An automorphism on product measures

Let $n \geq 2$ and consider an n-tuple of sample spaces $\{X_\nu = (\Omega_\nu, \mathcal{B}_\nu)\}$. We get the product space

$$Y = \prod X_\nu$$

whose sample space is the set-theoretic product $\prod \Omega_\nu$ and Boolean σ-algebra $\mathcal{B}$ generated by $\{\mathcal{B}_\nu\}$.

0.1 Product measures. Let $\{\gamma_\nu\}$ be an n-tuple of signed measures on $X_1, \ldots, X_n$ where each γ_ν has a finite total variation. We get a unique measure γ^* on Y such that

$$\gamma^*(E_1 \times \ldots \times E_n) = \prod \gamma_\nu(E_\nu)$$

hold for every n-tuple of $\{\mathcal{B}_\nu\}$-measurable sets. We refer to γ^* as the product measure. It is uniquely determined because $\mathcal{B}$ is generated by product sets $E_1 \times \ldots \times E_n)$ with each $E_\nu \in \mathcal{B}_\nu$. When no confusion is possible we put

$$\gamma^* = \prod \gamma_\nu$$

0.2 Remark. The set of product measures is a proper non-linear subset of all measures on Y. This is already seen when $n = 2$ with two discrete sample spaces, i.e. X_1 and X_2 consists of N points for some integer N. A Every $N \times n$-matrix with non-negative elements $\{a_{jk}\}$ give a probability measure μ on $X_1 \times X_2$ when the double sum $\sum\sum a_{jk} = 1$ The condition that μ is a product measure is tha there exist N-tuples $\{\alpha_j$ and $\{\beta_k\}$ such that $\sum \alpha_\nu = \sum \beta_k = 1$ and $a_{jk} = \alpha_j \cdot \beta_k$.

The operator T_k. Consider a positive $\mathcal{B}$-measurable function k such that k and k^{-1} both are bounded functions. Let μ be a non-negative product measure on Y such that

$$\int_Y k \cdot d\mu = 1$$

Let $1 \leq \nu \leq n$ and g is some $\mathcal{B}_\nu$-measurable function. Then we have the integral

$$\text{(ii)} \qquad \int_Y g^* \cdot k \cdot d\mu$$

where g^* is the function on the product space defined by

$$g^*(x_1, \ldots, x_n)g(x_n)$$

This gives a measure denoted b $(k\mu)_\nu$ on X_ν such that (i) is equal to $\int g \cdot (k\mu)_\nu$ for all g as above. This gives the product measure

$$T_k(\mu) = \prod (k\mu)_\nu$$

It is clear that (i) entails that $T_k(\mu)$ is a probability measure on Y. denote by $\mathcal{S}_k^*$ the family of non-negative product measures satisfying (i) above, and similarly $\mathcal{S}_1^*$ is the set of product measures which at the same time are probability measures.

Theorem. *T yields a homeomorphism between S_k^* and S_1^*.*

Remark. Above we refer to the norm topology on the space of measure, i.e. if γ_1 and γ_2 are two measures on Y then the norm $||\gamma_1 - \gamma_2||$ is the total variation of the signed measure $\gamma_1 - \gamma_2$. Recall from XX that the space of meauures on Y is complete under this norm. In particular, let $\{\mu_\nu\}$ be a Cauchy sequence with respect to the norm where each $\mu_\nu \in \mathcal{S}_1$. Then there exists a strong limit μ^* where μ^* again belongs to $\mathcal{S}_1^*$ and

$$||\mu_\nu - \mu^*|| \to 0$$

0.4 A variational problem. The proof of Theorem 1 relies upon a variational problem which we begin to describe before Theorem 1 is proved in xx below. Denote by $\mathcal{A}$ the linear space of functions on Y whose elements are of the form

$$a = g_1^* + \ldots + g_n^*$$

where each g_ν^* comes from a function g_ν on X_ν as in (0.3 The exponential function e^a becomes

$$e^a = \prod e^{g_\nu^*}$$

If γ^* is a product measure with factors $\{\gamma_\nu\}$, it follows that $e^a \cdot \gamma^*$ is a product measures with factors $\{e^{g_\nu^*} \cdot \gamma_\nu\}$. Next, for every pair $\gamma \in \mathcal{S}_1^*$ and $a \in \mathcal{A}$ we set

$$W(a,\gamma) = \int_Y (e^a k - a) \cdot d\gamma$$

Keeping γ fixed we set

$$W_*(\gamma) = \min_{a \in \mathcal{A}} W(a,\gamma)$$

The main step towards the proof of Theorem xx is the following:

Proposition. *Let $\{a_\nu\}$ be a sequence in $\mathcal{A}$ such that*

$$\lim W(\gamma, a_\nu) = W_*(\gamma)$$

Then the sequence $\{e^{a_\nu} \cdot \gamma\}$ converges to a unique probability measure μ such that $T_k(\gamma) = \mu$.

The proof of Proposition xx is preceeded by the following two results.

