ON THE EXISTENCE OF PROBABILITY MEASURES WITH GIVEN MARGINALS.

Norbert Gaffke and Ludger Rüschendorf
Received: revised version: April 8, 1982

Abstract. As an immediate consequence of a minimax theorem
Ky Fan [2] proved a characterization of consistency of a system
of convex inequalities. The aim of this paper is to show that by
means of Ky Fan's theorem one can obtain simple and unified
proofs of some existence theorems on probability measures with
given marginals, avoiding in this way approximation and disinte-
gration techniques.

1. Preliminaries. In fact we shall make use of the following
special case of Theorem 1 of Ky Fan [2]:

THEOREM 1. Let K be a compact, convex set in a real topological
vector space X. Let \((F_s)_{s \in S}\) be a family of continuous, linear,
real functions on X and \((a_s)_{s \in S}\) be a family of real numbers.
Then the system

\[ F_s(x) = a_s, \ s \in S \]

is consistent on K (i.e. there exists an \(x_0 \in K\) satisfying (1)),
if and only if for any finite set of indices \(s_1, \ldots, s_m \in S\) and

AMS subject classification: 28 A 35, 46 B 27.

Key words and phrases: Convex inequalities, product space, polish
space, dual space, bounded additive set function, signed measure.
any real numbers \( a_1, \ldots, a_m \) it holds
\[
\min_{x \in \mathcal{K}} \sum_{j=1}^{m} a_j \mathcal{F}_j (x) \leq \sum_{j=1}^{m} a_j a_j \mathcal{S}_j
\]

In what follows, \((\mathcal{F}_i, \mathcal{X}_i), 1 \leq i \leq n\), are measurable spaces \((\mathcal{X}_i, \mathcal{F}_i)\) being a \( \sigma \)-field in \( \mathcal{X}_i \), and \((\mathcal{F}, \mathcal{X}) = \bigotimes_{i=1}^{n} (\mathcal{F}_i, \mathcal{X}_i)\) is their product space. The projection of \( \mathcal{F} \) onto \( \mathcal{F}_i \) is denoted by \( \pi_i \). There is given a set \( \mathcal{A} \) of nonempty subsets \( \Lambda \subset \{1, \ldots, n\} \). By \((\mathcal{F}_i, \mathcal{X}_i)\) we denote the product space of the \((\mathcal{F}_i, \mathcal{X}_i), i \in \Lambda\), and by \( \pi_i \) the projection of \( \mathcal{F} \) onto \( \mathcal{F}_i \). When \( \lambda \) is a signed measure on \((\mathcal{F}, \mathcal{X})\), then its marginals \( \pi_i \lambda \) on \((\mathcal{F}_i, \mathcal{X}_i)\) are defined by
\[
(\pi_i \lambda)(B_i) = \lambda(\pi_i^{-1}(B_i)), \quad B_i \in \mathcal{X}_i.
\]

In the first part of Section 2 (Theorem 2, Corollaries 1, 2) topological assumptions are needed. Here the \( \mathcal{F}_i \) are Polish spaces and \( \mathcal{X}_i \) are the Borel \( \sigma \)-fields. By \( \mathcal{F} \) and \( \mathcal{X} \) we denote the fields generated by the open sets in \( \mathcal{F}_i \) and \( \mathcal{X}_i \) respectively. Let \( C_B(\mathcal{F}) \) be the space of all bounded, continuous, real functions on \( \mathcal{F} \). As it is well-known, the topological dual \((C_B(\mathcal{F}))^* \) can be identified with \( rba(\mathcal{F}, \mathcal{X}) \) - the space of all regular, bounded, additive set functions on \((\mathcal{F}, \mathcal{X})\) (cf. Dunford-Schwartz [1], Theorem 2, p. 262).

The following lemma ensures countable additivity of certain elements of \( rba(\mathcal{F}, \mathcal{X}) \).

**Lemma 1.** Let \( \mathcal{F}_i \) be Polish spaces, \( 1 \leq i \leq n \), let \( u \in rba(\mathcal{F}, \mathcal{X}) \) be nonnegative, and let the marginals \( \pi_i u \) be countably additive. Then \( u \) is countably additive on \( \mathcal{F} \). Consequently, there exists a unique measure \( \overline{u} \) on \( \mathcal{X} \) that extends \( u \) and, especially, \( \int f d\overline{u} = \int f d\overline{u} \), for all \( f \in C_B(\mathcal{F}) \).

