathcal{A}). \tag{7.5}$$

We emphasize that for a fixed sublinear circuit λ, this only differs from the unconstrained case by the factor on the right hand side of the inequality. Moreover, in Theorem 6.1.8, it was shown that the amount of sublinear circuits is finite in the case of a polyhedral constraint set X. Throughout this section, we consider a fixed *rational* sublinear circuit $\lambda \in \Lambda_X(\mathcal{A}, \beta)$ for some $\beta \in \mathcal{A}$.

With this, we can proceed as we did in the unconstrained case — even obtaining the same sizes.

Fix a sublinear circuit $\lambda \in \Lambda_X(\mathcal{A}, \beta) \cap \mathbb{Q}^{\mathcal{A}}$. For $|\lambda^+| > 1$, let $p \in \mathbb{N}$ denote the smallest common denominator of the fractions λ_α for $\alpha \in \lambda^+$, i.e., $\lambda_\alpha = \frac{p_\alpha}{p}$ with $p_\alpha \in \mathbb{N}$ for all $\alpha \in \lambda^+$ and p is minimal. And for $|\lambda^+| = 1$, let $p = p_\alpha = 2$ for $\{\alpha\} = \lambda^+$. This way, we treat any sublinear circuit supported on more than two elements the same way as in the unconstrained case, but for the special case of a sublinear circuit supported on exactly two elements, we need a different procedure.

The associated set of *dual circuit variables* then is

$$(y_{k,i})_{k,i} \tag{7.6}$$

where $k \in [\lceil \log_2(p) \rceil]$ and $i \in [2^{\lceil \log_2(p) \rceil - k}]$. We denote the collection of these $\sum_{k=1}^{\lceil \log_2(p) \rceil} 2^{\lceil \log_2(p) \rceil - k} = 2^{\lceil \log_2(p) \rceil} - 1$ variables by y^λ or shortly by y. Further, we denote the restriction of a vector $v \in \mathbb{R}^{\mathcal{A}}$ to the components of $\lambda^+ \subseteq \mathcal{A}$ by $v_{|\lambda^+}$.

Definition 7.3.1. A *constrained dual circuit matrix* $C^*_{X,\mathcal{A},\lambda}(v_{|\lambda^+}, v_\beta, y)$ with respect to an X-circuit λ is a block diagonal matrix consisting of the blocks

$$\begin{pmatrix} y_{k-1,2i-1} & y_{k,i} \\ y_{k,i} & y_{k-1,2i} \end{pmatrix} \quad \text{for } k \in \{2, \ldots, \lceil \log_2(p) \rceil\} \text{ and } i \in [2^{\lceil \log_2(p) \rceil - k}], \tag{7.7}$$

the singleton blocks (v_β), and $(y_{\lceil \log_2(p) \rceil, 1} - v_\beta e^{-\sigma_X(-\mathcal{A}\lambda)})$ as well as $2^{\lceil \log_2(p) \rceil - 1}$ blocks of the form

$$\begin{pmatrix} u & y_{1,l} \\ y_{1,l} & w \end{pmatrix} \quad \text{for } l \in [2^{\lceil \log_2(p) \rceil - 1}] \tag{7.8}$$

where in each of these blocks u and w represent a variable of the set

$$\{v_\alpha \,:\, \alpha \in \lambda^+\} \cup \{v_\beta e^{-\sigma_X(-\mathcal{A}\lambda)}\}$$

such that altogether each v_α appears p_α times and $v_\beta e^{-\sigma_X(-\mathcal{A}\lambda)}$ appears $2^{\lceil \log_2(p) \rceil} - p$ times.

Remark 7.3.2. Note that for a sublinear circuit $\lambda \in \Lambda_X(\mathcal{A}, \beta)$ supported on exactly two elements with $\beta \in \mathcal{A}$ and α denoting the unique element in λ^+, the dual circuit matrix consists of the singleton blocks (v_β), $(y_{1,1} - v_\beta e^{-\sigma_X(-\mathcal{A}\lambda)})$, and the single 2×2 block

$$\begin{pmatrix} v_\alpha & y_{1,1} \\ y_{1,1} & v_\alpha \end{pmatrix},$$

yielding $v_\alpha \geq y_{1,1} \geq v_\beta e^{-\sigma_X(-\mathcal{A}\lambda)} \geq 0$.

The following theorem provides the analogous statement to Theorem 5.3.5.

Theorem 7.3.3. *The dual cone $C^*_X(\mathcal{A}, \lambda)$ of the λ-witnessed X-AGE-cone $C_X(\mathcal{A}, \lambda)$ is the projection of the spectrahedron*

$$\left\{(v, y) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{2^{\lceil \log_2(p) \rceil} - 1} \,:\, C^*_{X,\mathcal{A},\lambda}(v_{|\lambda^+}, v_\beta, y) \succcurlyeq 0\right\} \tag{7.9}$$

on $(v_{|\lambda^+}, v_\beta)$. $C^*_X(\mathcal{A}, \lambda)$ is second-order representable.

Proof. The first part of the proof can be deduced from the proof of Theorem 5.3.5 with the observation that for polyhedral X the number of X-circuits is finite. For the second part, suppose $X = \{x \in \mathbb{R}^n : Cx \leq d\}$ for some $C \in \mathbb{R}^{m\times n}$ and $d \in \mathbb{R}^m$. Then, $-\sigma_X(-\mathcal{A}\lambda)$ can be computed using a linear program with m inequalities. Now, again following the proof of Theorem 5.3.5, we obtain second-order representability. □

Naturally, for the size of the matrix $C^*_{X,A,\beta}(v_{|A}, v_\beta, y)$, we also find an analogous statement.

Corollary 7.3.4. *The matrix $C^*_{X,\mathcal{A},\lambda}(v_{|\lambda^+}, v_\beta, y)$ consists of $2^{\lceil \log_2(p)\rceil} - 1$ blocks of size 2×2 and two blocks of size 1×1.*

Following the discussion in Chapter 5, the same analogon works in the primal case for polyhedral constraint sets with *circuit variables*,

$$(x_\beta, (x_{k,i})_{k,i})$$

where $k \in [\lceil \log_2(p)\rceil]$ and $i \in [2^{\lceil \log_2(p)\rceil - k}]$. As in the dual case, we refer to these $1 + \sum_{k=1}^{\lceil \log_2(p)\rceil} 2^{\lceil \log_2(p)\rceil - k} = 2^{\lceil \log_2(p)\rceil}$ variables as $x^{A,\beta}$ or shortly as x.

Definition 7.3.5. The *circuit matrix* $C_{X,\mathcal{A},\lambda}(c_{|\lambda^+\cup\{\beta\}}, x_\beta, x)$ is the block diagonal matrix consisting of the blocks

$$\begin{pmatrix} x_{k-1,2i-1} & x_{k,i} \\ x_{k,i} & x_{k-1,2i} \end{pmatrix} \quad \text{for } k \in \{2, \dots, \lceil \log_2(p)\rceil\},\ i \in [2^{\lceil \log_2(p)\rceil - k}],$$

the two singleton blocks

$$\left(x_{\lceil \log_2(p)\rceil,1} - \left(\prod_{\alpha\in A} (\lambda_\alpha)^{\lambda_\alpha}\right) e^{\sigma_X(-\mathcal{A}\lambda)} x_\beta \right) \text{ and } \left(x_\beta + c_\beta \right), \tag{7.10}$$

as well as $2^{\lceil \log_2(p)\rceil - 1}$ blocks of the form

$$\begin{pmatrix} u & x_{1,l} \\ x_{1,l} & w \end{pmatrix} \quad \text{for } l \in [2^{\lceil \log_2(p)\rceil - 1}] \tag{7.11}$$

where $u, w \in \{c_\alpha : \alpha \in A\} \cup \{\left(\prod_{\alpha\in A} (\lambda_\alpha)^{\lambda_\alpha}\right) e^{\sigma_X(-\mathcal{A}\lambda)} x_\beta\}$ such that c_α appears p_α times for every $\alpha \in A$ and $\left(\prod_{\alpha\in A} (\lambda_\alpha)^{\lambda_\alpha}\right) e^{\sigma_X(-\mathcal{A}\lambda)} x_\beta$ appears $2^{\lceil \log_2(p)\rceil} - p$ times.

Theorem 7.3.6. *The set of coefficients of the cone $C_X(\mathcal{A}, \lambda)$ coincides with the projection of the spectrahedron*

$$\widehat{C_X(\mathcal{A}, \lambda)} := \left\{ (c, x) \in \mathbb{R}^{\mathcal{A}} \times \mathbb{R}^{2^{\lceil \log_2(p)\rceil}} : C_{X,\mathcal{A},\lambda}(c_{|\lambda^+\cup\{\beta\}}, x_\beta, x) \succcurlyeq 0,\ c_{|\mathcal{A}\setminus \mathrm{supp}(\lambda)} = 0 \right\}$$

on $(c_{|\lambda^+}, c_\beta)$. The cone $C_X(\mathcal{A}, \lambda)$ is second-order representable.

Again, the proof can be deduced from the one in the unconstrained case and the fact that X is polyhedral.

Corollary 7.3.7. *The matrix $C_{X,\mathcal{A},\lambda}(c_{|\,\mathrm{supp}(\lambda)}, x_\beta, x)$ consists of $2^{\lceil \log_2(p)\rceil} - 1$ blocks of size 2×2 and two blocks of size 1×1.*

We have obtained second-order representations of the λ-witnessed X-AGE-cones under the condition that the sublinear circuit λ is rational and X is polyhedral. We now assume that $\mathcal{A}^T X$ is rational, which — as already pointed out in the proof of Corollary 6.2.5 — immediately implies $\Lambda_X(\mathcal{A}) \subseteq \mathbb{Q}^{\mathcal{A}}$. We call an X-SAGE-cone with

$\mathcal{A}^T X$ rational the *rational* X-SAGE-cone with support in $\mathcal{A}$, and derive an explicit second-order representation of the rational X-SAGE-cone $C_X(\mathcal{A})$ and its dual.