0. x. Lemma. *Let $\epsilon > 0$ and $a \in \mathcal{A}$ be such that $W(a) \leq m_*(\gamma) + \epsilon$. Then it follows that*

$$\int e^a \cdot k \cdot \gamma \leq \frac{1+\epsilon}{1-e^{-1}}$$

Proof. For every real number s the function $a - s$ again belongs to $\mathcal{A}$ and by the hypothesis $W(a-s) \geq W(a) - \epsilon$. This entails that

$$\int e^a k \cdot d\gamma \leq \int_Y e^{a-s} \cdot k d\gamma + s \int k \cdot d\gamma + \epsilon \implies$$

$$\int (1 - e^{-s}) \cdot e^a \cdot k d\gamma \leq s + \epsilon$$

Lemma 1 follows if we take $s = 1$.

0.X Lemma. *Let γ_1 and γ_2 be a pair of probability measures on Y. Let $\epsilon > 0$ and suppose that*

$$\left| \int_Y G_\nu \cdot d\gamma_1 - \int_Y G_\nu \cdot d\gamma_2 \right| \leq \epsilon$$

hold for every $1 \leq \nu \leq n$ and every function g_ν on X_ν with maximum norm ≤ 1. Then the norm

$$||\gamma_1 - \gamma_2|| \leq n \cdot \epsilon$$

The proof is left to the reader where the hint is to make repeated use of Fubini's theorem.

Proof of Proposition XX Let $\epsilon > 0$ and consider a pair a, b in $\mathcal{A}$ such that $W(a)$ and $W(b)$ both are $\leq m_*(\gamma) + \epsilon$ where we also suppose that $\epsilon \leq 1$. Now $\frac{1}{2}(a+b)$ belongs to $\mathcal{A}$ and we get

$$2 \cdot W(\frac{1}{2}(a+b)) \geq 2 \cdot m_*(\gamma) \geq W(a) + W(b) - 2\epsilon$$

Next, notice that

$$W(a) + W(b) - 2 \cdot W(\frac{1}{2}(a+b)) = \int_Y [e^a + e^b - 2 \cdot e^{\frac{1}{2}(a+b))}] \cdot k d\gamma$$

Now we use the algebraic identity

$$e^a + e^b - 2 \cdot e^{\frac{1}{2}(a+b))}] = (e^{a/2} - e^{b/2})^2$$

It follows from (x-x) that

(iv) $$\int_Y (e^{a/2} - e^{b/2})^2 \cdot k \cdot d\gamma \leq 2\epsilon$$

Next, we notice the identity

$$|e^a - e^b)| = (e^{a/2} + e^{b/2}) \cdot |e^{a/2} - e^{b/2}|$$

Using the Cauchy-Schwarz inequality we get

$$[\int_Y |e^a - e^b| \cdot k \cdot d\gamma]^2 \leq 2\epsilon \cdot \int_Y (e^{a/2} + e^{b/2}) \cdot k \cdot d\gamma$$

By the remark in XX the last factor is bounded by a fixed constant and hence we have proved that

$$\int_Y |e^a - e^b| \cdot k \cdot d\gamma \leq C \cdot \sqrt{\epsilon}$$

where C is a fixed constant. Replacing C by C/k_* where k_* is the minimum of k we get

$$\int_Y |e^a - e^b| \cdot d\gamma \leq C \cdot \sqrt{\epsilon}$$

Since the left hand side majorizes the total variation of the signed measures $e^a \cdot \gamma - e^b/cdot\gamma$ we get Cauchy sequences with respect to the strong norm and conclude that there exists a unique limit measure μ where $M(a_\nu) \to m_*(\gamma)$ implies that

$$||e^{a_\nu} \cdot \gamma - \mu|| \to 0$$

The equality $T(\mu) = \gamma$. To show this we study a-functions in the minimizing sequence. If $\rho \in \mathcal{A}$ is arbitrary we have

$$W(a_\nu + \rho) \geq W(a_\nu) - \epsilon_\nu$$

where $\epsilon_\nu \to 0$. This gives

$$\int_Y [ke^{a_\nu}(1-\rho) + \rho] \cdot d\gamma \leq \epsilon_\nu$$

Assuming that the maximum norm $|\rho|_Y \leq 1$ we can write

$$e^\rho = 1 + \rho + \rho_1$$

where $0 \leq \rho_1 \leq \rho^2$. Then we see that (xx) gives

$$\int_Y [\rho - ke^{a_\nu} \cdot \rho] \cdot d\gamma \leq \epsilon_\nu + \int \rho_1 \cdot \gamma \leq \epsilon + ||\rho||_Y^2$$

where the last inequality follows since γ is a probability measure. The same inequality holds with ρ replaced by $-\rho$ which entails that

$$|\int_Y (ke^{a_\nu} - 1) \cdot \rho \cdot d\gamma| \leq \epsilon_\nu + ||\rho||_Y^2$$

At this stage we apply Lemma xx to the measure $(ke^{a_\nu} - 1) \cdot d\gamma$ while we use ρ-functions in $\mathcal{A}$ of norm $\leq \sqrt{\epsilon_\nu}$. This gives the following inequality for the total variation:

$$||ke^{a_\nu} - 1) \cdot \gamma|| \leq n \cdot \frac{1}{\sqrt{\epsilon}} \cdot (\epsilon + \epsilon) = 2n \cdot \sqrt{\epsilon_\nu}$$

Remark. For every positive number q and every real number α one has the inequality

$$e^q \cdot \alpha - \alpha \geq 1 + \log q$$

Conclude that

$$W(a) \geq 1 + \log k_*$$

where k_* is the minium of the positive k-function.

