Real Analysis Exchange
Vol. 29(2), 2003/2004, pp. 607–629

C.-G. Ambrozie[*], Institute of Mathematics, Romanian Academy, P.O. Box 1-764, RO-70700 Bucharest, Romania. email: `Calin.Ambrozie@imar.ro`

MAXIMUM ENTROPY AND MOMENT PROBLEMS

Abstract

We study moment problems in which one searches for a probability density on $\mathbb{R}^n$ by using insufficient information given in an integral form. We characterize the existence of the representing densities for a finite multi–sequence of moments by the solvability of a concrete finite system of equations. Its solution provides the (unique) representing density of maximum entropy allowed by the given data, that turns out to be the exponential of a polynomial to be determined. For all densities of this form, the system to be solved can be taken linear to be provided sufficiently many moments are known.

1 Introduction

A truncated problem of moments can be stated as follows. Let T be a measurable space endowed with a fixed nonnegative measure m. Let u_α, $\alpha \in A$ be a finite family of known functions on T. Given a finite set $\gamma = (\gamma_\alpha)_\alpha$ of numbers γ_α $(\alpha \in A)$, one asks to establish if there exist measures $\mu \geq 0$ on T such that

$$\int_T u_\alpha(t) d\mu(t) = \gamma_\alpha \quad (\alpha \in A). \tag{1}$$

In particular, one looks for absolutely continuous measures $\mu = fm$ with $f \geq 0$. If (1) holds, then μ (resp. f) is called a *representing* measure (resp. density) for the sequence γ. The problem is then to characterize those sets γ which have nonnegative representing measures, study the set of the solutions and find or approximate such measures μ.

Key Words: moment problem, representing measure, entropy
Mathematical Reviews subject classification: 44A60, 49J99
Received by the editors May 12, 2003
Communicated by: Bernd Kirchheim

[*]The author expresses his thanks to the UFR de Mathématiques – Université de Lille 1 for its kind hospitality during the elaboration of the present work.

We shall assume that T is a compact subset of either n–dimensional real space, or the n–dimensional torus $\mathbb{T}^n$. To fix the notation, we start by stating the results in the case when $T \subset \mathbb{R}^n$ and m is the Lebesgue measure on T. However, our method applies with minor modifications to more general situations, including both cases mentioned here. We will consider only absolutely continuous representing measures fm, with nonnegative density f from $L^1(T)$ – the space of all (classes of) measurable functions that are Lebesgue integrable on T with respect to m. Set $a := \operatorname{card} A$. We characterize the existence of such representing densities by the solvability of the following system

$$\int_T u_\alpha(t)\, \mathrm{e}^{\sum_{\beta \in A} x_\beta\, u_\beta(t)}\, dm(t) \,=\, \gamma_\alpha \qquad (\alpha \in A) \tag{2}$$

of a equations with a unknowns x_α $(\alpha \in A)$. Therefore if our problem (1) has any absolutely continuous solution $\mu = fm$, then it will necessarily have also a solution of the form from above. The concrete form of (2) then should allow the study of the existence of (or approximate) the vector $x = (x_\alpha)_{\alpha \in A} \in \mathbb{R}^a$, see for instance [1], [3]; note also Remark 1 in this sense. When the system (2) (see (24) and (29), too) has a solution, it is unique and provides the (also unique) representing density f_* having maximal entropy, by the formula

$$f_*(t) = f_{*,x}(t) = \exp\,(\sum_{\alpha \in A} x_\alpha u_\alpha(t)) \quad (t \in T).$$

Namely, f_* maximizes the Boltzmann's integral $-\int f \ln f dm$ amongst all the absolutely continuous measures $\mu = fm \geq 0$ satisfying the equalities (1).

For powers moment problems, we show that if there exists an integrable representing density of the form $f_* = \exp\,(\sum_{\alpha \in A} x_\alpha u_\alpha)$ on the whole space $\mathbb{R}^n$ or $\mathbb{T}^n$, then knowing the moments γ_α, $\alpha \in A + A$ provides the values of x_α $(\alpha \in A)$ by solving a compatible and determined linear system (30). Note the following example. Let $n = 1$ and $\gamma_0,\, \gamma_1,\, \gamma_2 \in \mathbb{R}$. Set $u_\alpha(t) = t^\alpha$ $(\alpha = 0, 1, 2)$. In this case one can use (2) to compute x_α by hand. Namely, assume that $f_*(t) := \exp\,(x_0 + x_1 t + x_2 t^2)$, $t \in \mathbb{R}$ is integrable and satisfies (2). Since $f_* \in L^1(\mathbb{R})$, we have $x_2 < 0$. Hence by the Leibniz–Newton formula we have $\int f_*' dt = 0$ and $\int (t f_*(t))' dt = 0$, where f' denotes the derivative of f. It follows that $x_1\gamma_0 + 2x_2\gamma_1 = 0$ and $\gamma_0 + x_1\gamma_1 + 2x_2\gamma_2 = 0$. Then $x_1 = \gamma_0\gamma_1 d^{-1}$, $x_2 = -\gamma_0^2 d^{-1}$ and $x_0 = \ln(\gamma_0 / \int \exp\,(x_1 t + x_2 t^2)\, dt)$, where $d := \gamma_0\gamma_2 - \gamma_1^2$. Hence $f_*(t) = C \exp\,[-(t-s)^2/d\,]$ is a multiple of the Gauss distribution of mean $s = \gamma_1/2$ and dispersion d. Thus we get the well–known fact that the maximum entropy probability density of given mean and dispersion is the normal one, see [17] for instance. Similar computations providing x in terms of the known data γ_α, $\alpha \in A$ can be done also when $A = \{\alpha = (\alpha_1, \ldots, \alpha_n) \in$

$\mathbb{Z}_+^n \mid \alpha_1 + \cdots + \alpha_n \leq 2\}$. (This moment problem has been solved in [11] by different methods.)

The solution in the above example is valid if and only if $\gamma_0 > 0$ and $\gamma_0\gamma_2 - \gamma_1^2 > 0$. This agrees with the theory of the truncated problem of power moments, existing for $n = 1$. Namely, a set $(\gamma_\alpha)_{\alpha=0}^{2k}$ ($\neq 0$) has nonnegative representing measures (resp. densities) on $\mathbb{R} \Leftrightarrow$ the quadratic form $(\gamma_{\alpha+\beta})_{\alpha,\beta=0}^{k}$ is nonnegative (resp. positive) definite [4]. Similar characterizations hold for $A := \{0, 1, \ldots, k\}$ and $T :=$ an interval in $\mathbb{R}$, for $A := \{-k, \ldots, k\}$ and $T := \mathbb{T}$ (with $u_\alpha(e^{i\theta}) := e^{i\alpha\theta}$ and $m := d\theta$), as well as in other 1–dimensional cases [4], [6], [13], [24]. They are based on the possibility to represent any nonnegative polynomial (resp. trigonometrical polynomial) on $\mathbb{R}$ (resp. $\mathbb{T}$) as a sum of squares [21]. Namely the existence of μ can be characterized by the existence of a functional L_γ on polynomials such that $L_\gamma\, t^\alpha = \gamma_\alpha$ and $L_\gamma p \geq 0$ for any $p \geq 0$ on T [18]. Then by the Hahn–Banach theorem it suffices to ask $L_\gamma p \geq 0$ for those $p = \sum_{\alpha \in A} c_\alpha u_\alpha$ such that $p(t) \geq 0$, $t \in T$ [18], [24]. The problem is to describe these polynomials p. If they are (or can be expressed in terms of) sums of squares, then conditions like $L_\gamma(|q|^2) \geq 0$ for all q lead to characterizations as above. This method is not applicable for $n > 1$ when the set of nonnegative polynomials is more difficult to handle [8]. (For instance, not all of them can be written as sums of squares.) The same questions appear for trigonometric moment problems on $\mathbb{T}^n$ [15], [23].

There are also other approaches providing the existence of certain "maximum entropy"-type solutions [6], [13], but usually their proofs (operator theoretic or complex analytic) can work only in the case $n = 1$. Various results exist in certain multidimensional cases, too [7], [10], [11], [22], [25]. Note in particular a complete solution of the truncated moment problem for finitely-atomic measures and flat data in the sense of [11]. Also, the existence of a maximum entropy (in another sense) solution is proved in [7] for an operator-valued truncated trigonometric moment problem when $n = 2$ and $A =$ an infinite band in $\mathbb{Z}^2$.

The full moment problems (when $A = \mathbb{Z}_+^n$ or $\mathbb{Z}^n$) received satisfactory answers for a large class of supports T, see [10], [13], [14], [18], [24], [25]. In this case one can complete the space of all polynomials with respect to the inner product $(t^\alpha, t^\beta) := \gamma_{\alpha+\beta}$ to a Hilbert space H, then study the operators induced on H by the multiplication with the coordinate functions t_i [5], [12], [14], [25]. However, as noted in [11], this idea has a limited applicability in the truncated case, when H is not invariant under t_i.