We cannot follow the unconstrained approach as easily as we did in the case of the X-AGE-subcones. This is due to the fact that whereas there does exist a computable purely combinatorial criterion for an X-circuit λ to be *reduced* for $X = \mathbb{R}^n$, such a thing is not known for general constraint sets X — and not even for polyhedric sets X. Nevertheless, we are able to obtain a second-order representation — just without the advantage of reducing the size of the program using the set of reduced X-circuits.

As we only consider exponential sums and not generalized polynomial functions as in Chapter 5, we do not need to introduce additional variables for covering the *odd case*. Hence, we consider the dual variables

$$y^{\mathcal{A}} = \left\{(y^{\lambda}) \,:\, \lambda \in \Lambda_X(\mathcal{A})\right\},$$

consisting of $\sum_{\lambda \in \Lambda_X(\mathcal{A})}(2^{\lceil \log_2(p_\lambda)\rceil} - 1)$ components, where p_λ denotes the maximum of 2 and the smallest common denominator of λ We similarly consider the primal variables

$$x^{\mathcal{A}} = \left\{(x^{\lambda}) \,:\, \lambda \in \Lambda_X(\mathcal{A})\right\},$$

also consisting of $\sum_{\lambda \in \Lambda_X(\mathcal{A})}\left(2^{\lceil \log_2(p_{A,\beta})\rceil} - 1\right)$ components.

We define

$$\widehat{C}^*(\mathcal{A}) = \left\{(v, y^{\mathcal{A}}) \;: C^*_{X,\mathcal{A},\lambda}(v_{|\lambda^+}, v_\beta, y) \succcurlyeq 0 \text{ for all } \lambda \in \Lambda_X(\mathcal{A})\right\}$$

where the variable vector v lives in the space $\mathbb{R}^{\mathcal{A}}$ and

$$\widehat{C}(\mathcal{A}) = \sum_{\lambda \in \Lambda_X(\mathcal{A})} \widehat{C_X(\mathcal{A}, \lambda)}.$$

Here, for every $\lambda \in \Lambda_X(\mathcal{A}, \beta)$, $\widehat{C_X(\mathcal{A}, \lambda)}$ is the set from Theorem 7.3.6.

Corollary 7.3.8.

(1) *The dual of the rational X-SAGE-cone $C^*_X(\mathcal{A})$ is the projection of $\widehat{C}^*(\mathcal{A})$ on the coordinates $v \in \mathbb{R}^{\mathcal{A}}$.*

(2) *The primal rational X-SAGE-cone $C_X(\mathcal{A})$ is the projection of $\widehat{C}(\mathcal{A})$ on the coordinates $v \in \mathbb{R}^{\mathcal{A}}$.*

We apply this to the second-order representations from Theorems 7.3.6 and 7.3.3.

Corollary 7.3.9 (Second-Order Representation of the Dual Rational X-SAGE-Cone)**.** *Let X be polyhedral and $\mathcal{A}^T X$ rational. A vector $v \in \mathbb{R}^{\mathcal{A}}$ is contained in the rational dual X-SAGE-cone $(C^*_X(\mathcal{A}))$ if and only if the circuit vector $y^{\mathcal{A}}$ satisfies the inequalities of Theorem 7.3.3 for every $\lambda \in \Lambda_X(\mathcal{A})$.*

Corollary 7.3.10 (A second-order representation of the rational X-SAGE-cone)**.** *Let X be polyhedral and $\mathcal{A}^T X$ rational. A function $f \in \mathbb{R}^{\mathcal{A}}$ with coefficient vector c is contained in the rational X-SAGE-cone $C_X(\mathcal{A})$ if and only if there exists c^λ for $\lambda \in \Lambda_X(\mathcal{A})$ with $c = \sum_{\lambda \in \Lambda_X(\mathcal{A})} c^\lambda$ and for the circuit vector $x^{\mathcal{A}}$ and the conditions of Theorem 7.3.6 hold for every $\lambda \in \Lambda_X(\mathcal{A})$.*

We demonstrate the results of this subsection using a modified version of Example 5.3.13.

Example 7.3.11. Let $\mathcal{A} = \{0, 1, 2\}$ and $X = \mathbb{R}_+$. As the (normalized) reduced circuits in the case $X = \mathbb{R}_+$ are completely determined by Theorem 6.3.12, namely, they either have the form $(-1, 1, 0)$ or $(1/2, -1, 1/2)$, we can even sharpen our result for this particular example:

In the first case, as the sublinear circuit is supported on exactly two elements, we set $p = 2 = p_0$. According to Remark 7.3.2, we obtain $v_\alpha \geq y_{1,1} \geq v_\beta e^{-\sigma_X(-\mathcal{A}\lambda)} \geq 0$, which simplifies to $v_\alpha \geq v_\beta \geq 0$, as for any $\mathbb{R}_+$-circuit, the support function evaluates to 0.

With the same argumentation, the second case is completely analogous to the unconstrained case.

Hence, a vector (v_0, v_1, v_2) is contained in $C_X(\mathcal{A})^*$ if and only if $v_\alpha \geq 0$ for all $\alpha \in \{0, 1, 2\}$, $v_1 \leq v_0$, and there exists some $y \in \mathbb{R}$ such that $y \geq v_1$ and

$$\begin{pmatrix} v_0 & y \\ y & v_2 \end{pmatrix} \succcurlyeq 0. \tag{7.12}$$

Chapter 8

Résumé and Open Problems

8.1 The SAGE-Cone, Extremality, and its Duality Theory

In Chapter 3, we have provided characterizations of the dual cone of SAGE exponentials and presented several new and several improved results associated with the dual viewpoint. Prominently, we have shown a complete characterization of the extreme rays of the SAGE-cone.

It remains a future task to further understand the relation of the SAGE-cone and its specializations to the underlying class of all nonnegative polynomials or exponential sums, both from the primal and the dual point of view. Specifically, the relation of the cone of SONC polynomials to the cone of sparse nonnegative polynomials and of the dual SONC-cone to sparse moment cones (as studied by Nie [Nie14]) deserve further study. It is an open question whether SONC polynomials are dense inside the nonnegative ones.

In Chapter 4, we have studied two optimization approximations using the SAGE-cone and its dual.

In the case of symmetric exponential sums, we showed that our orbit reduction allows for substantial computational gains both theoretically and practically. This motivates a theoretical study of the strength of the AM/GM bounds in this framework. In particular, it encourages the comparison of the symmetric SAGE-cone with the cone of symmetric nonnegative exponential sums.

Moreover, building upon the projection-free descriptions of the dual cones from Chapter 3, we examined further computational aspects. The results presented in this section provide an effective algorithm for optimizing over a cone induced by the dual SAGE-cone. Recall that this resulting cone is a proper subset of the corresponding primal cone; see Subsection 4.3.1. In particular, we get a bound using an algorithm which is independent of the existing primal SONC, SAGE, and SOS algorithms.

A possibility for future research would be to investigate the polyhedron we discovered in Proposition 4.3.4. From duality theory, we know that there has to exist a primal polyhedron as well. The primal SAGE-cone itself is, however, not polyhedral; see, e.g., [FW19]. Hence, it might be interesting to examine the relation of this primal polyhedron to the SAGE-cone.

8.2 The $\mathcal{S}$-Cone and Second-Order Representations

In Chapter 5, we have introduced the $\mathcal{S}$-cone as a unified framework for the classes of SAGE exponentials and SONC polynomials; thus, it exhibits a prominent computationally tractable class within the class of sparse nonnegative polynomials.

We have provided second-order representations for primal and dual rational $\mathcal{S}$-cones. These statements also remain valid for non-rational sets $\mathcal{A}$ as long as all the relevant barycentric coordinates are still rational. It is an open question whether an $\mathcal{S}$-cone and its dual are also second-order representable in the general non-rational case. Also, despite the use of reduced circuits, the second-order representation of the $\mathcal{S}$-cone is still rather large. The question remains whether smaller second-order representations for the $\mathcal{S}$-cone exist.

8.3 The Conditional SAGE-Cone and Sublinear Circuits

In the third part of the thesis, we examined the conditional SAGE-cone. Building upon results by Murray, Chandrasekaran, and Wierman, we introduced the notion of sublinear circuits as objects inducing the X-SAGE-cone in Chapter 6. We showed a variety of properties of these X-circuits including affine independency of the positive support and that the X-SAGE-cone can indeed be constructed using these X-circuits as the only witnesses. We constructed several results to determine whether a vector $\nu \in N_\beta$ is an X-circuit, which are either necessary or sufficient.

For extremality of the set of X-circuits and, in particular, the conditional SAGE-cone, we introduced the concept of reducedness based on the notion of the circuit graph. Here, we were also able to provide either necessary or sufficient results for this characterization.

Throughout this part of the thesis, we put a particular emphasis on the case of polyhedral constraint sets. We have studied the connection of sublinear circuits and their supports as well as the sublinear circuits for polyhedral sets X. Since the number of X-circuits is finite for polyhedral sets, this allows to apply polyhedral and combinatorial techniques. In particular, the X-SAGE-cones can be decomposed into a finite number of power cones, which arise from the reduced sublinear circuits.

In the last chapter of this thesis, we have explored optimization and representation approaches for the conditional SAGE-cone. We have shown that for symmetric constraint sets, we can exploit existing symmetries in exponential sums to reduce the size of the relative entropy program which certifies containment in the X-SAGE-cone (and also serves as an approximation of the constrained optimization problem).

Following the approach in the first part of the thesis, we then examined a variation of the dual X-SAGE-cone in Chapter 7 and presented a linear program for polyhedral conic constraint sets X to approximate the constrained optimization problem.

Finally, we presented a second-order representation of the X-SAGE-cone and its dual as well as of the subcones of AGE exponentials (and their respective duals). Those representations work for polyhedral constraint sets and under the assumption that $\mathcal{A}^T X$ is rational. The former is due to the fact that, only for polyhedral constraint sets X, we could prove that the number of X-circuits is finite, the latter restriction ensures that the set of X-circuits is contained in $\mathbb{Q}^{\mathcal{A}}$. This allows us to define the second-order-cone programs in terms of the coordinates of the sublinear circuits.

