

SELF-CONCORDANT BARRIERS AND CHEBYSHEV SYSTEMS*

LEONID FAYBUSOVICH†

Abstract. We explicitly calculate characteristic functions of cones of generalized polynomials corresponding to Chebyshev systems on intervals of the real line and the circle. Thus, in principle, we calculate homogeneous self-concordant barriers for this class of cones. This class includes almost all "cones of squares" considered in [4]. Our construction, however, does not use this structure and is applicable to a much broader class of cones. Even for "cones of squares" within the considered class our results are new.

Key words. interior-point algorithms, characteristic functions of convex cones, T-systems

AMS subject classifications. 90C51, 90C34

1. Introduction. To apply a modern interior-point technique as it is developed in [5], it is necessary to know a self-concordant barrier for a feasible domain of a given convex optimization problem. Given a convex domain in a finite-dimensional vector space, there exists an explicit formula for at least one such a barrier, the so-called universal barrier function [5]. For example, let K be a closed convex pointed cone in $\mathbf{R}^n$ with a nonempty interior. Consider

$$(1.1) \quad \Phi(p) = \ln \int_{K^*} e^{-\langle c, p \rangle} d\mu(c),$$

where $p \in \text{int}(K)$, K^* is the cone dual to K and μ is the standard Lebesgue measure on $\mathbf{R}^n$. Then Φ after an appropriate normalization is the so-called homogeneous self-concordant barrier function. The knowledge of such a function in a "computable" form, enables one, in principle, to develop interior-point algorithms (along with complexity estimates) for optimization problems whose feasibility domain is the intersection of K with an affine subspace in $\mathbf{R}^n$ and for many other related problems (through the barrier calculus). Unfortunately, the expression (1.1) requires the evaluation of multidimensional integrals over geometrically complicated domains for the computation of the value of Φ , its gradient and the Hessian at a given point $p \in \text{int}(K)$. This is, in general, computationally too expensive taking into account the original task in question, i.e. solving a convex optimization problem. There are a number of situations where (1.1) can be more or less explicitly calculated. Most of the corresponding cones belong to the class of symmetric cones and (1.1) is then easily expressed in terms of the attached Jordan algebra (see e.g. [2]). A part of the theory of interior-point algorithms admits an infinite-dimensional generalization [6] but the concept of the universal barrier function seems to be essentially finite-dimensional.

In the present paper we significantly expand the class of cones for which (1.1) can be explicitly calculated. Respectively, we expand the class of optimization problems to which the modern interior-point technique can be applied. Namely, we consider cones of generalized nonnegative polynomials generated by Chebyshev systems on the intervals of the real line or the unit circle. For such cones we find more or less explicit expressions for (1.1) only slightly more complicated (in computational sense) than for symmetric cones. In particular, practically all cones considered in [4] can

*This work was supported in part by NSF grant DMS98-03191 and DMS01-02628

†Department of Mathematics, University of Notre Dame, Notre Dame, In , 46556, USA
(leonid.faybusovich.1@nd.edu)

be treated from our viewpoint. Note, however, that the representation of a given cone as a "cone of squares" (and hence the reducibility of a given problem to the semidefinite programming) which is crucial for Nesterov's construction, does not play any role in our approach. Thus, our results are applied to a broader class of cones. The calculation of (1.1) is new even for most of the cones considered in [4].

2. Chebyshev systems. We start with several examples of Chebyshev systems. We then formulate several important for us properties of such systems.

DEFINITION 1. A system of real functions $u_0, \dots, u_n$ defined on an abstract set E is called a Chebyshev system (T -system) of order n on E if the determinant

$$\det(u_i(t_j)),$$

$i, j = 0, 1, \dots, n$, does not vanish for any pairwise distinct $t_0, \dots, t_n \in E$. If the set E is endowed with a topology, one usually assumes that the functions $u_0, \dots, u_n$ are continuous on E . In this paper we are mostly interested in the cases where $E = [a, b] \subset \mathbf{R}$ or $E = \mathbf{S}^1$ (unit circle). In the latter case, $\mathbf{S}^1$ may be viewed as an interval $[a, b]$ with identified endpoints. A T -system on a circle is a T -system of functions on $[a, b]$ with the additional property that $u_k(a) = u_k(b)$, $k = 0, 1, \dots, n$.