**Proof.** Let \( \mathcal{F}_i \) be the field generated by the product \( \mathcal{F} \) and \( \mathcal{X}_i \), where \( \mathcal{F}_i \in \mathcal{F}_i \), \( 1 \leq i \leq n \), and \( u_0 \) be the restriction of \( u \) on \( \mathcal{F}_i \). As in [9], Lemma 2, inner compact regularity of \( u_0 \) can be proved, and hence countable additivity of \( u_0 \). Therefore, \( u_0 \) possesses a unique extension to a measure \( \overline{u} \) on \( \mathcal{X} \). Using regularity of \( u \) and \( \overline{u} \) it is easy to see that \( u(R) = \overline{u}(R) \) for all \( R \in \mathcal{R} \).

2. Probability measures with given marginals. The problem we want to consider first is the following. Let \( \mathcal{F}_i \) be Polish spaces, \( 1 \leq i \leq n \), and \( \lambda_A \) be given probability measures on \((\mathcal{F}_A, \mathcal{X}_A)\), \( A \in \Lambda \). Furthermore, let \( \Lambda \) be a subset of \( M_1(\mathcal{F}, \mathcal{X}) \) - the set of all probability measures on \((\mathcal{F}, \mathcal{X})\). The problem is to characterize the existence of a \( \lambda^* \in \Lambda \) with marginals \( \pi_i \lambda^* = \lambda_A \), \( A \in \Lambda \).

Considering \( M_1(\mathcal{F}, \mathcal{X}) \) as a subset of \((C_B(\mathcal{F}))^* \) we assume that
\[
(3) \quad \Lambda \text{ is a convex and relatively closed subset of } M_1(\mathcal{F}, \mathcal{X}) \text{ w.r.t. the weak*-topology.}
\]

Another assumption is more technical:
\[
(4) \quad \bigcup_{A \in \Lambda} A = \{1, \ldots, n\}.
\]

**Theorem 2.** Under assumptions (3) and (4) there exists a \( \lambda^* \in \Lambda \) with marginals \( \lambda_A^*, A \in \Lambda \), if and only if
\[
(5) \quad \inf \{ \int \left( \sum_{A \in \Lambda} f_A^* \lambda_A \right) d\lambda; \lambda \in \Lambda \} \leq \sum_{A \in \Lambda} \int f_A^* d\lambda_A^*,
\]

for all \( f_A^* \in C_B(\Lambda) \), \( A \in \Lambda \).

**Proof.** The necessity of (5) is obvious. To prove sufficiency, apply Theorem 1 with \( X = rba(\mathcal{F}, \mathcal{X}), K = \Lambda \) - the weak*-closure of \( \Lambda \) in \( X \) (K is by Alaoglu's Theorem compact) - and with the system of linear equalities
\[
(1') \quad \int f_A^* \lambda_A d\lambda = \int f_A^* d\lambda_A^*, \quad f_A \in C_B(\Lambda), \quad A \in \Lambda.
\]
Then condition (2) becomes

$$(2) \quad \min \left\{ \int \left( \sum_{A \in \mathcal{A}} f_A \cdot \pi_A \right) \lambda; \lambda \in \mathcal{A} \right\} \leq \sum_{A \in \mathcal{A}} f_A d\lambda_A'$$

for all $f_A \in C_D(\mathcal{Q}_A)$, $A \in \mathcal{A}$, which is fulfilled by (5).