The obvious pending line of further research is an advanced combinatorial understanding of X-circuits and, in particular, reduced X-circuits. Although we provided necessary and sufficient criteria for being a (reduced) X-circuit, we do not have a criterion that is *both* necessary and sufficient — not even for the case of polyhedral X. These results would be particularly useful for practical optimization techniques as they significantly reduce the number of elements we have to consider in any X-SAGE decomposition.

Another line of further research are non-polyhedral sets X, where in general the number of X-circuits is not finite anymore. It remains a future task to study necessary and sufficient criteria for sublinear circuits of structured non-polyhedral sets, such as sets with symmetry. In a different direction, Forsgård and de Wolff have characterized the boundary of the SAGE-cone through a connection between circuits and tropical geometry [FW19]. It also remains for future work to establish a generalization of this work, aiming at connecting the conditional SAGE-cone and sublinear circuits to tropical geometry.

All three presented optimization and representation methods using the X-SAGE-cone (or its dual) have severe restrictions on the sets X they work for. It remains of further interest to understand what happens if we have constraint sets X that do not satisfy these restrictions and whether we can derive other computationally tractable optimization approaches then.

Bibliography

[ADB19] A. Agrawal, S. Diamond, and S. Boyd. “Disciplined Geometric Programming”. In: *Optimization Letters* 13 (5 2019), pp. 961–976.

[Agr+18] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd. “A Rewriting System for Convex Optimization Problems”. In: *Journal of Control and Decision* 5.1 (2018), pp. 42–60.

[AM16] A. Ahmadi and A. Majumdar. “Some applications of polynomial optimization in operations research and real-time decision making”. In: *Optimization Letters* 10 (2016), pp. 709–729.

[AM19] A. A. Ahmadi and A. Majumdar. “DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization”. In: *SIAM J. Appl. Algebra Geom.* 3.2 (2019), pp. 193–230.

[Ave19] G. Averkov. “Optimal size of linear matrix inequalities in semidefinite approaches to polynomial optimization”. In: *SIAM J. Appl. Algebra and Geometry* 3.1 (2019), pp. 128–151.

[Bac+12] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin. “Invariant semidefinite programs”. In: *Handbook on Semidefinite, Conic and Polynomial Optimization.* Springer, 2012, pp. 219–269.

[BN01] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization.* SIAM, Philadelphia, 2001.

[Bos20] S. Bosch. *Algebra.* Springer Berlin Heidelberg, 2020.

[Boy+07] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi. “A tutorial on geometric programming”. In: *Optim. Eng.* 8.1 (2007), pp. 67–127.

[BPR06] S. Basu, R. Pollack, and M-F. Roy. *Algorithms in Real Algebraic Geometry.* Springer-Verlag Berlin Heidelberg, 2006.

[BR21] G. Blekherman and C. Riener. “Symmetric non-negative forms and sums of squares”. In: *Discrete & Computational Geometry* 65.3 (2021), pp. 764–799.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization.* MOS-SIAM Series on Optimization. Cambridge University Press, 2004.

[BV08] C. Bachoc and F. Vallentin. “New upper bounds for kissing numbers from semidefinite programming”. In: *J. Amer. Math. Soc.* 21.3 (2008), pp. 909–924.

[CS16] V. Chandrasekaran and P. Shah. “Relative entropy relaxations for signomial optimization”. In: *SIAM J. Optim.* 26.2 (2016), pp. 1147–1173.

[CS17] V. Chandrasekaran and P. Shah. “Relative entropy optimization and its applications”. In: *Math. Program., Ser. A* 161.1-2 (2017), pp. 1–32.

[DA21] J. Dahl and E. Andersen. *A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization.* 2021.

[DB16] S. Diamond and S. Boyd. “CVXPY: A Python-Embedded Modeling Language for Convex Optimization”. In: *Journal of Machine Learning Research* 17.83 (2016), pp. 1–5.

[DCB13] A. Domahidi, E. Chu, and S. Boyd. “ECOS: An SOCP solver for embedded systems”. In: *European Control Conference (ECC)*. 2013, pp. 3071–3076.

[DG08] P.J.C. Dickinson and L. Gijben. “On the computational complexity of membership problems for the completely positive cone and its dual”. In: *Comput Optim Appl* 57.2 (2008), pp. 403–415.

[DIW17] M. Dressler, S. Iliman, and T. de Wolff. “A Positivstellensatz for sums of nonnegative circuit polynomials”. In: *SIAM J. Appl. Algebra and Geometry* 1.1 (2017), pp. 536–555.

[DIW19] M. Dressler, S. Iliman, and T. de Wolff. “An approach to constrained polynomial optimization via nonnegative circuit polynomials and geometric programming”. In: *J. Symb. Comp.* 91 (2019), pp. 149–172.

[DKW21] M. Dressler, A. Kurpisz, and T. de Wolff. *Optimization over the Boolean hypercube via sums of nonnegative circuit polynomials*. 2021.

[DNT21] M. Dressler, H. Naumann, and T. Theobald. “The dual cone of sums of non-negative circuit polynomials”. In: *Advances in Geometry* 21 (2021), pp. 227–236.

[DP15] P. J. C. Dickinson and J. Povh. “On an extension of Pólya’s Positivstellensatz”. In: *Journal of Global Optimization* 61 (2015), pp. 615–625.

[DR20] S. Debus and C. Riener. *Reflection groups and cones of sums of squares*. Preprint, arXiv:2011.09997. 2020.

[Dre+20] M. Dressler, J. Heuer, H. Naumann, and T. de Wolff. “Global optimization via the dual SONC cone and linear programming”. In: *Proc. 45th International Symposium on Symbolic and Algebraic Computation*. 2020, pp. 138–145.

[Dre18] M. Dressler. “Sums of nonnegative circuit polynomials : geometry and optimization”. PhD thesis. Goethe-Universität Frankfurt/Main, 2018. URL: `http://publikationen.ub.uni-frankfurt.de/files/46971/Thesis_Mareike_Dressler.pdf`.

[DV15] C. Dobre and J. Vera. “Exploiting symmetry in copositive programs via semidefinite hierarchies”. In: *Math. Program.* 151.2 (2015), pp. 659–680.

[EPR20] A. A Ergür, G. Paouris, and J. M. Rojas. “Tropical varieties for exponential sums”. In: *Math. Annalen* 377 (2020), pp. 863–882.

[Fel+20] E. Feliu, N. Kaihnsa, O. Yürük, and T. de Wolff. “The kinetic space of multistationarity in dual phosphorylation”. In: *Journal of Dynamics and Differential Equations* (2020).

[FI17] F. Fillastre and I. Izmestiev. “Shapes of polyhedra, mixed volumes and hyperbolic geometry”. In: *Mathematika* 63.1 (2017), pp. 124–183.

[FRS18] T. Friedl, C. Riener, and R. Sanyal. “Reflection groups, reflection arrangements, and invariant real varieties”. In: *Proc. Amer. Math. Soc.* 146.3 (2018), pp. 1031–1045.

[FW19] J. Forsgård and T. de Wolff. *The algebraic boundary of the SONC cone*. Preprint, arXiv:1905.04776. 2019.

[GKZ94] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants.* Springer, 1994, pp. 271–296.

[GP04] K. Gatermann and P. A Parrilo. "Symmetry groups, semidefinite programs, and sums of squares". In: *J. Pure & Applied Algebra* 192.1-3 (2004), pp. 95–128.

[Hal18] G. Hall. "Optimization over nonnegative and convex polynomials with and without semidefinite programming". PhD thesis. Dept. of Operations Research and Financial Engineering, Princeton University, 2018. URL: `https://scholar.princeton.edu/sites/default/files/ghall/files/georgina_hall_proquest_version.pdf`.

[HG05] D. Henrion and A. Garulli, eds. *Positive Polynomials in Control.* Vol. 312. Lecture Notes in Control and Information Sciences. Springer, 2005.

[HHS21] A. Heaton, S. Hosten, and I. Shankar. "Symmetry Adapted Gram Spectrahedra". In: *SIAM J. Appl. Algebra Geom.* 5 (2021), pp. 140–164.

[HLP52] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities.* 2nd ed. Cambridge, at the University Press, 1952.

[IW16a] S. Iliman and T. de Wolff. "Amoebas, nonnegative polynomials and sums of squares supported on circuits". In: *Res. Math. Sci.* 3.9 (2016).

[IW16b] S. Iliman and T. de Wolff. "Lower Bounds for Polynomials with Simplex Newton Polytopes Based on Geometric Programming". In: *SIAM J. Optim.* 26.2 (2016), pp. 1128–1146.

[Jos16] C. Josz. "Application of Polynomial Optimization to Electricity Transmission Networks". PhD thesis. University Paris VI, 2016. URL: `https://tel.archives-ouvertes.fr/tel-01478431`.

[Kar+17] O. Karaca, G. Darivianakis, P. Beuchat, A. Georghiou, and J. Lygeros. "The REPOP Toolbox: Tackling polynomial optimization using relative entropy relaxations". In: *20th IFAC World Congress, IFAC PapersOnLine.* Vol. 50(1). Elsevier, 2017, pp. 11652–11657.

[KKW05] M. Kojima, S. Kim, and H. Waki. "Sparsity in sums of squares of polynomials". In: *Math. Program.* 103.1, Ser. A (2005), pp. 45–62.

[KNT21] L. Katthän, H. Naumann, and T. Theobald. "A unified framework of SAGE and SONC polynomials and its duality theory". In: *Math. Comput.* 90 (2021), pp. 1297–1322.

[Kri64] J.-L. Krivine. "Anneaux préordonnés". In: *J. Anal. Math.* 12 (1964), pp. 307–326.

[KS10] E. de Klerk and R. Sotirov. "Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem". In: *Math. Program.* 122.2 (2010), p. 225.

[Las00] J.B. Lasserre. "Global optimization with polynomials and the problem of moments". In: *SIAM J. Optim.* 11.3 (2000), pp. 796–817.

[Las10] J.B. Lasserre. *Moments, Positive Polynomials and their Applications.* Imperial College Press, London, 2010.

[Las15] J.B. Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization.* Vol. 1. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2015.