Consider several examples of T -systems.

EXAMPLE 1. Let $u_i(t) = t^i$, $i = 0, 1, \dots, n$, $t \in [a, b]$. This is a T -system as it easily follows from the properties of the Vandermonde determinant.

EXAMPLE 2. The system of functions $t^m, t^{m-1}, \dots, t, 1, (t-x_1)_+^m, (t-x_2)_+^m, \dots, (t-x_r)_+^m$, form the so-called WT -system on the interval $[-1, 1]$, provided $-1 < x_1 < \dots < x_r < 1$. Here $x_+ = \max\{x, 0\}$. The requirement here is that all determinants from the Definition 1 are nonnegative.

EXAMPLE 3. The functions

$$1, \sin t, \dots, \sin(nt), \cos t, \dots, \cos(nt)$$

form a periodic T -system on $[0, 2\pi]$ of the order $2n$. One can show that every periodic T -system has an even order. For a detailed discussion of examples given above and many more examples see e.g. [3].

Given a T -system $u_0, \dots, u_n$ on the interval $[a, b]$, consider the cone K of nonnegative generalized polynomials associated with this system:

$$K = \{p = \sum_{i=0}^n a_i u_i : p(t) \geq 0, \forall t \in [a, b]\}.$$

We can associate with K the dual cone

$$K^* = \{(c_0, \dots, c_n)^T \in \mathbf{R}^{n+1} : \sum_{i=0}^n c_i a_i \geq 0,$$

$$\forall p = \sum_{i=0}^n a_i u_i \in K\}.$$

THEOREM 1. *We have :*

$$\text{int}(K) = \{p \in K : p(t) > 0, \forall t \in [a, b]\} \neq \emptyset.$$

The vector $(c_0, \dots, c_n)^T \in K^$ if and only if there exists a Borel measure σ on $[a, b]$ such that*

$$(2.1) \quad c_i = \int_a^b u_i(t) d\sigma(t), i = 0, 1, \dots, n.$$

For a proof of Theorem 1 see e.g. [3]. If in the representation (2.1) the corresponding measure σ is concentrated in a finite number of points

$$a \leq \xi_1 < \xi_2 < \dots < \xi_m \leq b,$$

then (2.1) takes the form:

$$(2.2) \quad c_i = \sum_{j=0}^m \rho_j u_i(\xi_j), i = 0, 1, \dots, m,$$

$\rho_j > 0$. Following [3], the points ξ_j involved in the representation (2.2) will be called the roots and the coefficients ρ_j will be called the weights. We further introduce the notation $\epsilon(t), a \leq t \leq b$, where $\epsilon(t) = 2, a < t < b, \epsilon(a) = \epsilon(b) = 1$. The sum

$$\sum_{j=1}^m \epsilon(\xi_j)$$

will be called the index of the representation (2.2). A representation (2.2) is called principal if its index is equal to $n + 1$, where n is the order of the Chebyshev system $u_0, \dots, u_n$. Consider the possible types of principal representations. If $n = 2\nu - 1, \nu = 1, 2, \dots$, then either all $\xi_j \in (a, b), m = \nu$, or $\xi_j \in (a, b), j = 2, 3, \dots, \nu, \xi_1 = a, \xi_{\nu+1} = b, m = \nu + 1$. In the former case the corresponding representation (2.2) is called the lower principal representation and in the latter case the representation (2.2) is called upper principal representation. If $n = 2\nu$, then either ν roots $\xi_j, j = 2, 3, \dots, \nu + 1$ belong to (a, b) and $\xi_1 = a, m = \nu + 1$, or ν roots $\xi_j, j = 1, 2, \dots, \nu$ belong to (a, b) and $\xi_{\nu+1} = b, m = \nu + 1$. In the former case the representation (2.2) is called the lower principal representation and in the latter case the representation (2.2) is called upper principal representation. Thus, a principal representation is upper or lower according to whether it has or has not a root at the right end point b of the interval $[a, b]$.

THEOREM 2. *Given a T -system $u_0, \dots, u_n$ on the interval $[a, b]$, each point $c \in \text{int}(K^*)$ (see (2.1)) has exactly one lower principal representation and exactly one upper principal representation.*

This result admits the following modification for the case of a periodic T -system on the interval $[a, b], n = 2\nu$.