Thus the moment problems for $n > 1$ and $A =$ finite have received rather partial answers. For heuristic reasons and without claim of rigor, we sketch below the idea of our approach, based on Shannon's idea. Suppose that

$m(T) = \gamma_0 = 1$ and set $F_\alpha(f) = \int_T u_\alpha(t) f(t) dm - \gamma_\alpha$ $(\alpha \in A)$. Assume the existence of the representing densities $f \geq 0$ of γ with $\int_T f dm = 1$. Then among them there exists one probability density f_* having the maximum degree of randomness allowed by the conditions (1). (This implies in particular that f_* is strictly positive and smooth on T.) Namely, this density maximizes the entropy functional $H(f) := -\int_T f \ln f \, dm$ with the restrictions $F_\alpha(f) = 0$ $(\alpha \in A)$. Since $f_* > 0$ on T, it belongs in a certain sense to the interior of the domain of H. Hence we may apply the method of the Lagrange multipliers for the conditioned extremum. Then there are $x_\alpha \in \mathbb{R}$ $(\alpha \in A)$ such that f_* be a critical point of the function $L := H + \sum_{\alpha \in A} x_\alpha F_\alpha$; namely, $L'(f_*) = 0$. Thus $L'(f_*)g = 0$ for all g. Note that $L(f) = \int_T G(f) \, dm$ where $G(f) := -f \ln f + \sum_{\alpha \in A} (u_\alpha f - \gamma_\alpha)$. By using the formula $L'(f)g = \lim_{s \to 0} s^{-1}(L(f + sg) - L(f))$, it follows that $\int_T G'(f_*) g \, dm = 0$ for all g. Hence $G'(f_*) = 0$. Now $G'(f) = -\ln f - 1 + \sum_{\alpha \in A} x_\alpha u_\alpha$. We obtain that ef_* is the exponential of a linear combination of the functions $u_\alpha = u_\alpha(t)$ $(\alpha \in A)$. Writing the conditions (1) for $\mu := f_* m$, we obtain the system (2) which must then have a solution $x = (x_\alpha)_{\alpha \in A}$. In the 2nd section we will rigorously state and prove these considerations.

The idea from above is a known natural approach to this type of problems, at least in the case $T =$finite. One way or another, it was also used or suggested in several problems in which maximum entropy distributions naturally arise [9], [17], [19], [20]. However, to our knowledge there are no proofs of the existence of the maximum entropy solution f_* in the present context. The form of $f_* = f_{*,x}$ was known for $n = 1$ and $T \subset \mathbb{R}$ in some particular cases [9], [17], [20] (see Theorem 0 below), but the existence of the corresponding x satisfying a system of the form (2) is always—at least implicitly— assumed by hypothesis. Then it is verified that the (already existing) solution $f_* = f_{*,x}$ defined by $f_{*,x}(t) := \sum_{\alpha \in A} x_\alpha u_\alpha(t)$ $(t \in T)$ maximizes H among all other possible (nonnegative) solutions ([20], see the proof of (3)$\Rightarrow$(2) in Theorem 14).

The novelty in the present paper, for any $n \geq 1$, is that Theorems 12, 14, 16 are mainly results of existence of the maximum entropy representing density $f_* = f_{*,x}(t)$ given by the parameters x. Namely, we assume only that there is at least one (arbitrary) nonnegative representing density. Then we prove the existence of the solution x of (2), and hence of the associated density $f_{*,x}$ of maximum entropy. It suffices actually (see Theorem 12) to assume that there exist representing measures whose absolutely continuous parts are $\not\equiv 0$.

To briefly recall the significance of the maximum entropy solution [9], [17], [19], let $V : (\Omega, \mathcal{A}, P) \to (T, m)$ be a random variable with values in T and absolutely continuous repartition $P \circ V^{-1} = \mu = fm$, where $(\Omega, \mathcal{A}, P)$ is a probability field. Let T be finite with $m :=$ the normalized cardinal measure. The average of the minimum amount of information necessary to determine the position of V in T proves then to be equal to Shannon's entropy

$$H(f) := -\int_{\Omega} \log_2 f(V(\omega))\, dP(\omega) \quad (= -\sum_{t\in T} f(t)\log_2 f(t)),$$

see for instance [17]. In general, if T is endowed with some arbitrary non-negative measure m, then the corresponding degree of randomness of V is measured by

$$H(V) := -\int_{\Omega} \ln f \circ V\, dP \quad (= -\int_T f \ln f\, dm).$$

Suppose that the repartition f of V is unknown, but we can find the mean values of some quantities u_α, $\alpha \in A$ depending on V. The available data on V are thus given by the knowledge of the numbers

$$\gamma_\alpha := \int_{\Omega} u_\alpha(V(\omega))\, dP(\omega) \quad (= \int_T u_\alpha(t) f(t)\, dm(t)) \quad (\alpha \in A).$$

The problem is now to choose the most reliable f by using all this (and only this) information. The repartition f_* of the highest degree of randomness allowed by the conditions (1) is then the natural choice for f, see for instance [17], [19] for details. Note also in this sense the very interesting result from below.

Theorem 0. [9] *Let $n := 1$ and $T := [a,b] \subset \mathbb{R}$. Let V be a random variable with uniform distribution on T. If $V_1, V_2, \ldots$ are independent copies of V, then the conditional probability of V given the observation*

$$k^{-1}\sum_{i=1}^{k} u_\alpha(V_i) = \gamma_\alpha \qquad (\alpha \in A,\ k = 1, 2, \ldots)$$

converges to $f_{*,x}$ as $k \to \infty$.

Therefore in certain moment-type problems it could be of interest to approximate $f_{*,x}$ (that is, $x \in \mathbb{R}^a$). To this aim, one could minimize numerically the function V from Remark 1, using projected gradients for example. In practice, for a large class of such problems x can be obtained from a linear system of na equations if sufficiently many moments are given, see Theorem 17. The present approach seems to present certain similarities to various problems of reconstructing the shape or probability density of an object from indirect measurements providing its moments.

2 Main Results

Remark 1. If $\gamma_\alpha \in \mathbb{R}$, and $u_\alpha : T \to \mathbb{R}$ are linearly independent continuous functions, then the functional $V = V(x)$ defined for $x = (x_\alpha)_{\alpha \in A}$ by

$$V(x) = \int_T \exp\big(\sum_\alpha x_\alpha u_\alpha(t)\big)\, dm(t) - \sum_\alpha \gamma_\alpha x_\alpha$$

is smooth, strictly convex and $\frac{\partial V}{\partial x_\beta}(x) = \int_T u_\beta(t) \exp(\sum_\alpha x_\alpha u_\alpha(t)) dm(t) - \gamma_\beta$. Moreover, a solution x of (2) exists if and only if V is bounded from below and attaints its minimum – in which case the (unique) minimum point of V is x.

Proof. For every real numbers c_α we have

$$\sum_{\beta,\delta} \frac{\partial^2 V}{\partial x_\beta \partial x_\delta}(x) c_\beta c_\delta = \int_T \big(\sum_\beta c_\beta u_\beta(t)\big)^2 \exp\big(\sum_\alpha x_\alpha\, u_\alpha(t)\big)\, dm(t) \ \geq 0,$$

with equality iff all $c_\alpha = 0$. Use also the equality

$$V(y) - V(x) = V'(x)(y-x) + \frac{1}{2} V''(x)(y-x, y-x) + o(\|y-x\|^2),$$

showing as it is known that any critical point x of the (strictly convex) function V is a local (and hence global) minimum point. We omit the details. $\square$

Let then T be a compact subset of $\mathbb{R}^n$, with nonempty interior. Let m denote n–dimensional Lebesgue measure. We assume that $m(T) = 1$ and $m(\partial T) = 0$, where ∂T denotes the boundary of T. Let $A \subset \mathbb{Z}_+^n$ be a finite set with $0 \in A$. Let $\gamma_\alpha \in \mathbb{R}$ for $\alpha \in A$, and set $\gamma := (\gamma_\alpha)_{\alpha \in A}$. We assume $\gamma_0 = 1$. Set $u_\alpha(t) := t^\alpha = t_1^{\alpha_1} \ldots t_n^{\alpha_n}$ for $t = (t_1, \ldots t_n) \in T$ and $\alpha = (\alpha_1, \ldots, \alpha_n) \in A$.

Notation. Let P_A be the linear space of polynomial functions on T generated by the monomials u_α, $\alpha \in A$. Consider on P_A the restriction of the supremum norm $\mid p \mid_\infty := \max_{t \in T} \mid p(t) \mid$ $(p \in \mathbb{R}[X])$. Let P_A^* denote the dual of P_A. For $t \in T$, let $\delta_t \in P_A^*$ denote the Dirac functional $\delta_t p := p(t)$, $p \in P_A$. For $b = (d_\alpha)_{\alpha \in A} \in (P_A^*)^a$, define the operator B_b on P_A by $B_b u_\alpha = \sum_{\beta \in A} d_\alpha(u_\beta) u_\beta$ $(\alpha \in A)$. If $d_\alpha = \delta_{t(\alpha)}$ for some set $\tau := (t(\alpha))_{\alpha \in A}$ of points $t(\alpha) \in T$ $(\alpha \in A)$, then we set $\det(\tau)$ =the determinant of B_b.