[Lau09] M. Laurent. "Sums of squares, moment matrices and optimization over polynomials". In: *Emerging Applications Of Algebraic Geometry*. Ed. by M. Putinar and S. Sullivant. Vol. 149. IMA Vol. Math. Appl. New York: Springer, 2009, pp. 157–270.

[MCD17] V. Magron, G. Constantinides, and A. Donaldson. "Certified roundoff error bounds using semidefinite programming". In: *ACM Trans. Math. Software* 43.4 (2017), Art. 34, 31.

[MCW21a] R. Murray, V. Chandrasekaran, and A. Wierman. *Newton polytopes and relative entropy optimization. Found. Comput. Math.* 2021.

[MCW21b] R. Murray, V. Chandrasekaran, and A. Wierman. "Signomial and polynomial optimization via relative entropy and partial dualization". In: *Math. Program. Comput.* (2021).

[MHR19] S. Müller, J. Hofbauer, and G. Regensburger. "On the Bijectivity of Families of Exponential/Generalized Polynomial Maps". In: *SIAM J. Appl. Algebra and Geometry* 3.3 (2019), pp. 412–438.

[MK87] K. G. Murty and S. N. Kabadi. "Some NP-complete problems in quadratic and nonlinear programming". In: *Math. Program.* 39.2 (1987), pp. 117–129.

[MNT20] R. Murray, H. Naumann, and T. Theobald. *Sublinear circuits and the constrained signomial nonnegativity problem.* Preprint, arXiv:2006.06811. 2020.

[Mou+21] P. Moustrou, H. Naumann, C. Riener, T. Theobald, and H. Verdure. *Symmetry reduction in AM/GM-based optimization.* Preprint, arXiv:2102.12913. 2021.

[MRV21] P. Moustrou, C. Riener, and H. Verdure. "Symmetric ideals, Specht polynomials and solutions to symmetric systems of equations". In: *J. Symb. Comp.* 107 (2021), pp. 106–121.

[MSW19] V. Magron, H. Seidler, and T. de Wolff. "Exact Optimization via Sums of Nonnegative Circuits and Arithmetic-geometric-mean-exponentials". In: *Proceedings of the 2019 on International Symposium on Symbolic and Algebraic Computation.* ACM, July 2019.

[Mül+15] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. "Sign Conditions for Injectivity of Generalized Polynomial Maps with Applications to Chemical Reaction Networks and Real Algebraic Geometry". In: *Found. Comp. Math.* 16.1 (2015), pp. 69–97.

[Mur20] R. Murray. *Sageopt 0.5.3.* 2020.

[Nie14] J. Nie. "The $\mathcal{A}$-truncated K-moment problem". In: *Found. Comput. Math.* 14.6 (2014), pp. 1243–1276.

[NN94] Y. Nesterov and A. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming.* SIAM, Philadelphia, 1994.

[NPP17] C. Nowzari, V. M. Preciado, and G. J. Pappas. "Optimal Resource Allocation for Control of Networked Epidemic Models". In: *IEEE Transactions on Control of Network Systems* 4.2 (2017), pp. 159–169.

[NT21a] H. Naumann and T. Theobald. *Sublinear circuits for polyhedral constraint sets.* Preprint, arXiv:2103.09102. 2021.

[NT21b] H. Naumann and T. Theobald. "The $\mathcal{S}$-cone and a primal-dual view on second-order representability". In: *Beiträge Algebra Geom. (Special issue on the 50th anniversary of the journal)* 62.1 (2021), pp. 229–249.

[ÖS19] B. Öztürk and A. Saab. "Optimal Aircraft Design Decisions under Uncertainty via Robust Signomial Programming". In: *AIAA Aviation 2019 Forum*. American Institute of Aeronautics and Astronautics, 2019–3351, 2019.

[Pap19] D. Papp. *Duality of sum of nonnegative circuit polynomials and optimal SONC bounds.* Preprint, arXiv:1912.04718. 2019.

[Par00] P.A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD Thesis, California Institute of Technology. 2000.

[PKC12] C. Pantea, H. Koeppl, and G. Craciun. "Global injectivity and multiple equilibria in uni- and bi-molecular reaction networks". In: *Discrete and Continuous Dynamical Systems - Series B* 17.6 (May 2012), pp. 2153–2170.

[Pre+14] V. M. Preciado, M. Zargham, C. Enyioha, A. Jadbabaie, and G. J. Pappas. "Optimal Resource Allocation for Network Protection Against Spreading Processes". In: *IEEE Transactions on Control of Network Systems* 1.1 (2014), pp. 99–108.

[Put93] M. Putinar. "Positive polynomials on compact semi-algebraic sets". In: *Indiana Univ. Math. J.* 42.3 (1993), pp. 969–984.

[Ray+18] A. Raymond, J. Saunderson, M. Singh, and R. R. Thomas. "Symmetric sums of squares over k-subset hypercubes". In: *Math. Program., Ser. A* 167.2 (2018), pp. 315–354.

[Rez89] B. Reznick. "Forms derived from the arithmetic-geometric inequality". In: *Math. Annalen* 283.3 (1989), pp. 431–464.

[Rie+13] C. Riener, T. Theobald, L. Jansson-Andrén, and J. B. Lasserre. "Exploiting symmetries in SDP-relaxations for polynomial optimization". In: *Math. Oper. Res.* 38.1 (2013), pp. 122–141.

[Rie12] C. Riener. "On the degree and half-degree principle for symmetric polynomials". In: *J. Pure & Applied Algebra* 216.4 (2012), pp. 850–856.

[Roc97] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1997.

[Sch14] R. Schneider. *Convex Bodies: the Brunn–Minkowski Theory.* Cambridge University Press, 2014.

[Sd18] H. Seidler and T. de Wolff. *An Experimental Comparison of SONC and SOS Certificates for Unconstrained Optimization.* Preprint, arXiv:1808.08431. 2018.

[Sd19] H. Seidler and T. de Wolff. *POEM: Effective Methods in Polynomial Optimization, version 0.2.1.0.* `http://www.iaa.tu-bs.de/AppliedAlgebra/POEM/`. 2019.

[Sla14] M. Slater. "Lagrange Multipliers Revisited". In: *Giorgi G., Kjeldsen T. (eds) Traces and Emergence of Nonlinear Programming* (2014).

[Sta99] R. P. Stanley. *Enumerative Combinatorics.* Vol. 2. Cambridge University Press, 1999.

[Ste74] G. Stengle. “A Nullstellensatz and a Positivstellensatz in semialgebraic geometry”. In: *Mathematische Annalen* 207.2 (1974), pp. 87–97.

[Stu96] B. Sturmfels. *Gröbner Bases and Convex Polytopes.* Amer. Math. Soc., 1996.

[Tim03] V. Timofte. “On the positivity of symmetric polynomial functions.: Part i: General results”. In: *J. Math. Analysis and Applications* 284.1 (2003), pp. 174–190.

[Wan+20] A. H. Wang, P. Jaini, Y. Yu, and P. Poupart. *A Positivstellensatz for conditional SAGE signomials.* Preprint, arXiv:2003.03731. 2020.

[Wan18a] J. Wang. *Nonnegative polynomials and circuit polynomials.* Preprint, arXiv:1804.09455. 2018.

[Wan18b] J. Wang. *On supports of sums of nonnegative circuit polynomials.* Preprint, arXiv:1809.10608. 2018.

[Wat83] W. C. Waterhouse. “Do symmetric problems have symmetric solutions?” In: *Amer. Math. Monthly* 90.6 (1983), pp. 378–387.

[WLT18] T. Weisser, J. B. Lasserre, and K.-C. Toh. “Sparse-BSOS: a bounded degree SOS hierarchy for large scale polynomial optimization with sparsity”. In: *Math. Program. Comput.* 10.1 (2018), pp. 1–32.

[WM20a] J. Wang and V. Magron. “A second order cone characterization for sums of nonnegative circuits”. In: *Proc. 45th International Symposium on Symbolic and Algebraic Computation.* 2020, pp. 450–457.

[WM20b] J. Wang and V. Magron. “A second order cone characterization for sums of nonnegative circuits”. In: *Proc. 45th International Symposium on Symbolic and Algebraic Computation.* 2020, pp. 450–457.

[WML21a] J. Wang, V. Magron, and J. Lasserre. “Chordal-TSSOS: A Moment-SOS Hierarchy That Exploits Term Sparsity with Chordal Extension”. In: *SIAM J. Optim.* 31 (2021), pp. 114–141.

[WML21b] J. Wang, V. Magron, and J. Lasserre. “TSSOS: A Moment-SOS Hierarchy That Exploits Term Sparsity”. In: *SIAM J. Optim.* 31 (2021), pp. 30–58.

[YHD18] M. York, W. Hoburg, and M. Drela. “Turbofan Engine Sizing and Tradeoff Analysis via Signomial Programming”. In: *Journal of Aircraft* 55.3 (2018), pp. 988–1003.

[Zie95] G. M. Ziegler. *Lectures on Polytopes.* Springer, New York, 1995.