THEOREM 3. *Each point $c \in \text{int}(K^*)$ admits a unique representation (2.2) with $m = \nu + 1$ one of whose roots $\xi_1, \dots, \xi_{\nu+1}$ is a prescribed point $\xi \in [a, b]$. For a proof of Theorems 2,3 see e.g. [3].*

3. Calculation of characteristic functions. We are using now principal representations of elements of K^* to calculate the characteristic function of the cone K generated by a Chebyshev system $u_0, \dots, u_n$. We assume that $u_0, \dots, u_n$ are continuously differentiable functions on the interval $[a, b]$. Let us start with the case $n = 2\nu - 1$. Given $p \in K$, we wish to calculate

$$(3.1) \quad F(p) = \int_{K^*} e^{-\langle c, p \rangle} d\mu(c),$$

where μ is the standard Lebesgue measure on $\mathbf{R}^{n+1}$. We use the lower principal representation (2.2) to parametrize $\text{int}(K^*)$:

$$(3.2) \quad c_i = \sum_{j=1}^{\nu} \rho_j u_j(\xi_i),$$

$i = 0, 1, \dots, 2\nu - 1$. According to Theorem 2 the map (3.2) gives a one-to-one correspondence between

$$\mathbf{R}_+^\nu \times \{\xi \in \mathbf{R}^\nu : a < \xi_1 < \xi_2 < \dots < \xi_\nu < b\}$$

and $\text{int}(K^*)$. Here $\mathbf{R}_+ = \{x \in \mathbf{R} : x > 0\}$. Denote this map by $\Phi = \Phi(\rho_1, \dots, \rho_\nu, \xi_1, \dots, \xi_\nu)$. We obviously have:

$$\frac{\partial \Phi}{\partial \rho_j} = u(\xi_j),$$

$j = 1, 2, \dots, \nu$, where

$$u(\xi_j) = (u_0(\xi_j), \dots, u_{2\nu-1}(\xi_j))^T \in \mathbf{R}^{2\nu},$$

$$\frac{\partial \Phi}{\partial \xi_j} = \rho_j u'(\xi_j),$$

$j = 1, 2, \dots, \nu$. Thus, the Jacobian of this map is equal to:

$$|\det(u(\xi_1), \dots, u(\xi_\nu), \rho_1 u'(\xi_1), \dots, \rho_\nu u'(\xi_\nu))| =$$

$$\left(\prod_{k=1}^{\nu} \rho_k \right) |\det(u(\xi_1), u'(\xi_1), \dots, u(\xi_\nu), u'(\xi_\nu))|.$$

Making the change of variables in (3.1) and using the Fubini theorem, we obtain:

$$F(p) = \int_{a \leq \xi_1 < \xi_2 < \dots < \xi_\nu \leq b} |\det(u(\xi_1), u'(\xi_1), \dots, u(\xi_\nu), u'(\xi_\nu))| \times$$

$$\left(\int_{\mathbf{R}^+} \rho_1 e^{-p(\xi_1)\rho_1} d\rho_1 \dots \int_{\mathbf{R}^+} \rho_\nu e^{-p(\xi_\nu)\rho_\nu} d\rho_\nu \right) d\xi_1 \dots d\xi_\nu.$$

Here we used that if $p = a_0 u_0 + \dots a_n u_n$, then

$$\langle c, p \rangle = \sum_{i=0}^n c_i a_i = \sum_{i=0}^n \sum_{j=1}^{\nu} \rho_j u_i(\xi_j) a_i =$$

$$\sum_{j=1}^{\nu} \rho_j \sum_{i=0}^n a_i u_i(\xi_j) = \sum_{j=1}^{\nu} \rho_j p(\xi_j).$$

Let

$$V(\xi_1, \xi_2, \dots, \xi_\nu) = \det(u(\xi_1), u'(\xi_1) \dots u(\xi_\nu), u'(\xi_\nu)).$$

Since

$$\int_0^{+\infty} x e^{-\alpha x} dx = \frac{1}{\alpha^2}, \alpha > 0,$$

we obtain

$$(3.3) \quad F(p) = \int_{a < \xi_1 < \dots < \xi_\nu < b} \left(\prod_{j=1}^{\nu} \frac{1}{p(\xi_j)^2} \right) |V(\xi_1, \dots, \xi_\nu)| d\xi_1 \dots d\xi_\nu.$$