Therefore, Theorem 1 yields the existence of $\lambda^* \in \mathcal{A}$ satisfying

$$(1') \quad \text{Assumption (4) implies that } \pi_i^{\lambda^*} \text{, } 1 \leq i \leq n, \text{ are countably additive on } \mathcal{P}_i \text{ and from Lemma 1 and assumption (3) it follows that } \lambda^* \text{ can be chosen in } \mathcal{A} \cap M_1(\mathcal{Q}, \lambda) = \mathcal{A}.$$  

**REMARK 1.**

a) For $A$ being the set of singletons $(1)$, $1 \leq i \leq n$, Theorem 2 has been proved by Strassen \[10\], Theorem 7, under a somewhat weaker topological assumption on $\lambda$.

b) For $A = M_i(\mathcal{Q}, \lambda)$ the left hand side of (5) equals

$$\inf \left\{ \sum_{A \in \mathcal{A}} f_A \pi_A(\omega); \omega \in \mathcal{Q} \right\}, \text{ and in this case Theorem 2 has been proved by Kellerer [8], Satz 2.2, for } G_i = \mathbb{R}, 1 \leq i \leq n.$$  

**COROLLARY 1.** Let (4) be satisfied and let $F \subset \mathcal{Q}$ be closed and $0 < \varepsilon < 1$. Then there exists a probability measure $\lambda^* \in M_1(\mathcal{Q}, \lambda)$ with marginals $\lambda_A$, $A \in \mathcal{A}$, and $\lambda^*(F) \geq 1 - \varepsilon$ if and only if

$$(6) \quad (1 - \varepsilon) \inf \left\{ \sum_{A \in \mathcal{A}} f_A \pi_A(\omega); \omega \in F \right\} + \varepsilon \inf \left\{ \sum_{A \in \mathcal{A}} f_A \pi_A(\omega); \omega \in \mathcal{Q} \right\} \leq \sum_{A \in \mathcal{A}} f_A d\lambda_A' \text{ for all } f_A \in C_D(\mathcal{Q}_A), A \in \mathcal{A}.$$  

**Proof.** The set $\mathcal{A} = \{ \lambda \in M_1(\mathcal{Q}, \lambda) \geq 1 - \varepsilon \}$ is convex and relatively closed in $M_1(\mathcal{Q}, \lambda)$ (by Portmanteau's Theorem). With $f_A := \sum_{A \in \mathcal{A}} f_A \pi_A$ and $\inf f_A := \inf f_A(\omega)$ the left hand side of (5) equals

$$\inf \left\{ (\inf f_A)(\lambda(F)) + (\inf f_A)(1 - \lambda(F)) \right\} \leq$$

$$\inf \left\{ (1 - \varepsilon) \inf f_A + \varepsilon \inf f_A, \text{ if } \lambda^*(F) \geq (1 - \varepsilon) \inf f_A + \varepsilon \inf f_A \right\} \leq$$

Let now $(\mathcal{Q}_1, \mathcal{A}_1), 1 \leq i \leq n$, be arbitrary measurable spaces. By $B(\mathcal{Q}, \lambda)$ we denote the space of all bounded, $\lambda$-measurable, real functions on $\mathcal{Q}$. The topological dual $(B(\mathcal{Q}, \lambda))^*$ can be identified with $ba(\mathcal{Q}, \lambda)$ - the space of all bounded, additive set functions on $(\mathcal{Q}, \lambda)$ (cf. Dunford-Schwarz [1], Theorem 1, p. 258).

The proof of the following lemma can be omitted.

**LEMMA 2.** Let $v, u \in ba(\mathcal{Q}, \lambda)$, $u \leq v$, and let $f \in B(\mathcal{Q}, \lambda)$. Then

$$(7) \quad \max \left\{ fd\lambda; \lambda \in ba(\mathcal{Q}, \lambda), u \leq \lambda \leq v \right\} = f^+ du - f^- dv,$$

where $f^+$ and $f^-$ are the positive and negative parts of $f$.

The following theorem is due to Kellerer [7], Satz (4.2), Satz (4.3).

**THEOREM 3.** Let $v$ and $u$ be bounded signed measures on $(\mathcal{Q}, \mathcal{A})$ with $v \leq u$ and $\lambda^*_A$ be bounded signed measures on $(\mathcal{Q}, \mathcal{A}_A)$, $A \in \mathcal{A}$. Then there exists a signed measure $\lambda^*$ on $(\mathcal{Q}, \mathcal{A})$ with marginals $\lambda^*_A$, $A \in \mathcal{A}$, and $v \leq \lambda^* \leq u$, if and only if
GAFFKE-RÜSCENDORF

We now consider the problem treated in Theorem 3 in more detail, when $\mathcal{A}$ is the set of singletons $\{i\}$, $1 \leq i \leq n$, $\lambda_{\{i\}} = \lambda_i$ are probability measures and the lower bound $\nu$ is zero. This problem has a very interesting history. It has been solved for $n = 2$ by Fréchet [3], Kellerer [7], Satz 4.4, Strassen [10], Theorem 6, Hansel-Trojalic [4].