Appendix A

Deutsche Zusammenfassung

Globale Optimierung und Optimierung unter Nebenbedingungen untersucht die Frage, welchen Minimalwert eine gegebene reelle Funktion $f : \mathbb{R}^n \to \mathbb{R}$ entweder über $\mathbb{R}^n$ oder über einer Teilmenge $X \subsetneq \mathbb{R}^n$ annehmen kann. Dieses Problem tritt in vielen Zweigen der Mathematik und deren Anwendungsgebieten auf. Die ähnliche Frage, ob eine reelle Funktion nur nichtnegative Werte annimmt, ist eine grundlegende Frage in der reell-algebraischen Geometrie. Beide Probleme können als äquivalent behandelt werden: Das Infimum einer Funktion f ist der größte reelle Wert λ, sodass die Funktion $f - \lambda$, die man durch Subtraktion des Skalars von f erhält, global nichtnegativ ist, d.h,

$$f^* = \inf\{f(x) \ : \ x \in \mathbb{R}^n\} \ = \ \sup\{\lambda \in \mathbb{R} \ : \ f - \lambda \text{ ist nichtnegativ auf } \mathbb{R}^n\}. \quad \text{(A.1)}$$

In dieser Arbeit werden vor allem Exponentialsummen $f = \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle}$ und ggf. Polynome betrachtet, die als Spezialfall von Exponentialsummen angesehen werden können: Für $\mathcal{A} \subseteq \mathbb{N}^n$ ergibt die Substitution $x_i = \ln y_i$ Polynomfunktionen $y \mapsto \sum_{\alpha \in \mathcal{A}} c_\alpha y^\alpha$ auf dem positiven Orthanten $\mathbb{R}^n_{>0}$. Nichtnegative Polynome oder Exponentialsummen und Optimierung über beide Funktionen sind allgegenwärtig in Anwendungen, und Dünnbesetztheit ist eine der zentralen Struktureigenschaften, die Potenzial für effizient berechenbare Approximationen bieten. Neben klassischen Anwendungen in Kontrolltheorie und Robotik (siehe z.B., [HG05], [AM19]), treten neuere Anwendungen von nichtnegativen Polynomen und polynomieller Optimierung in dem Problem der Lastflussoptimierung [Jos16], Kollisionsvermeidung [AM16], Regressionsproblemen mit formbeschreibenden Nebenbedingungen [Hal18], chemischen Reaktionsnetzwerken [Mül+15; MHR19], Optimierung von Flugzeugdesigns [ÖS19; YHD18], ebenso wie in epidemiologischer Prozesskontrolle [NPP17; Pre+14] auf; siehe auch [EPR20] und die dortigen Verweise.

Die Berechnung des Minimalwerts f^* aus dem Einführungsproblem sowie die Entscheidung über die Nichtnegativität einer gegebenen reellen Funktion ist NP-schwer [MK87], selbst im Fall polynomieller Optimierung [Lau09]. Daher besteht die Idee darin, nach effizient berechenbaren hinreichenden Bedingungen für die Nichtnegativität zu suchen – sogenannten *Nichtnegativitätszertifikaten*. Im Idealfall erfüllt eine große Teilmenge der Elemente im Nichtnegativitätskegel diese Bedingungen. Eine bekannte und große Teilmenge nichtnegativer Polynome sind *SOS-Polynome* – Summen von Quadraten anderer Polynome. Sie bieten eine Relaxation für die Suche nach dem Minimalwert eines gegebenen Polynoms. Die Nichtnegativität *dieser* Polynome kann mittels semidefiniter Programmierung verifiziert werden [Las00; Par00]; für einen breiteren Überblick über dieses Thema siehe Kapitel 2. Es stellt sich jedoch das Problem, dass die Elemente der Zerlegung einen möglicherweise unbeschränkten Grad haben können, der eine möglicherweise unendliche Größe des semidefiniten Programms verursacht.

In jüngster Vergangenheit haben mehrere Forscher hinreichende Bedingungen für Nichtnegativität basierend auf der arithmetisch-geometrischen (AM/GM) Ungleichung entwickelt. Während SOS-Nichtnegativitätszertifikate mit dem Grad eines Polynoms arbeiten und insbesondere ganzzahlige Exponenten erfordern, ist das bei diesen auf der AM/GM-Ungleichung basierenden Verfahren nicht der Fall. Daher funktionieren sie auch für Exponentialsummen. Aufbauend auf früheren Ergebnissen von Reznick [Rez89] wurden die Verfahren in den Arbeiten von Pantea, Koeppl und Craciun [PKC12], Iliman und de Wolff [IW16a], und Chandrasekaran und Shah [CS16] weiterentwickelt. Im Gegensatz zu SOS-basierten Zertifikaten erhalten diese AM/GM-basierten Nichtnegativitätszertifikate die Dünnbesetztheit einer gegebenen Exponentialsumme oder eines Polynoms [MCW21a; Wan18a].

Formal ist eine Exponentialsumme (oder ein *Signom*) mit endlicher Trägermenge $\mathcal{A} \subseteq \mathbb{R}^n$ definiert als Summe $\sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, x \rangle}$ mit reellen Koeffizienten c_α. Betrachtet man beispielsweise eine Menge von Trägerpunkten $A \subseteq \mathbb{R}^n, \beta \in \mathbb{R}^n$ und Koeffizienten $\lambda \in \mathbb{R}_+^A$ mit $\mathbf{1}^T \lambda = 1$ und $\sum_{\alpha \in A} \lambda_\alpha \alpha = \beta$, dann ergibt die gewichtete AM/GM-Ungleichung

$$\sum_{\alpha \in A} \lambda_\alpha e^{\langle \alpha, x \rangle} \geqslant \prod_{\alpha \in A} e^{\lambda_\alpha \langle \alpha, x \rangle} = e^{\langle \beta, x \rangle}.$$

Demzufolge kann die Nichtnegativität auf $\mathbb{R}^n$ von Exponentialsummen der Form

$$\sum_{\alpha \in A} \lambda_\alpha e^{\langle \alpha, x \rangle} - e^{\langle \beta, x \rangle}$$

mittels der AM/GM-Ungleichung verifiziert werden.

Das für diese Arbeit relevante Rahmenwerk wurde unter verschiedenen Namen eingeführt: In 2016 führten Iliman und de Wolff das Konzept der *Summen nichtnegativer Kreispolynome* (SONC) ein [IW16a]. Ein *Kreis* ist ein Tupel (A, β) mit affin unabhängiger Menge $A \subseteq \mathbb{R}^n$ und einem Element $\beta \in \operatorname{relint} \operatorname{conv}(A)$. Ein *Kreispolynom* ist ein Polynom mit Exponenten in $A \cup \{\beta\}$, so dass die Koeffizienten, die zu A korrespondieren, nichtnegativ sind. Ein Polynom, dessen Träger ein Kreis ist, kann nur dann *global nichtnegativ* sein, wenn $A \subseteq (2\mathbb{N})^n$; für einen Beweis siehe z.B. [Fel+20].

Später im selben Jahr führten Chandrasekaran und Shah das Konzept der *Summen arithmetisch-geometrischer Exponentiale* (SAGE) ein [CS16]. Ein *arithmetisch-geometrisches Exponential (AGE-Exponential)* ist eine Exponentialsumme mit höchstens einem negativen Koeffizienten. Diese Definition umfasst genau die zuvor erwähnte Unterklasse von Exponentialsummen, deren Nichtnegativität durch die AM/GM-Ungleichung verifiziert werden kann (was den Namen dieser Klasse von Exponentialsummen erklärt).

Der SONC-Ansatz wurde in verschiedenen Arbeiten zur Relaxation globaler Optimierungsprobleme verwendet, d.h. die Nichtnegativitätsbedingung wurde durch eine SONC-Bedingung

$$f^{\text{sonc}} = \sup\{\lambda \in \mathbb{R} \, : \, f - \lambda \text{ ist SONC}\} \leq \, \inf\{f(x) \, : \, x \in \mathbb{R}^n\} \tag{A.2}$$

ersetzt. Iliman, de Wolff und Dressler et al. untersuchten Nichtnegativitätszertifikate, die zu einem geometrischen Programm für bestimmte Unterklassen von SONC-Polynomen führen [IW16a; IW16b; DIW17; DKW21]. Karaca et al. kombinierten den SONC-Ansatz mit dem Sums-of-Squares-Ansatz zur polynomiellen Optimierung auf dem nichtnegativen Orthanten [Kar+17]. In einem verwandten Rahmen konzentrierten sich Chandrasekaran, Shah, Murray und Wierman auf Ansätze, die auf ein relatives Entropie-Programm sowohl zur Optimierung von Exponentialsummen als auch von Polynomen führen [CS16; CS17; MCW21a]. Diese relativen Entropie-Programme haben den großen Vorteil, dass sie effizient und für allgemeine Klassen von Exponentialsummen berechnet werden können. Seidler und de Wolff stellten einen Algorithmus zur Berechnung von Kreiszerlegungen vor [Sd18], und zusammen mit Magron verglichen sie die bestehenden Ansätze SONC-, SAGE- und SOS-basierter Relaxationen globaler Optimierungsprobleme [MSW19]. Unter Verwendung ähnlicher Ansätze wie sie in Kapitel 3 dieser Arbeit diskutiert werden entwickelte Papp einen Algorithmus zur Berechnung der optimalen Kreisszerlegung eines gegebenen Polynoms [Pap19]. Viele dieser Optimierungsansätze sind in [Sd19] implementiert.

In 2019 stellten Forsgård und de Wolff eine neue Sprache zur Untersuchung von Exponentialsummen mit Kreisen als Trägermenge vor. In dieser neuen Sprache untersuchten sie den algebraischen Rand des SONC-Kegels (für Exponentialsummen, [FW19]) unter Verwendung der Theorie regulärer Unterteilungen, A-Diskriminanten und tropischer Geometrie.

Neben den beiden zuvor besprochenen Optimierungsansätzen gibt es auch einen Bezug zur semidefiniten Programmierung. Ähnlich wie in [Kar+17] kombinierte Averkov SONC-Ansätze mit semidefiniten Ansätzen zur polynomiellen Optimierung. Er zeigte, dass der Kegel der SONC-Polynome als Projektion eines Spektraeders dargestellt werden kann [Ave19]. Wang und Magron lieferten später eine explizite Darstellung des primalen SONC-Kegels als Projektion eines Kegels zweiter Ordnung [WM20a].

Bereits 2016 zeigten Iliman und de Wolff, dass die globale Nichtnegativität eines Kreispolynoms auf die Nichtnegativität einer Exponentialsumme mit derselben Trägermenge und möglicherweise leicht modifizierten Koeffizienten reduziert werden kann [IW16a]. In 2018 bewiesen Murray, Chandrasekaran und Wierman, dass dies für allgemeine Polynome mit höchstens einem negativen Term (d. h., höchstens einem Term $c_\beta x^\beta$ mit $c_\beta < 0$ oder $\beta \notin (2\mathbb{N})^n$) gilt [MCW21a]. Wang bemerkte schon zuvor im selben Jahr, dass die polynomielle Form eines AGE-Exponentials ein SONC-Polynom ist [Wan18a], und Murray, Chandrasekaran und Wierman führten einen unabhängigen Beweis für die explizite Aussage, dass die Trägermengen von Extremalstrahlen des Kegels der AGE-Exponentiale Kreise sind (wieder [MCW21a]). Beide zeigten somit die Äquivalenz von SONC und SAGE in Bezug auf die gewählte Sprache der Polynome oder Exponentialsummen.