Remark 2. [2] Let $b := (d_\alpha)_{\alpha \in A} \in (P_A^*)^a$. Then b is a basis iff B_b is invertible. There exist algebraic bases b of the form $b = (\delta_{t(\alpha)})_{\alpha \in A}$ with $t(\alpha) \in T \setminus \partial T$, and so $\{\tau \mid \det(\tau) \neq 0\}$ is dense in T^a.

Theorem 3. [2] *Let $f \in L^1(T) \setminus \{0\}$ be nonnegative almost everywhere on T. Let $(\gamma_\alpha)_{\alpha \in A}$ be a finite set of power moments of the measure $\mu = fm + \nu$, where $\nu \geq 0$ is a singular measure. Namely, $\gamma_\alpha = \int_T u_\alpha \, d\mu \quad (\alpha \in A)$. Then there exists $g \in C^\infty(\mathbb{R}^n)$ with $\inf_T g > 0$ such that $\gamma_\alpha = \int_T u_\alpha g \, dm \quad (\alpha \in A)$.*

Notation. Let $\gamma = (\gamma_\alpha)_{\alpha \in A}$ be a fixed sequence of power moments, having at least one representing density $f \in L^1_+(T)$ (nonnegative and Lebesgue integrable on T). We shall prove then the existence of a representing density of maximum entropy. To this aim, we may assume by Theorem 3 that f is Riemann integrable and bounded (from below and above) on T by some finite and strictly positive constants. From now on, we fix such a function f. In what follows, the symbol c (or c_1, $c_2 \ldots$) will always stand for a finite positive constant depending only on the initial data of the problem (like n, T, f, A, γ). We endow $\mathbb{R}^n$ with the norm $\mid t \mid := \max_{l=1}^n \mid t_l \mid$ for $t = (t_l)_{l=1}^n \in \mathbb{R}^n$. Then call *cube* any set $K = K(t_0, s) := \{t; \mid t - t_0 \mid \leq s\}$ where $t_0 \in \mathbb{R}^n$ and $s > 0$; its *side length* σ_K, resp. *diameter* are of course $2s$, resp. $2s\sqrt{n}$.

Lemma 4. *There exists $c_1 = c_1(n, T) > 0$ such that for any $\epsilon > 0$ there are:*

(1) *a finite set of cubes $K_i \subset T$, $i \in \{1, \ldots, N\}$ of the same size, such that $m(K_i \cap K_j) = 0$ for any $i \neq j$, the diameter of each K_i is less than ϵ and $m(T \setminus (\cup_{i=1}^N K_i)) < \epsilon$;*

(2) *a set $\tau = \{t(i)\}_{i=1}^N$ of points $t(i) \in K_i$, such that for any $i \neq j$ we have*

$$\mid t(i) - t(j) \mid > c_1 \, N^{-1/n}, \tag{3}$$

and for any distinct points $t^1, \ldots, t^a$ of the set τ, $\delta_{t^1}, \ldots, \delta_{t^a}$ is a basis of P_A^*. Moreover, $c_2 \leq N^{1/n} \, \sigma_{K_i} \leq c_3$ for some constants $c_2 = c_2(T)$, $c_3 = c_3(T)$.

PROOF. Take two cubes K, K', that we fix from now on, with $T \subset K$ and K' in the interior of T. More precisely, take $K' = \prod_{i=1}^n [a_i, a_i + b] \subset int\, T$. Let (e_i) be the canonical basis of $\mathbb{R}^n$. Then let K be the union of the $(2k_0 + 1)^n$ translates $K' + b \sum_{i=1}^n j_i e_i$ of K' over all $j_1, \ldots, j_n \in \mathbb{Z}$ with $|j_i| \leq k_0$, where $k_0 \in \mathbb{N}$ is fixed large enough so that $T \subset K$. Let $\epsilon > 0$ be arbitrary. Then take an integer $l' = l'_\epsilon \geq 1$ sufficiently big such that the cubes of sides b/l' have diameters less than ϵ. We divide each side of K into $l := (2k_0 + 1)l'$ equal compact intervals. Consider the corresponding cubic subdivision $\mathcal{K} = \mathcal{K}_\epsilon$ of K consisting of all the l^n products of n such intervals. Then choose those cubes of $\mathcal{K}$, denoted by $K_i = K_{i,\epsilon}$ for $i \in \{1, \ldots, N\}$ with $N = N_\epsilon$, that are included in T. If l' is sufficiently large (1) is fulfilled. Moreover we can assume $K' \subset \cup_{i=1}^N K_i$. For each $i \in \{1, \ldots, N\}$ let $t(i)$ be, for the moment, the center of K_i. We have $N \geq l'^n$, since here exist at least those cubes K_i covering K'. Set $c_1 = b/2$. Hence

$$\inf_{i \neq j} \mid t(i) - t(j) \mid = b/l' \geq b N^{-1/n} > c_1 \, N^{-1/n}.$$

Now let S^1, $S^2, \ldots$ be an enumeration of all $N!/\,a!(N{-}a)!$ subsets $\{\delta_{t^1}, \ldots, \delta_{t^a}\}$ of $\{\delta_{t(1)}, \ldots, \delta_{t(N)}\}$ having a elements. Apply successively Remark 2 for each S^j. That is, by successive small perturbations of the tuple $(t(1), \ldots, t(N))$ we can insure that all S^j are basis of P_A^*, and the (strict) inequality (3) still holds. Now the side length $\sigma_{K_i} = b/l'$ of the cubes K_i in the subdivision $\mathcal{K}$ of K is $\geq 2^{-1}c_1\,N^{-1/n}$, see (3) and note that $t(i) \in K_i$. Thus $N^{1/n}\sigma_{K_i} \geq c_1/2$. Also, N is less or equal than the cardinal l^n of $\mathcal{K}$, which gives an estimate of the form $N^{1/n}\sigma_{K_i} \leq c_3$ using $l = (2k_0+1)l'$. □

Notation. For any $\epsilon > 0$, we let τ be a set $(t(i))_{i=1}^N$ of points $t(i) \in K_i$ given by Lemma 4. (we shall sometimes omit stating the dependence on ϵ.) Let $P = P_\epsilon$ be the measure on $\mathbb{R}^n$ defined by

$$P = N^{-1}\sum_{i=1}^{N}\delta_{t(i)}. \tag{4}$$

Lemma 5. *There exists $c_4 = c_4(T)$ such that for any $\epsilon > 0$ and any approximation $\cup_{i=1}^N K_i$ of T, with points $t(i) \in K_i$ as in Lemma 4, we have*

$$P(K(t,r)) \leq c_4\,r^n \quad (t \in \mathbb{R}^n,\, r > \sigma_{K_i}). \tag{5}$$

Proof. We follow the proof of Lemma 4. The side length of $K(t,r)$ is $2r$. Also, the side length of any K_i is $\geq c_2N^{-1/n}$. Hence the number ν of those cubes K_i in the subdivision $\mathcal{K}$ of K such that $K_i \cap K(t,r) \neq \emptyset$ is of order $(2r/c_2N^{-1/n})^n$; namely, we have an estimate of the form $\nu \leq c\,(2r/c_2N^{-1/n})^n$. By (4) we get

$$P(K(t,r)) \leq N^{-1}\nu \leq N^{-1}c\,(2r/c_2N^{-1/n})^n = 2^n c c_2^{-n}\,r^n. \qquad \square$$

Lemma 6. *There are $c_5 = c_5(T)$ and $c_6 = c_6(T,a)$ such that for any $(\mathcal{K},\tau)$ as in Lemma 4 and any subset T' of $\{t(1), \ldots t(N)\}$ with $P(T') > c_6N^{-1}$ there exist distinct points $t^i \in T'$, $i = 1, \ldots, a$ with the property that for any $i \neq j$ there is a coordinate index $l = l_{i,j} \in \{1, \ldots, n\}$ with*

$$\mid (t^i)_l - (t^j)_l \mid \geq c_5\,(P(T')/a)^{1/n}. \tag{6}$$

Proof. We have $\sigma_{K_i} \leq c_3N^{-1/n}$, see Lemma 4. Set $c_5 = c_4^{-1/n}$, see Lemma 5, and $c_6 = ac_3^nc_4$. For $p := P(T')$, set $r = (p/ac_4)^{1/n}$. Since $p > c_6N^{-1}$, $r > \sigma_{K_i}$. Then (5) holds for any ball of radius r. We can assume $a \geq 2$. We prove by induction on $2 \leq m \leq a$ that there are $t^1, \ldots, t^m \in T'$ such that the estimate (6) holds for any $i,\ j \in \{1, \ldots, m\}$ with $t^i \neq t^j$, for some $l_{i,j}$. Let