COROLLARY 2. Let $\lambda_{\{i\}}$ be probability measures on $(\Omega_i, \mathcal{X}_i)$, $1 \leq i \leq n$, and $\mu$ be a bounded measure on $(\Omega, \mathcal{X})$. Then there exists a probability measure $\lambda^*$ on $(\Omega, \mathcal{X})$ with marginals $\lambda_{\{i\}}$, $1 \leq i \leq n$, and $\lambda^* \leq \mu$, if and only if

\[
\frac{1}{n} \sum_{i=1}^{n} f_i \circ \sigma_{\{i\}} - \tau_{\{i\}} d\lambda_i - \frac{1}{n} \sum_{i=1}^{n} f_i d\lambda_i \leq 0
\]

for all nonnegative $f_i \in B(\Omega_i, \mathcal{X}_i)$, $1 \leq i \leq n$, or, equivalently, for all $f_i \in B(\Omega_i, \mathcal{X}_i)$ with values in $[0, 1]$, $1 \leq i \leq n$.

Proof. By Theorem 3 there exists a $\lambda^*$ satisfying the conditions of the corollary, if and only if

\[
\frac{1}{n} \sum_{i=1}^{n} f_i d\lambda_i \leq \frac{1}{n} \sum_{i=1}^{n} f_i \circ \sigma_{\{i\}} d\mu_i
\]

for all $f_i \in B(\Omega_i, \mathcal{X}_i)$, $1 \leq i \leq n$. Clearly (10) implies (9). For the converse implication let

\[
c_1 = \inf f_i, \quad 1 \leq i \leq n, \quad c = \sum_{i=1}^{n} c_i.
\]

With $g_i = f_i - c_i$, $1 \leq i \leq n$, (10) is equivalent to

\[
\frac{1}{n} \sum_{i=1}^{n} g_i d\lambda_i + c \leq \sum_{i=1}^{n} g_i \circ \sigma_{\{i\}} + c_i d\mu
\]

for all nonnegative $g_i \in B(\Omega_i, \mathcal{X}_i)$, $1 \leq i \leq n$, and $c \in \mathbb{R}$.

But (9) implies $\tau_{\{i\}} \circ \sigma_{\{i\}} \geq \lambda_{\{i\}}$, $1 \leq i \leq n$, and hence (11) is satisfied if $c \geq 0$. For $c < 0$ multiplication of both sides of (9) by $|c|$.
yields (11).
Finally note that the left hand side of (9) is decreased by replacing $f_i > 0$ by $\min \{1, f_i\} \leq 1 \leq n$.

**Remark 3.** As pointed out in Remark 2 the assumption that $u$ is bounded can be replaced by the weaker assumption: There exists an $i \leq n$ such that $\pi_i u$ is $\sigma$-finite (or $u$ is "rechtecksnormalem").

**Remark 4.** Of course, in (9) we can restrict to step-functions
\[ f_i = \frac{m_i}{k_1 \ldots k_n} a_{k_1}^{(i)} \ldots a_{k_n}^{(i)} \]
where $B_{k}^{(i)} \leq m_i$, $1 \leq i \leq n$. We may assume that for fixed $i$ the $B_k^{(i)}$ are disjoint and that their union is $\mathbb{H}_i$. Then, avoiding positive part operation, (9) can be reformulated as
\[ (9') \quad \min \left\{ \sum_{i=1}^{n} \delta_{k_1}^{(i)} \ldots \delta_{k_n}^{(i)} B_{k_1}^{(i)} \ldots B_{k_n}^{(i)} : H_1 \geq -1 \right\} \]
where
\[ H_1 = ((\delta, a^{(1)}), \ldots, a^{(n)}); \quad \delta = (\delta_{k_1}, \ldots, \delta_{k_n}) \geq 0, \]
\[ a^{(i)} = (a^{(i)}_{k_1}, \ldots, a^{(i)}_{k_n}) \geq 0, 1 \leq i \leq n, \delta_{k_1}, \ldots, \delta_{k_n} \geq \frac{m_i}{k_1 \ldots k_n}, \]
for arbitrarily given partitions $\Omega_1 = \bigsqcup_{k=1}^{m_i} B_{k_1}^{(i)} \ldots B_{k_n}^{(i)} \in \mathbb{H}_1$.