Daher ist es naheliegend, Ergebnisse, die sowohl für SONC als auch für SAGE relevant sind, nur für die einfachere Sprache der Exponentialsummen und unter einem gemeinsamen Namen zu diskutieren. Aufbauend auf dem von Chandrasekaran und Shah eingeführten Ansatz für Exponentialsummen wird in dieser Arbeit der Name *AGE-Eponentiale* für Exponentialsummen mit höchstens einem negativen Term und *SAGE-Exponentiale* für Summen dieser AGE-Exponentiale verwendet. Darüber hinaus werden, in Anlehnung an die von Iliman und de Wolff eingeführte Terminologie, AGE-Exponentiale, deren Träger einen Kreis bildet und bei dem der möglicherweise negative Term dem inneren Exponenten des Kreises entspricht, *Kreisexponentiale* genannt. Diese Funktionen sind von besonderem Interesse, weil – wie oben erwähnt – die Träger aller Extremalstrahlen des SAGE-Kegels Kreise sind. In dieser Arbeit wird ein besonderer Fokus auf Kreise gelegt: Zuerst wird die Dualitätstheorie untersucht, dann die Extremalstrahlen des SAGE-Kegels, verschiedene Optimierungsansätze und schließlich all diese Themen für die Optimierung unter Nebenbedingungen.

Für das Problem der Optimierung unter Nebenbedingungen haben die SONC- und SAGE-Ansätze einen gravierenden Nachteil gegenüber dem SOS-Ansatz, nämlich, dass es keinen Putinar-ähnlichen Positivstellensatz gibt, der ein Polynom in Summen von Quadraten und gegebene Nebenbedingungen zerlegt, sofern es nichtnegativ auf der gewählten Menge der Nebenbedingungen ist [DKW21]. Dressler et. al lieferten in der gleichen Arbeit einen anderen Positivstellensatz für SONC-Polynome, der aber später als ein Spezialfall des Krivine-Positivstellensatzes identifiziert wurde.

In 2019 untersuchten Murray, Chandrasekaran und Wierman einen anderen Ansatz zur Optimierung unter Nebenbedingungen, sowohl für Exponentialsummen als auch für Polynome [MCW21b], der wiederum zu einem Nichtnegativitätszertifikat unter Verwendung relativer Entropieprogrammierung führt. Unter Verwendung der Lagrangedualität kommen sie – ähnlich wie im Fall globaler Optimierungsprobleme – zu einem relativen Entropie-Programm. Die zugrundeliegende Menge der Nebenbedingungen wird durch eine sogenannte Stützfunktion repräsentiert. Mit diesem relativen Entropie-Programm lässt sich eine Hierarchie finden, um den optimalen Wert der gegebenen Funktion zu approximieren. Wang et. al konnten in der Tat zeigen, dass diese Hierarchie vollständig ist [Wan+20] (siehe auch [DP15]). Das implementierte Programm ist in [Mur20] zu finden, unter Verwendung des Solvers MOSEK [DA21].

Als weitere verwandte Arbeit sei hier die Ausnutzung von Dünnbesetztheit und Symmetrien zur Ableitung spezifischer SDP-Relaxationen für polynomielle Optimierung [KKW05; MCD17; Rie+13; WML21b; WML21a; WLT18] genannt.

In dieser Arbeit wird der SAGE-Kegel untersucht, seine Geometrie und Verallgemeinerungen davon. Die Arbeit besteht aus drei Hauptteilen:

1. Der erste Teil konzentriert sich auf den Kegel global nichtnegativer Exponentialsummen mit höchstens einem negativen Term. Insbesondere werden Dualitätstheorie, Extremalstrahlen des Kegels und zwei effiziente Optimierungsansätze über den SAGE-Kegel und seinen dualen Kegel untersucht.

2. Im zweiten Teil wird der sogenannte $\mathcal{S}$-Kegel eingeführt. Dieser Kegel bietet ein einheitliches Rahmenwerk für SAGE-Exponentiale und SONC-Polynome. Insbesondere konzentriert sich diese Arbeit auf Darstellungen zweiter Ordnung des $\mathcal{S}$-Kegels und seines Dualen unter Verwendung von Extremalitätsergebnissen aus Teil 1.

3. Der dritte und letzte Teil dieser Arbeit wendet sich der Untersuchung des *bedingten* SAGE-Kegels zu. Mittels *sublinearer Kreise* werden neue Dualitätsergebnisse und eine partielle Charakterisierung der Extremalität erarbeitet. Im Fall polyedrischer Nebenbedingungen vereinfacht sich diese Untersuchung und erlaubt es, sublineare Kreise und Extremalität für einige Fälle vollständig zu klassifizieren. Für Zulässigkeitsbereiche mit bestimmten Bedingungen, wie z. B. Mengen mit Symmetrien, konische oder polyedrische Mengen, können verschiedene Optimierungs- und Darstellungsergebnisse des unrestringierten auf den restringierten Kegel angewendet werden.

Extremalitäts- und Dualitätstheorie des SAGE-Kegels

Während viele Aspekte der Klassen der SAGE-Exponentiale und SONC-Polynome mit offenen Fragen und Forschungsanstrengungen verbunden sind, weisen sie eindeutig einige grundlegende strukturelle Phänomene auf, die sich gut im Rahmen der dünnbesetzten Trägermengen ausnutzen lassen. Aufbauend auf den früheren Arbeiten der Autorin in [DNT21] über den dualen Kegel der SONC-Polynome beginnt dieser Abschnitt mit der Untersuchung des dualen Kegels der SAGE-Exponentiale. Insbesondere wird eine projektionsfreie Darstellung des dualen Kegels der SAGE-Exponentiale hergeleitet und verwendet, um die Extremalstrahlen des *primalen* SAGE-Kegels vollständig zu charakterisieren. Diese Ergebnisse können anschließend verwendet werden, um effiziente Optimierungsansätze auf Basis des SAGE-Kegels zu untersuchen.

Im Rahmen dieses Abschnitts wird das Konzept *reduzierter Kreise* eingeführt. Reduzierte Kreise sind Kreise, die keine zusätzlichen Punkte der Gesamtträgermenge in ihrer konvexen Hülle enthalten. Unter Verwendung dieses Konzepts wird eine umfassende Charakterisierung des dualen SAGE-Kegels hergeleitet, siehe hierfür Theorem 3.1.5. Diese Charakterisierung liefert insbesondere projektionsfreie Charakterisierungen in Form von AGE-Exponentialen, deren Träger reduzierte Kreise sind. Die Charakterisierungen des dualen Kegels gehen weit über die Charakterisierungen des dualen Kegels der SAGE-Exponentiale aus [CS16] und des dualen Kegels der SONC-Polynome aus [DNT21] hinaus, wo die dualen Kegel in Form von Projektionen beschrieben werden.

Basierend auf den Charakterisierungen des Dualen des SAGE-Kegels wird gezeigt, dass jedes SAGE-Exponential als Summe nichtnegativer Kreisexponentiale geschrieben werden kann, deren Trägermengen reduzierte Kreise bilden, siehe Theorem 3.2.1, und eine exakte Charakterisierung der Extremalstrahlen des SAGE-Kegels hergeleitet, siehe Theorem 3.2.4. Diese Charakterisierung verschärft die notwendigen Bedingungen in [MCW21a] (siehe auch [Wan18a]) wesentlich.

Symmetriereduktion in AM/GM-Basierter Optimierung

Aus algebraischer Sicht ist ein Problem *symmetrisch*, wenn es invariant unter einer Gruppenoperation ist. Symmetrien sind im Kontext von Polynomen oder Exponentialsummen und Optimierung allgegenwärtig, da sie sich sowohl in der Problemformulierung als auch in der Lösungsmenge manifestieren. Dadurch lässt sich oft die Komplexität der entsprechenden algorithmischen Fragestellungen reduzieren. Bezüglich der Lösungsmenge wurde bereits 1840 von Terquem beobachtet, dass ein symmetrisches Polynom nicht immer einen vollständig symmetrischen Minimierer hat (siehe auch die Übersicht von Waterhouse [Wat83]). In vielen Fällen enthält die Menge der Minimierer jedoch symmetrische Punkte, siehe z. B. [FRS18; MRV21; Rie12; Tim03]. Im Hinblick auf Problemformulierungen hat die Symmetriereduktion in vielen Situationen wesentliche Fortschritte gebracht, siehe z.B. [BV08; KS10; DV15], insbesondere im Kontext von Summen von Quadraten, siehe z.B. [Bac+12; BR21; DR20; GP04; HHS21; Ray+18; Rie+13].

Dieser Abschnitt untersucht, inwieweit Symmetrien in der AM/GM-basierten Optimierung ausgenutzt werden können, unter der Annahme, dass das Problem selbst Symmetrien aufweist. Damit wird eine erste systematische Untersuchung AM/GM-basierter Ansätze in G-invarianten Situationen für eine gewählte Gruppe G betrachtet.

Es wird ein Symmetrie-adaptiertes Zerlegungstheorem bewiesen und eine angepasste relative Entropieformulierung von G-invarianten SAGE-Exponentialen entwickelt. Diese Adaption reduziert die Größe der resultierenden relativen Entropie-Programme oder geometrischen Programme, siehe Theorem 4.2.1, Theorem 4.2.3 und Korollar 4.2.6. Wie aus diesen Aussagen hervorgeht, hängt die Verbesserung von der Orbitstruktur der Gruppenoperation ab.

Die strukturellen Ergebnisse in dieser Arbeit werden in Form von Berechnungen ausgewertet. In Situationen mit starker Symmetriestruktur verringert sich die

Anzahl der Variablen und die Anzahl der Gleichungen und Ungleichungen wesentlich. Dementsprechend reduziert sich die Rechenzeit der Inneren-Punkte-Verfahren, die der Berechnung von SAGE-Schranken zugrunde liegen. In verschiedenen Fällen ist die symmetrieangepasste Berechnung sogar erfolgreich, wenn das konventionelle SAGE-Programm versagt.