Observe that the function under the integral sign in (3.3) is symmetric with respect to variables $\xi_1 \dots \xi_\nu$. Hence,

$$F(p) = \frac{1}{\nu!} \int_a^b \dots \int_a^b \left(\prod_{j=1}^{\nu} \frac{1}{p(\xi_j)^2} \right) |V(\xi_1, \dots, \xi_\nu)| d\xi_1 \dots d\xi_\nu.$$

LEMMA 1. *The function $V(\xi_1, \dots, \xi_\nu)$ does not change the sign on $\mathbf{R}^\nu$.*

Proof

Consider, first, the case where $a < \xi_1 < \xi_2 \dots < \xi_\nu < b$. Let $\eta_1, \dots, \eta_\nu$ be such that

$$(3.4) \quad \xi_1 < \eta_1 < \xi_2 < \eta_2 \dots \xi_\nu < \eta_\nu \leq b.$$

Since $u_0, \dots, u_{2\nu-1}$ is a T -system, we can assume without loss of generality that:

$$\gamma(\xi_1, \dots, \xi_\nu, \eta_1, \dots, \eta_\nu) = \det[u(\xi_1), u(\eta_1), u(\xi_2), u(\eta_2), \dots, u(\xi_\nu), u(\eta_\nu)] > 0$$

for any ξ, η_i satisfying (3.4). By the mean value theorem we have

$$\gamma(\xi_1, \dots, \xi_\nu, \eta_1, \dots, \eta_\nu) =$$

$$\prod_{i=1}^{\nu} (\eta_i - \xi_i) \det[u(\xi_1), u'(\theta_1), u(\xi_2), u'(\theta_2), \dots, u(\xi_\nu), u'(\theta_\nu)] > 0$$

for some $\xi_i < \theta_i < \eta_i, i = 1, 2, \dots, \nu$. Hence,

$$\det[u(\xi_1), u'(\theta_1), u(\xi_2), u'(\theta_2), \dots, u(\xi_\nu), u'(\theta_\nu)] > 0.$$

Taking limit when $\eta_i \rightarrow \xi_i, i = 1, 2, \dots, \nu$, we obtain:

$$V(\xi_1, \dots, \xi_\nu) \geq 0$$

for all $a < \xi_1 < \xi_2 < \dots < \xi_\nu < b$. Using continuity of V , we obtain:

$$V(\xi_1, \dots, \xi_\nu) \geq 0$$

for $a \leq \xi_1 \leq \xi_2 \leq \dots \leq \xi_\nu \leq b$. Our final observation is that:

$$V(\xi_{\sigma(1)}, \xi_{\sigma(2)}, \dots, \xi_{\sigma(\nu)}) = V(\xi_1, \dots, \xi_\nu)$$

for any permutation σ of the set $\{1, 2, \dots, \nu\}$. Hence, $V(\xi_1, \xi_2, \dots, \xi_\nu) \geq 0$ for all ξ_i satisfying $a \leq \xi_i \leq b, i = 1, 2, \dots, \nu$.

Using Lemma 1, we obtain:

$$(3.5) \quad F(p) = \frac{\epsilon}{\nu!} \int_a^b \dots \int_a^b \det[\tilde{u}(\xi_1), \tilde{u}'(\xi_1), \dots, \tilde{u}(\xi_\nu), \tilde{u}'(\xi_\nu)] d\xi_1 \dots d\xi_\nu,$$

where

$$\tilde{u}(\xi) = \frac{u(\xi)}{p(\xi)}, \epsilon = \pm 1.$$

The next Proposition which is due to N.G. de Bruijn (see [1]) is crucial for the evaluation of the characteristic function.

PROPOSITION 1.