Clearly the minimum of the left hand side of (9') is attained at some extreme point of the polyhedral set $H_1$. The following lemma gives the extreme points of $H_1$ in the case $n = 2$.

**Lemma 3.** $(a^*, b^*, c^*)$ is an extreme point of
\[ H_1 = ((\delta, a^*), b^*); \quad \delta = (\delta_{k_1}^{(1)}, \ldots, \delta_{k_1}^{(m_1)}) \geq 0, a^{(1)} = (a^{(1)}_{k_1}, \ldots, a^{(1)}_{k_n}) \geq 0, \]
\[ b^{(1)} = (b^{(1)}_{k_1}, \ldots, b^{(1)}_{k_n}) \geq 0, \delta_{k_1}^{(1)} \geq \alpha_{k_1}^{(1)} + \beta_{k_1} - 1, \]
if and only if there are subsets $J \subseteq \mathbb{N}_p := \{1, \ldots, p\}$,
$K \subseteq \mathbb{N}_q := \{1, \ldots, q\}$, where either $J + \mathbb{N}_p$, $K + \mathbb{N}_q$, or $J = \emptyset$, $K = \emptyset$, such that
\[ (12) \quad \delta_{k_1}^{(1)} = 1 \text{ on } J \times K, a^{(1)} = 1 \text{ on } J, b^{(1)} = 1 \text{ on } K, \]
and all other $\delta_{k_1}^{(1)}, a^{(1)}, b^{(1)}$ are zero.

**Proof.** Let $(\delta, a^*, b^*)$ be an extreme point of $H_2$. Then clearly
\[ \delta_{k_1}^{(1)} = (a^{(1)}_{k_1} + \beta_{k_1} - 1) \quad \text{for all } j, k. \]
Let $J_0 := \{j \in \mathbb{N}_p : a^{(1)}_j \in \{0, 1\}\}$,
$K_0 := \{k \in \mathbb{N}_q : b^{(1)}_k \in \{0, 1\}\}$, and suppose that $J_0 \neq \emptyset$ or $K_0 \neq \emptyset$.
Then choose an $\varepsilon > 0$ with $\varepsilon < \min \{a^{(1)} - 1, b^{(1)} - 1\}$.
Let $j \in J_0$, $k \in K_0$, and define two distinct points
\[ \tilde{a}_{j} = \begin{cases} a^{(1)}_j + \varepsilon, & j \in J_0 \\ a^{(1)}_j, & j \notin J_0 \end{cases}, \quad \tilde{b}_{k} = \begin{cases} b^{(1)}_k + \varepsilon, & k \in K_0 \\ b^{(1)}_k, & k \notin K_0 \end{cases} \]
and
\[ \tilde{z}_{jk} = \begin{cases} (a^{(1)} + \beta_{k} - 1)_{j, k}, & \tilde{z}_{jk} = (a^{(1)} + \beta_{k} - 1)_{j, k} \end{cases} \]
Then \( (\delta, a, \delta) = \frac{1}{2}(\hat{\delta}, a, \hat{\delta}) + \frac{1}{2}(\bar{\delta}, a, \bar{\delta}) \), since \( a = \frac{1}{2}(a + \bar{a}) \), \( \delta = \frac{1}{2}(\hat{\delta} + \bar{\delta}) \), and

\[
\delta_{jk} = \hat{\delta}_{jk} = \bar{\delta}_{jk} \text{ if } j \in J_0, k \in K_0 \text{ or } j \notin J_0, k \notin K_0,
\]

and in the case \( j \in J_0, k \notin K_0 \) or \( j \notin J_0, k \in K_0 \) we have by the choice of \( \varepsilon \) that

\[
\frac{1}{2}(a_j + \delta_k + \varepsilon - 1) + \frac{1}{2}(a_j + \delta_k - \varepsilon - 1) + = (a_j + \delta_k - 1) + .
\]