$m := 2$. Let $t^1 \in T'$ be arbitrary. If $T' \subset K(t^1, r)$, then by the inequalities (5) we have

$$p = P(T') \le P(K(t^1, r)) \le c_4\, r^n = p/a,$$

whence $a \le 1$ which is false. Hence there is $t^2 \in T' \setminus K(t^1, r)$. Since $\mid t^2 - t^1 \mid \ge r$, there exists an index $l = l_{1,2}$ such that

$$\mid (t^2)_l - (t^1)_l \mid \ge (p/ac_4)^{1/n} = c_5(p/a)^{1/n}$$

Now let $m < a$. Assume that we have found $t^1, \dots, t^m \in T'$ such that the estimates (6) hold for $i, j = 1, \dots, m$ with $t^i \ne t^j$, then show that they hold for any $i,\, j \in 1, \dots, m+1$ with $t^i \ne t^j$. If $T' \subset \cup_{i=1}^m K(t^i, r)$, then again (5) gives

$$p = P(T') \le P(\cup_{i=1}^m K(t^i, r)) \le m \max_{i=1}^m P(K(t^i, r)) \le m\, c_4\, r^n = mp/a,$$

whence $a \le m$ which is false. Hence there is $t^{m+1} \in T' \setminus \cup_{i=1}^m K(t^i, r)$. Since all $\mid t^{m+1} - t^i \mid \ge r$, for each $i \in \{1, \dots, m\}$ there exists an index $l = l_{i,m+1}$ with

$$\mid (t^{m+1})_l - (t^i)_l \mid \ge (p/ac_4)^{1/n} = c_5(p/a)^{1/n}. \qquad \square$$

Lemma 7. *There exists $c_7 = c_7(T)$ such that the following holds. For any $(\mathcal{K}, \tau)$ as in Lemma 4 and T' as in Lemma 6, let $t^1, \dots, t^a \in T'$ satisfying (6). Let $v_i \in \mathbb{R}$ for $i \in \{1, \dots, a\}$. For any $i, i_0 \in \{1, \dots, a\}$ with $i \ne i_0$, fix $l = l_{i,i_0}$ such that $\mid (t^{i_0})_l - (t^i)_l \mid \ge c_5\,(P(T')/a)^{1/n}$. Then set*

$$p_{i,i_0}(x) = (x_l - (t^i)_l)(t^{i_0})_l - (t^i)_l)^{-1} \quad (i \ne i_0,\, x \in T,\, l = l_{i,i_0}),$$

$$L_{i_0}(x) = \prod_{1 \le i \le a,\, i \ne i_0} p_{i,i_0}(x) \quad (x \in T).$$

Let $L = \sum_{i=1}^a v_i L_i$. Then $L(t^i) = v_i$ for $i \in \{1, \dots, a\}$ and

$$|L|_\infty \le c_7\, P(T')^{(1-a)/n} \cdot \max_{i=1}^a |v_i|. \tag{7}$$

PROOF. One can assume $\mid v_i \mid \le 1$. We have $L_i(t^j) = \delta_{ij}$ for $i, j = 1, \dots, a$. Then $L(t^i) = v_i$. Applying the estimates (6) to the denominators of p_{i,i_0}, we obtain

$$|p_{i,i_0}|_\infty \le 2 \max_{t \in T} \mid t \mid\; c_5^{-1}(a/P(T'))^{1/n}.$$

Hence $|L_{i_0}|_\infty \le c_5^{1-a} a^{\frac{a-1}{n}} P(T')^{(1-a)/n}$. Use also $|L| \le a \max_i |L_i|$. $\square$

Notation. We will apply Lemmas 4–7 for $\epsilon = \epsilon_k := 1/k$ $(k = 1, 2, \ldots)$. Then for each $k \geq 1$ we fix some cubes $K_i = K_{i,k}$ $(i \in \{1, \ldots, N\})$, where $N = N_k$, as well as a set $\tau = \tau_k = (t(i))_{i=1}^N$ of points $t(i) = t(i,k) \in K_i$ satisfying (3) with c_1 independent of k. Let $P = P_k$ be defined by (4). Set $m_k := m(K_i)$ and $x_i^0 = x_i^0(k) := f(t(i,k))$ for $i \in \{1, \ldots, N\}$. Set $c_8 = 2^{-1} \inf_T f$ (> 0) and $c_8' = \sup_T f$. Then $c_8 \leq x_i^0 \leq c_8'$. Set $\rho = \rho(k) := m_k \sum_{i=1}^{N_k} x_i^0(k)$,

$$a_{i\alpha} = a_{i\alpha}(k) := t(i,k)^\alpha m_k \quad (\alpha \in A).$$

Let S denote the N–dimensional simplex

$$S = S_k := \{x = (x_i)_{i=1}^N \in \mathbb{R}^N \mid x_i \geq 0, \sum_{i=1}^N x_i = \rho\}.$$

Set $D = D(k) := \{x \in S \mid x_i > 0, i = 1, \ldots, N\}$. Define $H = H_k : S \to \mathbb{R}$ by

$$H(x) := -\sum_{i=1}^N x_i \ln x_i \quad (x \in S)$$

(with $0 \ln 0 := 0$). Then $\sum_{i=1}^N a_{i\alpha} x_i^0$ is the Riemann sum of the function $u_\alpha f$ corresponding to $(\mathcal{K}, \tau)$ and so

$$\left| \sum_{i=1}^N a_{i\alpha} x_i^0 - \gamma_\alpha \right| \to 0 \text{ as } k \to \infty \tag{8}$$

because $\int_T u_\alpha f dt = \gamma_\alpha$. If k_0 is sufficiently large, then for $k \geq k_0$ we have

$$\rho(k) \leq 2; \quad 2^{-1} \leq m_k N_k \leq 2; \tag{9}$$

the first estimate holds by (8) for $\alpha := 0$ since $\gamma_0 = 1$ and $\rho = \rho(k)$ is the Riemann sum of f, whence $\rho(k) \to 1$. For the second one, note that by Lemma 4 (1) we get, for $\epsilon := 1/k$, that $m_k N_k \to m(T) = 1$.

The index $k \geq 1$ from above will be fixed and omitted in the Lemmas 8–11.

Lemma 8. *Define the functions F_α $(\alpha \in A)$ on S by $F_\alpha(x) = \sum_{i=1}^N a_{i\alpha}(x_i - x_i^0)$. Take $Z_\alpha = \{x \in S \mid F_\alpha(x) = 0\}$. Set $K = \bigcap_{\alpha \in A} Z_\alpha$. Then there are $y = (y_i)_{i=1,\ldots,N} \in K \cap D$ and $\lambda_\alpha \in \mathbb{R}$ $(\alpha \in A)$ such that $H(y) = \max_{x \in K} H(x)$ and*

$$y_i = \exp\left(\sum_{\alpha \in A} \lambda_\alpha a_{i\alpha}\right) \quad (i = 1, \ldots, N). \tag{10}$$

PROOF. Note that $x^0 = (x_i^0)_i \in K \cap D$. Since H is continuous and $K \neq \emptyset$ is compact, there is $y \in K$ with $H(y) = \max_K H$. We prove that $y \in D$. Since $x^0 \in D$, we can assume $y \neq x^0$. Set $h(s) = H(sx^0 + (1-s)y)$ for $s \in [0,1]$. Then h is of class C^1 on $(0,1)$ and continuous on $[0,1]$. For $0 < s < 1$,

$$h'(s) = -\sum_{i=1}^{N} (\ln(sx_i^0 + (1-s)y_i) + 1)(x_i^0 - y_i). \tag{11}$$

Then $h'(s) = \sigma_1(s) + \sigma_2(s)$, where $\sigma_1(s) = -\sum_{i \mid y_i = 0} (\ln sx_i^0 + 1)x_i^0$ (≥ 0) and

$$\sigma_2(s) = -\sum_{i \mid y_i > 0} (\ln(sx_i^0 + (1-s)y_i) + 1)(x_i^0 - y_i).$$

Suppose $y \notin D$. There is i_0 with $y_{i_0} = 0$. Then the sums $\sigma_1(s)$, $0 < s < \epsilon$ have at least one term > 0, if $\epsilon > 0$ is sufficiently small. Hence $\lim_{s\to 0} \sigma_1(s) = +\infty$. Also, $\lim_{s\to 0} \sigma_2(s) = \sigma_2(0)$. Then $\lim_{s\to 0} h'(s) = +\infty$. Hence $(h(s) - h(0))/s > 0$ if $s > 0$ is enough small. Then $sx^0 + (1-s)y \neq y$ since $y \neq x^0$ and

$$H(sx^0 + (1-s)y) - H(y) = s(h(s) - h(0))/s > 0.$$

This is impossible since $sx^0 + (1-s)y \in K$ and $H(y) = \max_T H$. Thus $y \in D \cap K$. Then $H(y) = \max_K H \geq \max_{D\cap K} H \geq H(y)$, whence

$$H(y) = \max\{H(x) \mid x_i > 0, i = 1, \ldots, N;\ F_\alpha(x) = 0, \alpha \in A\}.$$

Then we may apply the method of Lagrange multipliers on the (open) domain $\{x \in \mathbb{R}^N \mid x_i > 0\}$ containing y. Thus there are $\lambda_\alpha \in \mathbb{R}$ $(\alpha \in A)$ such that y is a critical point of the function $L := H + \sum_{\alpha\in A} \lambda_\alpha F_\alpha$ defined on $\{x \in \mathbb{R}^N \mid x_i > 0\}$. Using $\partial L/\partial x_i(y) = 0$ we obtain $y_i = \exp\,(-1 + \sum_{\alpha\in A} \lambda_\alpha a_{i\alpha})$, $i = 1, \ldots, N$. Then denote $(-1 + \lambda_0 a_{i0})/a_{i0}$ by λ_0. □

Remark 9. Let y, x^0 and h be as stated in the proof of Lemma 8. Then for any s in a neighborhood of the interval $[0,1]$ we have

$$h''(s) = -\sum_{i=1,\ldots,N} ((x_i^0 - y_i)^2 (sx_i^0 + (1-s)y_i)^{-1} \leq 0.$$

Thus y (resp. 0) is a point of maximum for H (resp. h) and $h'(0) = 0$. By (11),

$$-\sum_{i=1,\ldots,N} (\ln y_i + 1)(x_i^0 - y_i) = 0. \tag{12}$$

Notation. Let $\{\pi\}$ denote the number of $i \in \overline{1,N}$ for which a property π holds.