Let (X, μ) be a measurable space with a finite positive measure μ on X . Suppose that $\psi_1, \dots, \psi_{2n}, \phi_1, \dots, \phi_{2n}$ are integrable functions on X . Let

$$D = D(t_1, \dots, t_n)$$

be the determinant of the matrix with k -th row

$$\phi_k(t_1), \psi_k(t_1), \phi_k(t_2), \psi_k(t_2) \dots \phi_k(t_n), \psi_k(t_n),$$

$k = 1, 2, \dots, 2n, t_1, t_2, \dots, t_n \in X$. Then

$$\Lambda = \int_X \dots \int_X Dd\mu(t_1) \dots d\mu(t_n) = n! Pf(B).$$

Here $B = (b_{ij})$ is $2n \times 2n$ skew-symmetric matrix with

$$b_{ij} = \int_X [\phi_i(x)\psi_j(x) - \phi_j(x)\psi_i(x)]d\mu(x)$$

and $Pf(B)$ stands for the Pfaffian of B .

Proof.

Using the definition of the determinant, we have:

$$D(t_1, \dots, t_n) =$$

$$\sum_{\sigma \in \Sigma(2n)} (-1)^{\text{sign}\sigma} \phi_{\sigma(1)}(t_1)\psi_{\sigma(2)}(t_1)\phi_{\sigma(3)}(t_2)\psi_{\sigma(4)}(t_2) \dots \phi_{\sigma(2n-1)}(t_n)\psi_{\sigma(2n)}(t_n).$$

Hence,

$$\Lambda = \sum_{\sigma \in \Sigma(2n)} (-1)^{\text{sign}\sigma} \tilde{k}_{\sigma(1)\sigma(2)} \tilde{k}_{\sigma(3)\sigma(4)} \dots \tilde{k}_{\sigma(2n-1)\sigma(2n)},$$

where

$$\tilde{k}_{ij} = \int_X \phi_i(t)\psi_j(t)d\mu(t).$$

Observe now that in the expression above $\tilde{k}_{ij}$ can be substituted by its skew-symmetric part $l_{ij} = (k_{ij} - k_{ji})/2$. Indeed, consider a two-form

$$\beta = \sum_{1 \leq i, j \leq 2n} \alpha_{ij} e_i \wedge e_j \in \bigwedge^2(\mathbf{R}^{2n}).$$

Here $e_1, \dots, e_{2n}$ is a canonical basis in $\mathbf{R}^{2n}$ and α_{ij} are some real numbers. Taking n times the wedge product of β with itself, we obtain:

$$(\beta \wedge \beta \dots \wedge \beta = (\sum_{\sigma \in \Sigma(2n)} (-1)^{\text{sign}\sigma} \alpha_{\sigma(1)\sigma(2)} \dots \alpha_{\sigma(2n-1)\sigma(2n)}) \omega,$$

$$\omega = e_1 \wedge e_2 \dots \wedge e_{2n-1} \wedge e_{2n}.$$

On the other hand,

$$\beta = \sum_{1 \leq i < j \leq 2n} (\alpha_{ij} - \alpha_{ji}) e_i \wedge e_j = \sum_{1 \leq i, j \leq 2n} \gamma_{ij} e_i \wedge e_j,$$

where $\gamma_{ij} = (\alpha_{ij} - \alpha_{ji})/2$. Hence,

$$\beta \wedge \beta \dots \wedge \beta = \left(\sum_{\sigma \in \Sigma(2n)} (-1)^{\text{sign} \sigma} \gamma_{\sigma(1)\sigma(2)} \dots \gamma_{\sigma(2n-1)\sigma(2n)} \right) \omega.$$

Applying this observation to our situation, we obtain:

$$\Lambda = \frac{1}{2^n} \sum_{\sigma \in \Sigma(2n)} (-1)^{\text{sign} \sigma} b_{\sigma(1)\sigma(2)} b_{\sigma(3)\sigma(4)} \dots b_{\sigma(2n-1)\sigma(2n)},$$

Recalling the definition of the Pfaffian of an even dimensional skew-symmetric matrix (see e.g [1])

$$Pf(B) = \frac{1}{n!2^n} \sum_{\sigma \in \Sigma(2n)} (-1)^{\text{sign} \sigma} b_{\sigma(1)\sigma(2)} b_{\sigma(3)\sigma(4)} \dots b_{\sigma(2n-1)\sigma(2n)},$$

we obtain:

$$\Lambda = n! Pf(B).$$

We are now in position to calculate the characteristic function of a cone generated by a Chebyshev system of odd order. Applying Proposition 1 to (3.5), we obtain.