It follows that \( J_0 = \emptyset \) and \( K_0 = \emptyset \), and, therefore, there are subsets \( J \subseteq \mathbb{N}_p, K \subseteq \mathbb{N}_q \) such that \( \delta_{jk} = 1 \) on \( J \times K \), \( a_j = 1 \) on \( J \), \( \delta_k = 1 \) on \( K \) and all other \( \delta_{jk}, a_j, \delta_k \) are zero. It can easily be seen that among these points the extreme points of \( H_2 \) are found as in the assertion.

From Remark 4 and Lemma 3 we obtain the following corollary (cf. Fréchet [3], Kellerer [7], Strassen [10], Hansel-Troalic [4]).

**COROLLARY 3.** Under the assumptions of Corollary 2 suppose \( n = 2 \). There exists a probability measure \( \lambda^* \) on \((\Omega, \mathcal{A})\) with marginals \( \lambda_1, \lambda_2 \), and \( \lambda^* \preceq \nu \), if and only if

\[
\mu(B_1 \times B_2) \geq \lambda_1(B_1) + \lambda_2(B_2) - 1
\]

for all sets \( B_i \in \mathcal{A}_i, i = 1, 2 \).

**Proof.** By Lemma 3 the inequality \((9')\) becomes

\[
\sum_{j \in J} \sum_{k \in K} \mu(B^{(1)}_j \times B^{(2)}_k) - \sum_{j \in J} \lambda_1(B^{(1)}_j) - \sum_{k \in K} \lambda_2(B^{(2)}_k) \geq -1,
\]

which is identical to \((13)\) with \( B_1 = \bigcup_{j \in J} B^{(1)}_j \), \( B_2 = \bigcup_{k \in K} B^{(2)}_k \).

**REMARK 5.** The "only if" part of Lemma 3 does not carry over to the case \( n \geq 3 \). But the points \((\delta^{(1)}, \ldots, \delta^{(n)}) \in H_n \) with \( \delta^{(1)} \in (0,1) \) and \( \delta_1^{(1)}, \ldots, \delta_n^{(1)} = \left( \sum_{i=1}^{n} a_i^{(1)} - 1 \right)_+ \) again are extreme points when \( n \geq 3 \). So a necessary condition for \((9')\) (resp. \((9)\)) in the case \( n = 3 \) is

\[
(14) \quad 2\nu(B_1 \times B_2 \times B_3) + \mu(B_1 \times B_2 \times B_3^C) + \mu(B_1 \times B_2^C \times B_3) + \mu(B_1^C \times B_2 \times B_3)
\]

\[
\geq \lambda_1(B_1) + \lambda_2(B_2) + \lambda_3(B_3) - 1, \text{ for all } B_i \in \mathcal{A}_i, 1 \leq i \leq 3.
\]

This again shows as observed by Kellerer [6], Satz 6.1, that for \( n = 3 \) it is not sufficient for \((9')\) that \( \mu(B_1 \times B_2 \times B_3) \) is not less than the lower Fréchet-bound:

\[
(15) \quad \mu(B_1 \times B_2 \times B_3) \geq \lambda_1(B_1) + \lambda_2(B_2) + \lambda_3(B_3) - 2.
\]

Condition \((15)\) is obtained from \((9')\) by choosing the points

\[
\delta_{j,k,l} = \frac{1}{2}, \quad (j,k,l) \in J \times K \times L, \quad a_j = \frac{1}{2} \text{ on } J, \quad \delta_k = \frac{1}{2} \text{ on } K,
\]

\[
\gamma_l = \frac{1}{2} \text{ on } L, \text{ and zero otherwise, which can easily be shown to be extreme points of } H_3.
\]

**References.**


Norbert Gaffke
Institut für Statistik und Wirtschaftsmathematik
RWTH Aachen
Wüllnerstr. 3
D-5100 Aachen
West-Germany

Ludger Rüschendorf
Institut für Mathematische Stochastik
Albert-Ludwig-Universität
Hebelstr. 27
D-7800 Freiburg
West-Germany