Globale Optimierung Mittels des Dualen SAGE-Kegels und Linearer Programmierung

Unter Verwendung des dualen Kegels der Summen von AGE-Exponentialen wird in diesem Abschnitt eine Relaxation globaler Optimierungsprobleme zur Minimierung einer Exponentialsumme und, als Spezialfall, eines multivariaten reellen Polynoms geliefert. Die Idee dieses Optimierungsansatzes besteht darin, das globale Optimierungsproblem (1.1) durch Optimierung über einen Kegel mit Koeffizienten zu relaxieren, der durch den dualen SONC-Kegel induziert wird. Dieser Ansatz ist motiviert durch die aktuellen Arbeiten [DNT21] und [MCW21a] sowie Kapitel 3 in dieser Arbeit und baut auf zwei Beobachtungen auf, die die wichtigsten theoretischen Beiträge liefern:

1. Der duale Kegel der AGE-Exponentiale ist im primalen enthalten, siehe Proposition 4.3.1. Auch eine Variante des dualen SAGE-Kegels ist im primalen SAGE-Kegel enthalten.

2. Die Optimierung über diesen modifizierten dualen Kegel kann durch Lösen eines linearen Programms erfolgen, siehe Proposition 4.3.4.

Es sei darauf hingewiesen, dass weder der primale noch der duale SAGE-Kegel polyedrisch ist; siehe in diesem Zusammenhang auch die Ergebnisse in [FW19]. Der Ansatz funktioniert wie folgt: Zunächst wird eine geliftete Version des dualen Kegels mit zusätzlichen linearen Hilfsvariablen untersucht (Theorem 4.1.6 (3)). Als Zweites wird gezeigt, dass die Koeffizienten einer gegebenen Exponentialsumme als durch Variablen im dualen Kegel induziert interpretiert werden können, siehe (4.21). Drittens wird erläutert, dass durch Fixierung dieser Koeffizientenvariablen ein Optimierungsproblem entsteht, das nur die linearen Hilfsvariablen beinhaltet, siehe Proposition 4.3.4.

Basierend auf den beiden oben genannten Schlüsselbeobachtungen werden zwei lineare Programme vorgestellt, die eine Relaxation des Problems (1.1) lösen.

Eine Primal-Duale Sicht auf Darstellbarkeit Zweiter Ordnung

Wie bereits erläutert bilden die Kegel der Summen arithmetisch-geometrischer Exponentiale und Summen nichtnegativer Kreispolynome Nichtnegativitätszertifikate auf Basis der arithmetisch-geometrischen Ungleichung und sind besonders nützlich im Zusammenhang mit dünnbesetzten Polynomen und Exponentialsummen.

In Kapitel 5 wird ein Kegel eingeführt und untersucht, der aus einer Klasse von verallgemeinerten Polynomfunktionen besteht und der einen gemeinsamen Rahmen für aktuelle Nichtnegativitätszertifikate von dünnbesetzten Polynomen und Exponentialsummen bietet. Dieser $\mathcal{S}$-Kegel verallgemeinert und vereinheitlicht sowohl den Kegel der SONC-Polynome als auch den Kegel der SAGE-Exponentiale. Insbesondere können mehrere Ergebnisse im Zusammenhang mit diesen Kegeln — wie die Charakterisierungen des dualen Kegels und der Extremalstrahlen — auf den $\mathcal{S}$-Kegel übertragen werden.

Da die Nichtnegativität einer Polynomfunktion $f(x_1, \ldots, x_n)$ auf $\mathbb{R}_+^n$ äquivalent zur Nichtnegativität von $f(|x_1|, \ldots, |x_n|)$ auf $\mathbb{R}^n$ ist, werden in diesem Abschnitt allgemeinere Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ der Form

$$f(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha + \sum_{\beta \in \mathcal{B}} d_\beta x^\beta, \tag{A.3}$$

mit Exponentenmengen $\mathcal{A} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathbb{N}^n \setminus (2\mathbb{N})^n$ betrachtet, welche auch die Exponentialsummen erfassen. Basierend auf einer Teilmenge dieser Funktionen wird der *$\mathcal{S}$-Kegel* $C_\mathcal{S}(\mathcal{A}, \mathcal{B})$ definiert, der eine Verallgemeinerung der oben genannten Kegel darstellt, siehe Definition 5.1.3. Dieser $\mathcal{S}$-Kegel enthält sogenannte *AG-Funktionen* und Summen dieser Elemente. AG-Funktionen sind Funktionen der Form (1.3) mit gewissen Trägerbedingungen. Sie können als eine (nicht-polynomielle) Verallgemeinerung von Polynomen angesehen werden.

Eine Motivation für die Einführung dieser Klasse von Funktionen ist, dass sie die gemeinsame Betrachtung der Nichtnegativität von Polynomen auf $\mathbb{R}^n$ und die Nichtnegativität von Polynomen auf dem nichtnegativen Orthanten $\mathbb{R}_+^n$ ermöglicht. Außerdem ist die globale Nichtnegativität der Summanden $\sum_{\alpha \in \mathcal{A}} c_\alpha |x|^\alpha$ äquivalent zur globalen Nichtnegativität der Exponentialsumme $y \mapsto \sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle \alpha, y \rangle}$.

Sowohl vom geometrischen als auch vom Optimierungsstandpunkt aus ist es von großem Interesse, zu verstehen, wie die verschiedenen Klassen von Kegeln miteinander verwandt sind und ob Techniken für verschiedene Kegel kombiniert werden können. Hier sei daran erinnert, dass Averkov bereits zeigte, dass der SONC-Kegel als Projektion eines Spektraeders dargestellt werden kann [Ave19]. In der Tat verwendet sein Beweis die Techniken von [BN01], die zeigen, dass der SONC-Kegel sogar mittels Kegel zweiter Ordnung darstellbar ist. Wang und Magron gaben einen alternativen Beweis basierend auf binomialen Quadraten und $\mathcal{A}$-gemittelten Mengen an [WM20b]. Beide Ansätze betrachten nur den primalen SONC-Kegel.

In dieser Arbeit wird der $\mathcal{S}$-Kegel und sein dualer Kegel unter dem Gesichtspunkt der Darstellbarkeit zweiter Ordnung betrachtet — und damit auch seine Spezialisierungen. Durch die Erweiterung der Ergebnisse von Averkov und von Wang und Magron, die den primalen SONC-Kegel betreffen, werden explizite verallgemeinerte Second-Order-Cone-Programme für rationale $\mathcal{S}$-Kegel und ihre Dualen unter Berücksichtigung der Extremalitätsergebnisse aus Kapitel 3 geliefert, um die Größe dieser Probleme zu reduzieren, siehe Korollare 5.3.18 und 5.3.19. Der Beweis in dieser Arbeit kombiniert die Techniken für die Darstellungen zweiter Ordnung aus [BN01] mit den Konzepten und der Dualitätstheorie aus Kapitel 3 (die entsprechenden Ergebnisse in der Sprache des $\mathcal{S}$-Kegels sind in [KNT21] zu finden). Die Herleitung der Ergebnisse in dieser Arbeit unterscheidet sich von dem Ansatz von Wang und Magron, und sie benötigt keine binomialenen Quadrate oder $\mathcal{A}$-gemittelte Mengen. Außerdem werden in den Second-Order-Cone Programmen, dank der Charakterisierung der Extremalstrahlen des $\mathcal{S}$-Kegels aus Kapitel 3 (in der Sprache des $\mathcal{S}$-Kegels siehe wiederum [KNT21]), keine redundanten Kreise berücksichtigt.

Sublineare Kreise und der Bedingte SAGE-Kegel

In diesem Teil der Arbeit wird der restringierte Fall untersucht, d. h. für eine konvexe und nichtleere Menge X wird das *Optimierungsproblem unter Nebenbedingungen*

$$f_X^* = \inf\{f(x) : x \in X\} = \sup\{\lambda : f - \lambda \geq 0 \text{ auf } X\}$$

für eine Exponentialsumme f mit dünnbesetzter Trägermenge $\mathcal{A} \subseteq \mathbb{R}^n$ betrachtet.

In der Vergangenheit wurden als bedingte SONC-Relaxation eines polynomiellen Optimierungsproblems unter Nebenbedingungen zumeist Zerlegungen von $f - \lambda$ in Summen nichtnegativer Polynome untersucht, deren Träger klassische $\mathbb{R}^n$-Kreise bilden [DKW21; DIW19]. In 2019 entwickelten Murray, Chandrasekaran und Wierman jedoch einen anderen Ansatz. Sie untersuchten, wann eine Exponentialsumme mit höchstens einem negativen Koeffizienten — deren Träger nicht notwendigerweise einen Kreis bilden muss — nichtnegativ auf einer gegebenen konvexen Menge X ist [MCW21b]. Sie nutzten die Tatsache, dass äquivalent zur Untersuchung der Nichtnegativität einer Exponentialsumme f auf X mit höchstens einem negativen Koeffizienten c_β, d.h.,

$$f \;=\; \sum_{\alpha \in \mathcal{A}\setminus\{\beta\}} c_\alpha e^{\langle \alpha, x\rangle} + c_\beta e^{\langle \beta, x\rangle} \quad \text{mit} \quad c_\alpha \geq 0 \;\text{ für alle } \alpha \text{ in } \mathcal{A} \setminus \{\beta\},$$

auch die Nichtnegativität von $\sum_{\alpha \in \mathcal{A}} c_\alpha e^{\langle (\alpha-\beta), x\rangle}$ untersucht werden kann. Dies ist per Konstruktion eine konvexe Funktion. Daher kann die X-Nichtnegativität dieser Funktion genau charakterisiert werden, indem man das Prinzip der starken Dualität aus der konvexen Optimierung anwendet. Dies führt zu einem relativen Entropie-Programm in einer dualen Variablen $\nu = (\nu_\alpha)_{\alpha \in \mathcal{A}}$ und den Koeffizienten der Exponentialsumme, unter Einbeziehung der Stützfunktion von X (siehe Proposition 2.4.12 für die explizite Formulierung). Das Gleiche gilt natürlich auch für die Optimierungsformulierung.