THEOREM 4. *Let $u_0, \dots, u_{2\nu-1}$ be a Chebyshev system of continuously differentiable functions on the interval $[a, b]$. Let p be a generalized polynomial strictly positive on $[a, b]$. Then*

$$F(p) = \epsilon Pf(B(p)),$$

where $B(p) = (b_{ij}(p))$,

$$b_{ij}(p) = \int_a^b \frac{u_i(t)u_j'(t) - u_j(t)u_i'(t)}{p(t)^2} dt,$$

$i, j = 0, 1, \dots, 2\nu - 1, \epsilon = \pm 1$.

The case of an even order Chebyshev system is slightly more complicated. Let $u_0, \dots, u_{2\nu}$ be a Chebyshev system of continuously differentiable functions on an interval $[a, b]$. Assume that

$$(3.6) \quad u_0(a) = 1, u_i(a) = 0, i = 1, \dots, 2\nu.$$

THEOREM 5. *Let $u_0, \dots, u_{2\nu}$ be a Chebyshev system of an even order of continuously differentiable functions on the interval $[a, b]$, such that $u(a) = e_1$. Let, further, p be a generalized polynomial strictly positive on $[a, b]$. Then*

$$F(p) = \epsilon \frac{Pf(B(p))}{p(a)},$$

where $B(p) = (b_{ij}(p))$,

$$b_{ij}(p) = \int_a^b \frac{u_i(\xi)u'_j(\xi) - u_j(\xi)u'_i(\xi)}{p(\xi)^2} d\xi,$$

$i, j = 1, 2, \dots, 2\nu$. Here $\epsilon = \pm 1$.

Similarly, using Theorem 3, we obtain.

THEOREM 6. *Let $u_0, \dots, u_{2\nu}$ be a periodic Chebyshev system of continuously differentiable functions on the interval $[a, b]$ such that $u(a) = e_1$. Let p be a generalized polynomial strictly positive on $[a, b]$. Then*

$$F(p) = \epsilon \frac{Pf(B(p))}{p(a)},$$

where $B(p) = (b_{ij}(p))$,

$$b_{ij}(p) = \int_a^b \frac{u_i(\xi)u'_j(\xi) - u_j(\xi)u'_i(\xi)}{p(\xi)^2} d\xi,$$

$i, j = 1, \dots, 2\nu$. Here $\epsilon = \pm 1$.

Observe now that the assumption made in Theorems 5, 6 does not restrict the generality of our approach.

LEMMA 2. *Let $u_0, u_1, \dots, u_n$ be a Chebyshev system on a set E . Let $a \in E$. One can always choose a basis $v_0, \dots, v_n$ in $\text{span}(u_0, \dots, u_n)$ such that $v_0(a) = 1, v_i(a) = 0, i = 1, 2, \dots, n$.*

Proof

Indeed, for any pairwise distinct points $t_i, i = 0, \dots, n$, there exists $v_i \in \text{span}(u_0, \dots, u_n)$ such that $v_i(t_j) = \delta_{ij}, i, j = 0, \dots, n$. (See e.g. [3]).

REMARK 1. *Since $F(p) > 0, p \in \text{int}(K)$ in Theorem 4-6, we conclude that $Pf(B(p))$ does not change the sign on $\text{int}(K)$. Furthermore, since $\det(B(p)) = Pf(B(p))^2$ (see e.g. [1]), we can easily rewrite $\ln F(p)$ in terms of $\ln \det B(p)$.*

REMARK 2. *The results of this paper can be extended to WT-systems.*

REFERENCES

- [1] N.G. de Bruijn, *On some multiple integrals involving determinants*, J. Indian Math. Soc., vol. 19, 1955, pp. 133-151.
- [2] L. Faybusovich, *Euclidean Jordan algebras and Interior-point algorithms*, J. Positivity, vol. 1, 1997, pp. 331-357.
- [3] S. Karlin and W.J. Studden, *Tchebycheff systems: With applications in analysis and statistics*, Interscience, New York, 1966.
- [4] Yu. Nesterov, *Squared Functional Systems and Optimization Problems*. In High performance optimization, Kluwer, Dordrecht, 2000, pp. 441-466.
- [5] Yu. Nesterov and Arkadii Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
- [6] J. Renegar, *Linear programming, complexity theory and elementary functional analysis*, Math. Programming, vol. 70, no 3, ser. A, pp.279-351.