Remark 10. If y is given by Lemma 8, then for sufficiently large v

$$N^{-1}\{y_i \leq v\} \geq 1 - 4/v. \tag{13}$$

PROOF. We have $\{y_i \leq v\} + \{y_i > v\} = N$. Using $m_k \sum_{i=1}^N (y_i - x_i^0) = F_0(y) = 0$ and (9), we get also

$$\begin{aligned} \frac{v}{2N}\{y_i > v\} &\leq m_k v \{y_i > v\} \leq m_k \sum_{i|y_i>v} y_i \\ &\leq m_k \sum_{i=1}^N y_i = m_k \sum_{i=1}^N x_i^0 = \rho \leq 2. \end{aligned} \qquad \square$$

Notation. For each $k \geq 1$, let $y = y(k)$ be a solution of the maximum problem solved by Lemma 8, given by some numbers $\lambda_\alpha = \lambda_\alpha(k)$ as in (10).

Lemma 11. *There exist constants c_9, c_{10}, c_{11} depending only on T, a, f such that for any $y = (y_i)_{i=1,\dots,N}$ given by Lemma 8*

$$N^{-1}\{c_9 \leq y_i \leq c_{10}\} \geq c_{11}. \tag{14}$$

PROOF. Set $c' = 2^{-1}\min\{1/e, \min_T f\}$ and $c'' = \max\{1/e, \max_T f\}$. Let $i \in \{1,\dots,N\}$; we consider the following cases concerning the index i:

1. $c' \leq y_i \leq c''$. Then we have an estimate of the form $|\,(\ln y_i + 1)(x_i^0 - y_i)\,| \leq c$ with $c = c(T, a, f)$. (Recall that $x_i^0 = f(t(i)) \in [c_8, c_8'].$) Hence

$$\Big|\sum_{i|c' \leq y_i \leq c''} (\ln y_i + 1)(x_i^0 - y_i)\,\Big| \leq c\,\{c' \leq y_i \leq c''\}. \tag{15}$$

2. $y_i > c''$. Then $(\ln y_i + 1)(x_i^0 - y_i) \leq 0$ and so

$$\sum_{i|y_i > c''} (\ln y_i + 1)(x_i^0 - y_i) \leq 0. \tag{16}$$

3. $y_i < c'$. Then $\ln y_i + 1 \leq \ln c' + 1 < 0$, $x_i^0 - y_i \geq \min_T f - c' \geq c'$ and $-(\ln y_i + 1)(x_i^0 - y_i) \geq -(\ln c' + 1)c'$ $(=: c''')$. Hence

$$-\sum_{i|y_i < c'} (\ln y_i + 1)(x_i^0 - y_i) \geq c'''\,\{y_i < c'\}. \tag{17}$$

By the inequalities (15), (16), (17) and the equation (12), we obtain

$$c''' \{y_i < c'\} \leq -\sum_{i|y_i<c'} (\ln y_i + 1)(x_i^0 - y_i) =$$

$$\sum_{i|c'\leq y_i \leq c''} (\ln y_i + 1)(x_i^0 - y_i) + \sum_{i|y_i>c''} (\ln y_i + 1)(x_i^0 - y_i) \leq c\{c' \leq y_i \leq c''\}.$$

Then $\{y_i < c'\} \leq c'/c''' \ \{c' \leq y_i \leq c''\}$. Also, $\{y_i < c'\} + \{c' \leq y_i \leq c''\} \leq N$. Hence $\{y_i < c'\} \leq c/c''' \,(N - \{y_i < c'\})$, and so $(c'/c''' + 1)\{y_i < c'\} \leq c/c''' \, N$. Then $N^{-1}\{y_i < c'\} \leq c/(c + c''')$ and hence

$$N^{-1}\{y_i \geq c'\} \geq c'''/(c + c'''). \tag{18}$$

Set $E := \{i = 1, \dots, N \mid \ y_i \geq c'\}$ and $F = F_v := \{i = 1, \dots, N \mid y_i \leq v\}$ for $v > 0$. By (13), we have $P(F) \geq 1 - 4/v$. By (18), $P(E) \geq c'''/(c + c''')$. Now

$$P(E) + P(F) - P(E \cap F) = P(E \cup F) \leq 1.$$

Consequently

$$P(E \cap F) \geq P(E) + P(F) - 1 \geq c'''/(c + c''') + 1 - 4/v - 1.$$

Then take for instance $v := 8(c/c''' + 1)$. We obtain $P(E \cap F) \geq 1/(2cc''' + 2)$. The desired conclusion follows for $c_9 = c'$, $c_{10} = v$ and $c_{11} = 1/(2c'/c''' + 2)$. □

Theorem 12. *Let $\gamma = (\gamma_\alpha)_{\alpha\in A}$ be a finite set of power moments of a measure $\mu = fm + \nu \geq 0$, with $f \in L^1(T) \setminus \{0\}$ and ν singular with respect to m. Namely, $\int_T t^\alpha \, d\mu(t) = \gamma_\alpha$ $(\alpha \in A)$. Then there exist $x_\alpha \in \mathbb{R}$ $(\alpha \in A)$ such that*

$$\int_T t^\alpha \exp\,(\sum_{\beta\in A} x_\beta t^\beta)\, dm(t) = \gamma_\alpha \quad (\alpha \in A).$$

Proof. By Theorem 3, we can assume $\mu = fm$ with f continuous and $\min_T f > 0$. For any $x = (x_\alpha)_{\alpha\in A} \in \mathbb{R}^n$, define the function $g_\alpha(\cdot, x)$ on T by $g_\alpha(t, x) = t^\alpha \exp\,(\sum_{\beta\in A} x_\beta t^\beta)$. We apply Lemma 4 for $\epsilon := 1/k$ $(k \geq 1)$. Consider then a sequence of approximations of T by unions of cubes $\cup_{i=1}^N K_i$, with $N = N_k$ and $K_i = K_i(k)$ and a sequence $\tau = \tau_k$ of sets $\tau = (t(i))_{i=1}^N$ of points $t(i) = t(i,k) \in K_i$, $i \in \{1, \dots, N\}$ satisfying the inequalities (3) etc. For each $k \geq 1$, let $\sigma_{\alpha k}(x)$ denote the corresponding Riemann sum of $g_\alpha(\cdot, x)$ $(x \in \mathbb{R}^n)$. Namely,

$$\sigma_{\alpha k}(x) := \sum_{i=1}^{N_k} g_\alpha(t(i,k), x) m_k = \sum_{i=1}^{N} t(i)^\alpha \exp\,(\sum_{\beta\in A} x_\beta t(i)^\beta) m_k.$$

By Lemma 8, for each $k \geq 1$ there are $\lambda_\alpha = \lambda_\alpha(k)$ $(\alpha \in A)$ such that the vector $y = (y_i)_{i=1}^N$ with $y_i = \exp\,(\sum_{\alpha \in A} \lambda_\alpha a_{i\alpha})$ is in K, see (10). Take $x_\alpha = x_{\alpha k} := \lambda_\alpha(k) m_k$. Set $x = x(k) = (x_{\alpha k})_{\alpha \in A}$. Thus

$$y_i = \exp\,(\sum_{\alpha \in A} x_\alpha t(i)^\alpha), \quad i = 1, \ldots, N. \tag{19}$$

Since $y \in K$, $y \in Z_\alpha$ for all $\alpha \in A$. Using (19) and $F_\alpha(y) = 0$ we obtain

$$\sum_{i=1}^N a_{i\alpha} x_i^0 = \sum_{i=1}^N a_{i\alpha} y_i = \sum_{i=1}^N t(i)^\alpha m_k \exp\,(\sum_{\beta \in A} x_\beta t(i)^\beta) = \sigma_{\alpha k}(x(k)). \tag{20}$$

By (8) and (20), we have then sets $x(k) = (x_{\alpha k})_{\alpha \in A}$ $(k \geq 1)$ such that

$$\mid \sigma_{\alpha k}(x(k)) - \gamma_\alpha \mid \to 0 \text{ as } k \to \infty, \;\; (\alpha \in A). \tag{21}$$