In Anlehnung an [MCW21b] wird dieser Ansatz *bedingtes SAGE* genannt. Wenn X fest gewählt ist, werden die X-nichtnegativen Exponentialsummen mit höchstens einem negativen Koeffizienten *X-AGE-Exponentiale* genannt; die Exponentialsummen, die sich in eine Summe solcher Funktionen zerlegen lassen, werden in dieser Arbeit *X-SAGE* genannt. Ähnlich wie im unrestringierten Fall kann mittels eines relativen Entropie-Programms entschieden werden, ob eine gegebene Funktion f X-SAGE ist.

Obwohl dieses relative Entropie-Programm auch ohne Berücksichtigung eines Kreis-ähnlichen Begriffs funktioniert, verrät es nichts über die Struktur des bedingten SAGE-Kegels. Ähnlich wie viele Ergebnisse für den unrestringierten Fall, die teilweise in Kapitel 3 behandelt werden, siehe auch [FW19; KNT21; MCW21a], dient dieser Abschnitt dem Versuch, Extremalität, Dualität und numerische Probleme des Kegels zu verstehen. Dazu wird das Konzept der *sublinearen Kreise* — manchmal auch *die X-Kreise von $\mathcal{A}$* genannt, wenn die betrachtete Träger- oder Nebenbedingungsmenge eine wichtige Rolle spielt — eingeführt. X-Kreise von $\mathcal{A}$ sind Nicht-Null-Vektoren $\nu^\star \in \mathbb{R}^{\mathcal{A}}$, bei denen die Stützfunktion $\nu \mapsto \sigma_X(-\mathcal{A}\nu)$ eine strikte Sublinearitätsbedingung aufweist (siehe Definition 6.1.2). Die Konstruktion stellt sicher, dass der Spezialfall der $\mathbb{R}^n$-Kreise auf die simplizialen Kreise des durch $\mathcal{A}$ induzierten affin-linearen Matroids reduziert wird. Auch für reelle Teilmengen $X \subsetneq \mathbb{R}^n$ haben sublineare Kreise affin unabhängige positive Trägermengen, vgl. Proposition 6.1.5.

Verbindet man die Theorie der sublinearen Kreise mit X-nichtnegativen AGE-Exponentialen, so zeigt Theorem 6.2.2, dass für jedes X-AGE-Exponential ein sublinearer Kreis existiert, der als duale Variable in der relativen Entropieformulierung dient. Außerdem induziert jeder normierte X-Kreis λ einen *λ-bezeugten X-AGE-Kegel* $C_X(\mathcal{A}, \lambda)$, der eine verallgemeinerte Variante der Kreiszahl-Bedingung ([IW16a]) vom unrestringierten Fall erfüllt. Die Verallgemeinerung bezieht wieder die Stützfunktion von X mit ein. Die Vereinigung all dieser Kegel ist wiederum der gesamte Kegel der X-AGE-Exponentiale, deren Träger in $\mathcal{A}$ enthalten ist, siehe Theorem 6.2.4. Abgesehen von diesen Ähnlichkeiten zum unrestringierten Fall zeigt Beispiel 6.2.9 jedoch, dass ein sublinearer Kreis im Allgemeinen nicht allein anhand seines Trägers identifizieren werden kann.

Abschnitt 6.3 geht der Frage nach, welche X-Kreise tatsächlich für die Darstellung des X-AGE-Kegels sowie des X-SAGE-Kegels notwendig sind. Dazu wird der Begriff der *reduzierten X-Kreise* entwickelt — diejenigen X-Kreise ν, für die $(\nu, \sigma_X(-\mathcal{A}\nu))$ einen Extremalstrahl des Kreisgraphen

$$\operatorname{pos}\left(\{(\lambda, \sigma_X(-\mathcal{A}\lambda)) \,:\, \lambda \text{ normierter } X\text{-Kreis von } \mathcal{A}\} \cup \{(\mathbf{0}, 1)\}\right)$$

erzeugt. Tatsächlich kann der bedingte SAGE-Kegel nur mit λ-bezeugten X-AGE-Kegeln konstruiert werden, sofern alle λ aus der Menge der reduzierten normierten X-Kreise betrachtet werden.

Sublineare Kreise für Polyedrische Zulässigkeitsbereiche

Im Fall polyedrischer Nebenbedingungen vereinfacht sich die Frage der X-Nichtnegativität erheblich und liefert interessante Ergebnisse. Daher wird ein besonderer Fokus auf die Situation polyedrischer Mengen X gelegt. In diesem Fall lassen sich die sublinearen Kreise exakt durch den Normalenfächer eines bestimmten Polyeders charakterisieren, siehe Theorem 6.1.8. Für polyedrische X ist die Anzahl der sublinearen Kreise endlich, und dies ergibt Zerlegungen der X-SAGE-Kegel in *endlich* viele Summanden, induziert von sublinearen Kreisen, siehe Theorem 6.3.6, im Folgenden paraphrasiert.

> *Sei X ein Polyeder. Es bezeichne $\Lambda_X^\star(\mathcal{A})$ die Menge der normierten reduzierten X-Kreise von $\mathcal{A}$. Sei $\Lambda_X^\star(\mathcal{A})$ nichtleer, und der Kegel der X-SAGE-Exponentiale bestehe aus mindestens einem nichtpositiven Term über X. Dann ist der Kegel gleich der Summe*
>
> $$\sum_{\lambda \in \Lambda_X^\star(\mathcal{A})} C_X(\mathcal{A}, \lambda).$$
>
> *Außerdem gibt es keine echte Teilmenge $\Lambda \subsetneq \Lambda_X^\star(\mathcal{A})$ der Menge der normierten reduzierten X-Kreise mit $\sum_{\lambda \in \Lambda_X^\star(\mathcal{A})} C_X(\mathcal{A}, \lambda) = \sum_{\lambda \in \Lambda} C_X(\mathcal{A}, \lambda)$.*

Theorem 6.3.6 liefert die effizienteste mögliche Beschreibung des X-SAGE-Kegels in Form von Potenzkegel-Ungleichungen.

Innerhalb der Klasse der Polyeder weisen polyedrische Kegel besonders schöne Eigenschaften auf. Man beachte, dass der unrestringierte Spezialfall $X = \mathbb{R}^n$, der in Kapitel 3 behandelt wird, ebenfalls in die Klasse der polyedrischen Kegel fällt. Jeder univariate Fall kann in einen der beiden Kegelfälle $\mathbb{R}$ (unrestringierter Fall), $\mathbb{R}_+$ (einseitiges Unendlichkeitsintervall), oder in den nicht-konischen Fall $[-1, 1]$ (kompaktes Intervall) überführt werden. Im multivariaten Fall sind die Polyeder $\mathbb{R}^n$, $\mathbb{R}_+^n$ (nichtnegativer Orthant) und der Würfel $[-1, 1]^n$ prominente Beispiele. Im Gegensatz zum unrestringierten Fall und zum nichtnegativen Orthanten liefert der Würfel $[-1, 1]^n$ einen nicht-konischen Fall.

Im gesamten Kapitel 6 werden Resultate mit dem univariaten kompakten Intervall $X = [-1, 1]$ und der Halbgeraden $X = [0, \infty)$ veranschaulicht. Des Weiteren werden die X-Kreise einer Punktmenge $\mathcal{A} \subseteq \mathbb{R}$ für beide Mengen X sowie die entsprechenden *reduzierten* X-Kreise untersucht. Dies mündet in einer vollständigen Charakterisierung der Extremstrahlen des X-SAGE-Kegels für $X = [-1, 1]$ und $X = [0, \infty)$ und $\mathcal{A} \subseteq \mathbb{R}$ (Theorem 6.4.1).

AM/GM-Basierte Optimierung unter Nebenbedingungen

Als Verbindung zu den im Rahmen dieser Arbeit erzielten Optimierungsergebnissen für den unrestringierten Fall zeigt der Abschluss dieser Arbeit, dass bestimmte Optimierungs- und Zerlegungsergebnisse für den unrestringierten Fall auch im restringierten gelten.

Insbesondere kann immer dann, wenn die Menge der Nebenbedingungen X für eine Gruppe G symmetrisch ist, ebenfalls eine symmetrische Zerlegung eines X-AGE-Exponentials gefunden werden, die zu einem relativen Entropie-Programm mit wesentlich reduzierter Größe führt, siehe Theorem 7.1.1 und Korollar 7.1.2.

In Abschnitt 7.2 wird die Optimierung über einen Kegel, der durch das Duale des X-SAGE-Kegels induziert wird, betrachtet. Hier müssen verschiedene Einschränkungen an die Menge X der Nebenbedingungen gemacht werden, nämlich, dass sie polyedrisch und kegelförmig ist. Die Notwendigkeit der Kegelförmigkeit ergibt sich aus der Tatsache, dass in diesem Fall die Stützfunktion $\sup_{x \in X} -(\mathcal{A}\nu)^T x$ immer dann zu 0 auswertet, wenn ihr Wert endlich ist. Dadurch wird sichergestellt, dass die X-AGE-artigen Exponentiale mit Koeffizienten im dualen X-SAGE-Kegel im primalen enthalten sind. Die Einschränkung, dass X polyedrisch ist, führ dazu, dass die Anzahl der Kreise endlich ist, was die Potenzkegeldarstellung induziert. Außerdem stellt sie sicher, dass das resultierende Optimierungsprogramm linear ist: Wäre X nicht polyedrisch, könnte die Menge X nicht durch das Lösen linearer Nebenbedingungen beschrieben werden.

Des Weiteren wird auch eine Kegeldarstellung zweiter Ordnung für den bedingten SAGE-Kegel und seinen dualen Kegel für polyedrische X und rationale Ausdrücke $\mathcal{A}^T X$ geliefert, siehe Abschnitt 7.3. Letzteres war bereits eine Einschränkung im unrestringierten Fall. Ersteres wird erneut für das Endlichkeitsresultat der Menge der X-Kreise sowie für die Potenzkegeldarstellbarkeit des X-SAGE-Kegels in diesem Fall benötigt.