We apply Lemma 11. Set

$$T' = T'_k := \{t(i) \mid i = 1, \ldots, N,\, c_9 \leq y_i \leq c_{10}\}.$$

Hence $P(T') \geq c_{11}$, see (14) and the definition (4) of P. There is $k_0 \geq 1$ such that $c_{11} > c_6 N_k^{-1}$ for all $k \geq k_0$. By the definition of T' and (10), we obtain $c_9 \leq \exp \sum_{\alpha \in A} \lambda_\alpha a_{i\alpha} \leq c_{10}$ for every $i = \overline{1, N}$ such that $t(i) \in T'$. Since $a_{i\alpha} = t(i,k)^\alpha m_k$ and $\lambda_\alpha(k) m_k = x_{\alpha k}$, we get $\sum_{\alpha \in A} x_{\alpha k} t(i,k)^\alpha \in [\ln c_9, \ln c_{10}]$. Define $q_k \in P_A$ by $q_k(t) = \sum_{\alpha \in A} x_{\alpha k} t^\alpha$ $(t \in T)$. Thus $\mid q_k(t) \mid \leq C$ for all $t \in T'_k$, where $C = \max(|\ln c_9|, |\ln c_{10}|)$. By Lemma 6, in each T'_k $(k \geq k_0)$ there exist some points $t^1 = t^1(k), \ldots, t^a = t^a(k)$ such that the estimates (6) hold. Let L be the Lagrange–type interpolation polynomial defined in Lemma 7 for $v_i := q_k(t^i)$. Then $L(t^i) = q_k(t^i)$ for $i = 1, \ldots, a$. Since δ_{t^i}, $i = 1, \ldots, a$ is a basis of P_A^* (see Lemma 4, (2)), then $L = q_k$. By (7), we obtain the estimates

$$|q_k|_\infty \leq c_7\, P(T')^{(1-a)/n} C \leq c_7 c_{11}^{(1-a)/n} C =: C' \tag{22}$$

where the $C' = C'(T, a, f)$ is independent of $k \geq k_0$. Now define ν on P_A by

$$\nu(p) = \max_{\alpha \in A} \mid x_\alpha \mid \qquad (p = p(t) = \sum_{\alpha \in A} x_\alpha t^\alpha).$$

Since T has nonempty interior, $\dim P_A = a$ and ν is a well–defined norm, see Remark 2. The norms $|\cdot|_\infty$ and $\nu(\cdot)$ are equivalent modulo some constants of type c. Then (22) implies that all the sequences $(x_{\alpha k})_{k \geq 1}$, $\alpha \in A$ are bounded. Hence by successively considering subsequences we can assume them convergent. Take $x_\alpha \in \mathbb{R}$ $(\alpha \in A)$ such that $\lim_{k \to \infty} x_{\alpha k} = x_\alpha$. Set $x = (x_\alpha)_{\alpha \in A}$.

Hence $g_\alpha(t, x(k)) \to g_\alpha(t, x)$ $(k \to \infty)$ for any $\alpha \in A$, uniformly with respect to $t \in T$ (= bounded). Since the convergence is uniform on T, we easily conclude $\sigma_{\alpha k}(x(k)) - \sigma_{\alpha k}(x) \to 0$ as $k \to \infty$ (see (9), too). Together with (21), this implies $\sigma_{\alpha k}(x) - \gamma_\alpha \to 0$. Now for each $\alpha \in A$, $\sigma_{\alpha k}(x)$ is the Riemann sum of the function $g_\alpha(\cdot, x)$. Letting $k \to \infty$, we obtain $\int_T g_\alpha(t, x)\, dm(t) = \gamma_\alpha$ for any $\alpha \in A$. Then $\int_T t^\alpha \exp\,(\sum_{\beta\in A} x_\beta t^\beta)\, dm(t) = \gamma_\alpha$ $(\alpha \in A)$. $\square$

Notation. For $f \in L^1_+(T)$ with $\int_T f\, dm = 1$, set $H(f) = -\int_T f \ln f dm \in [-\infty, +\infty)$ if either $\max(f \ln f, 0)$ or $\min(f \ln f, 0)$ has finite integral.

Lemma 13. [20] *Let $f, g \in L^1_+(T)$, $T_+ = \{t \in T | f(t) > 0\}$. If $\int_{T_+}(f-g)\, dm \geq 0$, then*

$$\int_{T_+} f\ \ln(f/g)\, dm \geq 0$$

with equality if and only if $f = g$ almost everywhere with respect to m.

Theorem 14. *The following statements are equivalent:*
(1) *there are representing densities $f \in L^1_+(T)$ of γ,*

$$\int_T t^\alpha f(t)\, dt = \gamma_\alpha \quad (\alpha \in A); \tag{23}$$

(2) *there exists a representing density $f_* \geq 0$ such that*

$$H(f_*) = \max\{\, H(f) \mid f = \text{solution of (23)}\,\};$$

(3) *there exist some real numbers x_β, $\beta \in A$ such that*

$$\int_T t^\alpha \exp\,(\sum_{\beta\in A} x_\beta t^\beta)\, dt = \gamma_\alpha \quad (\alpha \in A). \tag{24}$$

In this case:
(1') *the maximum entropy representing density f_* is unique;*
(2') *the set $x := (x_\alpha)_{\alpha\in A}$ is uniquely determined by the equations* (24)*;*
(3') *we have the identity*

$$f_*(t) = \exp\,(\sum_{\beta\in A} x_\beta t^\beta) \quad (t \in T). \tag{25}$$

PROOF. The main implication (1) $\Rightarrow$ (3) holds by Theorem 12. To prove (3) $\Rightarrow$ (2) assume the existence of a solution x of (24). We use the idea of the corresponding result from [20] concerning the case when T is an interval. Let f_*

be defined by the equality (25). By (24), we have $\int_T t^\alpha f_*(t)\,dm = \gamma_\alpha \ (\alpha \in A)$. Let f be an arbitrary nonnegative representing density for γ, such that f is Lebesgue integrable with respect to m. By Lemma 13 and the above equalities,

$$-\int_T f\,\ln f\,dm \le -\int_T f\,\ln f_*\,dm = -\int_T f(t)\sum_{\alpha\in A} x_\alpha t^\alpha dm(t) = -\sum_{\alpha\in A} x_\alpha\gamma_\alpha.$$

Then $-\sum_{\alpha\in A} x_\alpha\gamma_\alpha$ is a fixed upper bound for $H(f)$ over all f which satisfy the equalities (23). This upper bound is attained at least for $f = f_*$; namely, we have (2). Now (2) $\Rightarrow$ (1) is trivial. Therefore (1), (2) and (3) are equivalent. The uniqueness of the maximum entropy solution follows from the strict convexity of H [20]. This provides the equality (25) also. Now $x = (x_\alpha)_{\alpha\in A}$ is uniquely determined by γ, see (25) and (1') (or Proposition 15 from below). $\square$

Notation. Let Γ (resp. Γ') denote the subset of all $\gamma = (\gamma_\alpha)_{\alpha\in A} \in \mathbb{R}^a$ that have representing densities in $L^1_+(T)$ (resp. nonnegative representing measures). Assume $\gamma_0 > 0$. (The hypothesis $\gamma_0 = 1$ was not essential in Theorem 14.)

Proposition 15. *For each $\alpha \in \mathbb{Z}^n_+$ and $x = (x_\alpha)_{\alpha\in A} \in \mathbb{R}^a$, let*

$$g_\alpha(x) := \int_T u_\alpha(t)\exp\,(\sum_{\beta\in A} x_\beta u_\beta(t))\,dm(t) \tag{26}$$

and set $g := (g_\alpha)_{\alpha\in A}$. The function g is a real–analytic diffeomorphism of $\mathbb{R}^a$ onto the open convex cone Γ, and Γ' is the closure of Γ in $\mathbb{R}^a$.

PROOF. By Theorem 14, $g : \mathbb{R}^a \to \Gamma$ is surjective. Note that

$$\partial g_\alpha \,/\, \partial x_\beta = g_{\alpha+\beta}, \quad u_\alpha u_\beta = u_{\alpha+\beta}. \tag{27}$$

For $x,\, y \in \mathbb{R}^a$ we have $g(y) - g(x) = (\int_0^1 g'(sy + (1-s)x)\,ds)(y-x)\,ds$, where g' is the differential of g. If $\langle\cdot,\cdot\rangle$ denotes the scalar product in $\mathbb{R}^a$, then

$$\langle g(y) - g(x), y - x\rangle = \int_0^1 \sum_{\alpha,\beta\in A} \partial g_\alpha \,/\, \partial x_\beta(sx + (1-s)y)(y_\beta - x_\beta)(y_\alpha - x_\alpha)\,ds.$$

By using also (26) and (27), the previous equality provides

$$\langle g(y) - g(x), y - x\rangle =$$

$$\int_0^1\int_T \sum_{\alpha,\beta\in A} u_\alpha(t)u_\beta(t)\exp(\sum_{\delta\in A}(sy_\delta + (1-s)x_\delta)u_\delta(t))(y_\beta - x_\beta)(y_\alpha - x_\alpha)\,dm(t)\,ds$$

$$= \int_0^1 \int_T (\sum_{\alpha \in A} u_\alpha(t)(y_\alpha - x_\alpha))^2 \exp(\sum_{\delta \in A}(sy_\delta + (1-s)x_\delta)u_\delta(t))\, dm(t)\, ds \geq 0.$$

If $g(y) = g(x)$, the right hand side above is null. Hence the polynomial $p(t) := \sum_{\alpha \in A} t^\alpha (y_\alpha - x_\alpha)$ vanishes on T. Since $p(t) \equiv 0$, $y = x$. Thus g is injective. Let $v = (v_\alpha)_{\alpha \in A}$ with $g'(x)v = 0$. Then $\langle g'(x)v, v\rangle = 0$; namely, $\sum_{\alpha,\beta \in A} \partial g_\alpha \,/\, \partial x_\beta(x) v_\alpha v_\beta = 0$. Using (26) and (27) again, this implies

$$\int_T (\sum_{\alpha \in A} u_\alpha(t) v_\alpha)^2 \exp\,(\sum_{\delta \in A} x_\delta u_\delta(t))\, dm(t) = 0.$$

This is possible only if the polynomial $\sum_{\alpha \in A} t^\alpha v_\alpha$ is null. Hence $v = 0$. Therefore $g'(x)$ is injective for any x. Then we can apply the implicit function theorem to derive that g is a parametrization of Γ as claimed. By Alaoglu's theorem, Γ' is closed. Then $\overline{\Gamma} \subset \Gamma'$. Let $\gamma' \in \Gamma'$ be arbitrary. Then $\gamma' = (\gamma'_\alpha)_{\alpha \in A}$ with $\gamma'_\alpha = \int_T u_\alpha d\mu'$ for a nonnegative measure μ'. To prove that $\gamma' \in \overline{\Gamma}$, let $\epsilon > 0$ be arbitrary. By the Krein–Milman's theorem there exist $t'(k) \in T$ and $\lambda(k) > 0$, $k = 1, \ldots, K$ such that $\mid \sum_{k=1}^K \lambda(k) u_\alpha(t'(k)) - \gamma_\alpha \mid < \epsilon$ $(\alpha \in A)$. In particular for $\alpha := 0$ we obtain $\mid \sum_{k=1}^K \lambda(k) \mid < \gamma_0 + \epsilon$. There are $t(k) \in T \setminus \partial T$ such that $\mid t(k) - t'(k) \mid < \epsilon$ for $k = 1, \ldots, K$. We have also

$$\mid u_\alpha(t(k)) - u_\alpha(t'(k)) \mid \leq C \mid t(k) - t'(k) \mid$$

for a finite constant $C = C(T, A)$. The previous inequalities give

$$\mid \sum_{k=1}^K \lambda(k) u_\alpha(t(k)) - \gamma_\alpha \mid \leq \mid \sum_{k=1}^K \lambda(k)(u_\alpha(t(k)) - u_\alpha(t'(k)) \mid +$$

$$\mid \sum_{k=1}^K \lambda(k) u_\alpha(t'(k)) - \gamma_\alpha \mid < (\gamma_0 + \epsilon) C\epsilon + \epsilon.$$

By convoluting the measure $\sum_{k=1}^K \lambda(k) \delta_{t(k)}$ with a test function $\varphi \in C_0^\infty(\mathbb{R}^n)$ sufficiently close to δ in P_A^* and with enough small support, we obtain an absolutely continuous measure μ of support contained in $T \setminus \partial T$ and whose moments $\gamma_\alpha = \int_T u_\alpha d\mu$ satisfy an estimate of the form $\mid \gamma_\alpha - \gamma'_\alpha \mid \leq C\epsilon$. Set $\gamma := (\gamma_\alpha)_{\alpha \in A}$ $(\in \Gamma)$ and note that ϵ was arbitrary. □

Notation. Let $\mathbb{T}$ denote the unit circle. Let $(e^{i\theta_1}, \ldots, e^{i\theta_n})$ be the variable in $\mathbb{T}^n$. Let $\mathbb{T}^n$ be endowed with the measure $m := d\theta_1 \ldots d\theta_n$. Take $T \subset \mathbb{T}^n$ closed, with nonempty interior, such that $m(\partial T) = 0$. Let $A \subset \mathbb{Z}^n$ be a finite subset such that $0 \in A$ and $A = -A$. Let $\gamma = (\gamma_\alpha)_{\alpha \in A}$ with $\gamma_\alpha \in \mathbb{C}$ such that

$\gamma_{-\alpha} = \overline{\gamma}_\alpha$ for all $\alpha \in A$. Assume $\gamma_0 > 0$. Set $u_\alpha(z) := z^\alpha = z_1^{\alpha_1} \dots z_n^{\alpha_n}$ for $z = (e^{i\theta_1}, \dots, e^{i\theta_n})$ and $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n$.

Following word–for–word the lines of the case $T \subset \mathbb{R}^n$, one easily proves analogous of Theorems 12, 14 and Proposition 15 for trigonometric moments on $\mathbb{T}^n$, like Theorem 16 from below. We refer to [2] for Theorem 3 on $\mathbb{T}^n$. To deal with real moment functions, we take a partition $A = A. \cup (-A.) \cup \{0\}$. Then use in proofs the sine and cosine u_α–type functions 1, $(z^\alpha + z^{-\alpha})/2$, $(z^\alpha - z^{-\alpha})/\,2i$ $(\alpha \in A.)$. When $T \subset \mathbb{T}^n$, the notation $g(x) = \gamma$ for $x, \gamma \in \mathbb{R}^a$ of Proposition 15 means that x, γ belong to the a–dimensional $\mathbb{R}$–subspace of $\mathbb{C}^a$ defined by $x_{-\beta} = \overline{x}_\beta$, $\beta \in A$. Set $u_\alpha(e^{i\theta_1}, \dots, e^{i\theta_n}) = e^{i(\alpha_1\theta_1 + \dots + \alpha_n\theta_n)}$.

Theorem 16. *The following statements are equivalent:*

(1) *there exists a representing density $f \in L^1_+(T)$ of γ, namely*

$$\int_T e^{i(\alpha_1\theta_1 + \dots + \alpha_n\theta_n)} f(e^{i\theta_1}, \dots, e^{i\theta_n})\, d\theta_1 \dots d\theta_n = \gamma_\alpha \;\; (\alpha \in A); \tag{28}$$

(2) *there exists a representing density $f_* \geq 0$ such that*

$$H(f_*) = \max\{\, H(f) \mid f = \text{ solution of } (28)\,\};$$

(3) *there are some complex numbers x_β with $x_{-\beta} = \overline{x}_\beta$ $(\beta \in A)$ such that*

$$\int_T u_\alpha \exp\,(\sum_{\beta \in A} x_\beta u_\beta)\, d\theta_1 \dots d\theta_n = \gamma_\alpha \;\; (\alpha \in A). \tag{29}$$

In this case:

(1') *the maximum entropy representing density f_* is unique;*

(2') *the set $x := (x_\alpha)_{\alpha \in A}$ is uniquely determined by* (29) *and $x_{-\beta} = \overline{x}_\beta$;*

(3') *we have the equality $f_* = \exp\,(\sum_{\beta \in A} x_\beta u_\beta)$.*

Since the previous results were proved only for T compact, the existence of the representing density f_* on $\mathbb{R}^n$ or $\mathbb{T}^n$ is assumed below by hypotheses.

Theorem 17. *Let T be $\mathbb{R}^n$, resp. $\mathbb{T}^n$. Let γ_α $(\alpha \in \mathbb{Z}^n_+$, resp. $\mathbb{Z}^n)$ have a representing density $f_* = f_*(t)$ of the form $f_* = f_{*,x} = \exp\,(\sum_{\alpha \in A} x_\alpha u_\alpha)$ with A finite $(u_\alpha(t) = t^\alpha)$. If $T = \mathbb{T}^n$, suppose that $x_{-\alpha} = \overline{x}_\alpha$ $(\alpha \in A)$. If $T = \mathbb{R}^n$, assume $f_{*,x}$ rapidly decreasing. Set $x_\alpha = 0$ for $\alpha \notin A$. Then for any A' finite $\supset A$, the numbers x_α $(\alpha \neq 0)$ can be uniquely determined from the linear system*

$$\sum_{\beta \in A' \setminus \{0\}} \beta_j \gamma_{\alpha+\beta}\; x_\beta = v_{j\,\alpha} \quad (\alpha \in A' \setminus \{0\},\, 1 \leq j \leq n), \tag{30}$$

where $v_{j\,\alpha} = -(\alpha_j + 1)\gamma_\alpha$ *if* $T = \mathbb{R}^n$*, while* $v_{j\,\alpha} = -\alpha_j\gamma_\alpha$ *and* $A' = -A'$ *if* $T = \mathbb{T}^n$*. Also,* $x_0 = \ln(\gamma_0/\int_T \exp(\sum_{\alpha\in A'\setminus\{0\}} x_\alpha u_\alpha)dm)$*.*

PROOF. Let $T = \mathbb{R}^n$. Let $e_j \in \mathbb{Z}^n$, $j = 1, \dots, n$ denote the canonical basis. Set $x = (x_\alpha)_{\alpha\in\mathbb{Z}_+^n}$. Since $f_{*,x}(t) := \exp(\sum_{\alpha\in A'} x_\alpha u_\alpha(t))$ is a representing density of $(\gamma_\alpha)_{\alpha\in\mathbb{Z}_+^n}$ and $f_{*,x} \in \mathcal{S}(\mathbb{R}^n)$, then by the Leibniz–Newton formula

$$\int_T \frac{\partial}{\partial t_j}\,[\,u_\delta(t)\exp(\sum_{\beta\in A'} x_\beta u_\beta(t)\,]\,dt = 0 \quad (\delta \in \mathbb{Z}_+^n,\, 1 \le j \le n).$$

Compute the derivatives under the integral using $\partial u_\delta/\partial t_j = \delta_j u_{\delta - e_j}$. Then integrate term by term and use the notation $\alpha = \delta - e_j \in A' \setminus \{0\}$ to obtain (30), that therefore must be compatible. To prove that (30) is determined, it suffices to show that its only solution is 0. Let $y = (y_\alpha)_{\alpha\in A'\setminus\{0\}}$ such that

$$\sum_{\beta\in A'\setminus\{0\}} \beta_j\gamma_{\alpha+\beta}y_\beta = v_{j\,\alpha} \quad (\alpha \in A' \setminus \{0\},\, 1 \le j \le n). \tag{31}$$

For each j and α multiply (31) by $\beta_j y_\alpha$ and sum over $\alpha \in A' \setminus \{0\}$. Hence

$$\sum_{\alpha,\beta\in A'\setminus\{0\}} \gamma_{\alpha+\beta}\beta_j^2 y_\alpha y_\beta = 0 \quad (1 \le j \le n).$$

Using $\gamma_\alpha = \int_T u_\alpha f_{*,x}\,dm$ and $u_{\alpha+\beta} = u_\alpha u_\beta$, we get

$$\int_T (\sum_{\beta\in A'\setminus\{0\}} \beta_j y_\beta u_\beta(t))^2 \exp(\sum_{\epsilon\in A'} x_\epsilon u_\epsilon(t))\,dm(t) = 0.$$

Hence $\sum_{\beta\in A'\setminus\{0\}} \beta_j y_\beta u_\beta(t) = 0$ for all $t \in T$. Then $\beta_j y_\beta = 0$ for all j and β. For any $\beta \ne 0$ there is j with $\beta_j \ne 0$, and so $y_\beta = 0$. Thus $y = 0$. The same proof works as well for $T = \mathbb{T}^n$. Namely, use $\partial u_\delta/\partial\theta_j = i\delta_j u_\delta$ to obtain the corresponding system (30). We have to prove that the only set $y = (y_\alpha)_{\alpha\in A'\setminus\{0\}}$ satisfying the system (31) and the conditions $y_{-\beta} = \overline{y}_\beta$, $\beta \in A' = -A'$ is $y = 0$. Multiply (31) by $\beta_j y_\alpha$ and sum over $\alpha \in A' \setminus \{0\}$ to obtain

$$\int_T |\sum_{\beta\in A'\setminus\{0\}} \beta_j y_\beta u_\beta|^2 \exp(\sum_{\epsilon\in A'} x_\epsilon u_\epsilon)\,dm = 0 \quad (1 \le j \le n).$$

Use the linear independence of u_β in $P_{A'}$ and proceed as in the first part. $\square$

Remark 18. Suppose that we want to identify the joint repartition $fm := P \circ (V_1, \ldots, V_n)^{-1}$ of n random variables $V_1, \ldots, V_n$ with values in $\mathbb{R}$ (resp. $\mathbb{T}$), where m is the Lebesgue measure on $T := \mathbb{R}^n$ (resp. $\mathbb{T}^n$). We assume them to have an unknown degree of correlation which in some sense is finite. Namely, suppose that the density $f \in L^1(T)$ has the form $f = \exp\left(\sum_{\alpha \in A} x_\alpha u_\alpha\right)$ for a finite set of parameters x_α $(\alpha \in A)$, where $u_\alpha(t) = t^\alpha$ $(t \in T)$. More precisely, $x_\alpha \in \mathbb{R}$ for $\alpha \in \mathbb{Z}_+^n$ when $T = \mathbb{R}^n$, while $x_{-\alpha} = \overline{x}_\alpha \in \mathbb{C}$ for $\alpha \in \mathbb{Z}^n$ when $T = \mathbb{T}^n$. We can identify A (and so f) if we know sufficiently many moments

$$\gamma_\alpha := \int u_\alpha \circ (V_1, \ldots, V_n)\, dP \quad \left(= \int_T u_\alpha(t) f(t)\, dm\right)$$

of $V_1, \ldots, V_n$, by solving the (compatible and determined) linear system (30) corresponding to α, $\beta \in A' \setminus \{0\}$ for a set A' enough large (so that $A' \supset A$) and letting $x = (x_\alpha)_{\alpha \in A'}$ (due to the uniqueness we will get $x_\alpha = 0$ for all $\alpha \notin A$).

References

[1] C.-G. Ambrozie, *A maximum entropy approach to the n-dimensional truncated moment problems*, preprint, IRMA Lille, **42:14** (1997).

[2] C.-G. Ambrozie, *Representing densities of the multi-sequences of moments*, Hokkaido Math. J., to appear.

[3] C.-G. Ambrozie, *A note on moment problems*, Proc. Operator Theory Conf., **17** (2000), 21–27, Theta Found., Timisoara.

[4] N. I. Akhiezer, *The classical moment problem*, Hafner Publ. Co., New York, 1965.

[5] A. Atzmon, *A moment problem for positive measures in the unit disc*, Pacific J. Math., **59** (1975), 317–325.

[6] M. Bakonyi, T. Constantinescu, *Schur's algorithm and several applications*, Pitman Research Notes in Mathematics Series, **261**, Longman, 1992.

[7] M. Bakonyi, L. Rodman, I. M. Spitkovsky, H. J. Woerdeman, *Positive extensions of matrix functions of two variables with support in an infinite band*, C. R. Acad. Sci., Paris, Sér. **I, 323:8** (1996), 859–863.

[8] C. Berg, J. P. R. Christensen, P. Ressel, *Harmonic analysis on Semigroups*, Springer–Verlag, 1984.

[9] J. M. Van Campenhout, T. M. Cover, *Maximum entropy and conditional probability*, IEEE Trans. Inf. Theory, IT–**27** (1981), 483–489.

[10] G. Cassier, *Problme des moments sur un compact de* $\mathbb{R}^n$ *et décomposition de polinômes plusieurs variables*, J. Funct. Analysis, **58** (1984), 254–266.

[11] R. E. Curto, L. A. Fialkow, *Solution of the truncated complex moment problem for flat data*, Memoirs of the A.M.S., 1996.

[12] A. Devinatz, *A two–dimensional moment problem*, Duke Math. J., **24** (1957), 481.

[13] C. Foiaş, A. E. Frazho, *A commutant lifting approach to interpolation problems*, Operator Theory: Advances and Applications, Birkhäuser, 1990.

[14] B. Fuglede, *The multidimensional moment problem*, Exp. Math., **1** (1983), 47–65.

[15] J.–P. Gabardo, *Trigonometric moment problems for arbitrary finite subsets of* $\mathbb{Z}^n$, Trans. Am. Math. Soc., **350:11** (1998), 4473–4498 .

[16] I. V. Girsanov, *Lectures on mathematical theory of extremum problems*, Lecture Notes in Economics and Mathematical Sciences, **67**, Springer–Verlag, New York, 1972.

[17] S. Guiaşu, *Information theory with applications*, Mc–Graw Hill, London, 1977.

[18] E. K. Haviland, *On the momentum problem for distributions in more than one dimension*, I, Amer. J. Math., **57** (1935), 562–568.

[19] E. T. Jaynes, *On the rationale of maximum entropy methods*, Proc. IEEE, **70** (1982), 939–952.

[20] A. M. Kagan, Y. V. Linnuk, C. R. Rao, *Characterization problems in mathematical statistics*, New York, Wiley, 1983.

[21] G. Polya, G. Szegö, *Problems and theorems in Analysis*, II, Springer–Verlag, Berlin, 1971.

[22] M. Putinar, F.–H. Vasilescu, *Problme des moments sur les compacts semi–algébriques*, C. R. Acad. Sci. Paris, Série I, **323** (1996), 787–791.

[23] W. Rudin, *The extension problem for positive–definite functions*, Illinois J. Math. **7** (1963), 532–539.

[24] J. Shohat, J. Tamarkin, *The problem of moments*, Math. Surveys I, Amer. Math. Soc., Providence, RI, 1943.

[25] F.–H. Vasilescu, *Moment problems for multi–sequences of operators*, J. Math. Anal. Appl., **219:2** (1998), 246–259.