Introduction. Abstract measure theory is often convenient to achieve general results. Here we expose material from Beurling's article *An automorphism of product measures* where Theorem 1 is the main result. In this theorem appears a continuous function k defined on a product $Y = X_1, \ldots, X_n$ where each X_ν is a locally compact metric space. Under the assumption that there are positive real numbers $0 < a < b$ such that the range of k is confined to $[a, b]$ it will be proved that a certain operator $\mathcal{K}$ yields a homoeomorphism from the space of non-negative Riesz measures μ on Y normalized by the condition

$$\int k \cdot d\mu = 1$$

to the space of probability measures on Y. A much more involved case appears in the singular case, i.e. when $k(x)$ for example can attain arbitrary small positive values. In section 2 we discuss the singular case for a product of two locally compact metric spaces.

Schrödinger equations. A motivation for the abstract results in Section 1 come from the article *Théorie relativiste de l'electron et l'interprétation de la mécanique quantique* published 1932. In the introduction to [Beurling] the author ponits out that Schrödinger's raised a new and unorthodox question concerning Brownian motions leading to new mathematical problems of considerable interest. More precisely, consider a Brownian motion which takes place in a bounded region Ω of some euclidian space $\mathbf{R}^d$ for some $d \geq 2$. At time $t = 0$ the densities of particles under observation is given by some non-negative function $f_0(x)$ defined on Ω. The density at a later time $t > 0$ is then equal to a function $x \mapsto u(x, t)$ where $u(x, t)$ solves the heat equation

$$\frac{\partial u}{\partial t} = \Delta(u)$$

with boundary conditions $u(x, 0) = f_0(x)$ and

$$u(x, 0) = f_0(x) \quad \text{and} \quad \frac{\partial u}{\partial \mathbf{n}}(x, t) = 0 \quad \text{on} \quad \partial\Omega$$

Schrödinger took into the account the reality of quantum physics which means that in an actual experiment the observed density of particles at a time $t_1 > 0$ does not coincide with $u(x, t_1)$. He posed the problem to find the most probable development during the time interval $[0, t_1)$ which leads to the state at time t_1. His major conclusion was that the the requested density function which substitutes the heat-solution $u(x, t)$ should belong to a non-linear class of functions formed by products

$$w(x, t) = u_0(x, t) \cdot u_1(x, t)$$

where u_0 is a solution to (*) above defined for $t > 0$ while $u_1(x, t)$ is a solution to an adjoint equation

$$\frac{\partial u_1}{\partial t} = -\delta(u) \quad : \quad \frac{\partial u_1}{\partial \mathbf{n}}(x, t) = 0 \quad \text{on} \quad \partial\Omega$$

defined when $t < t_1$. This leads to a new type of Cauchy problems where one asks if there exists a unique w-function as above satisfying

$$w(x, 0) = f_0(x) \quad : \quad w(x, t_1) = f_1(x)$$

where f_0, f_1 are non-negative functions such that

$$\int_\Omega f_0 \cdot dx = \int_\Omega f_1 \cdot dx$$

The solvability of this non-linear boundary value problem was left open by Schrödinger and the search for solutions have been studied by many mathematicians. When Ω is a bounded set and has a smooth boundary one can use the Poisson-Greens function for the classical equation (*) and in this way rewrite Schrödinger's equation to a system of non-linear integral equations. We refer to page 190 in Beurling's article for details how one arrives at such integral equations and why this motivates the result in Theorem 1 below.

1. Product measures.

Let X be a locally compact metric space. Denote by $C^b(X)$ the linear space of bounded real valued functions on X which is a Banach a space equipped with the maximum norm. The linear space of real-valued Riesz measures on X with finite total variation is denoted by $\mathfrak{M}(X)$ and the subclass of non-negative measures of total mass one is denoted by $P^+(X)$. Next, consider an n-tuple $X_1, \ldots, X_n$ of locally compact spaces and let $Y = X_1 \times \ldots \times X_n$. be the product space. If $1 \leq \nu \leq n$ and $\phi \in C^b(X_\nu)$ we get the function Φ_ν on Y defined by

$$\Phi_\nu(x_1, \ldots, x_n) = \phi_\nu(x_\nu) \tag{1}$$

Then, if $\mu \in \mathfrak{M}(Y)$ we get the measure factors $\{\mu_\nu\}$ where

$$\mu(\Phi_\nu) = \mu_\nu(\phi) \tag{2}$$

hold for each $\phi \in C^b(X_\nu)$. Conversely, let $\{\mu_\nu\}$ be an n-tuple of measures on $X_1, \ldots, X_n$. Then we get their product measure μ_* where

$$\mu_*(E_1 \times \ldots \times E_n) = \prod \mu_\nu(E_\nu)$$

hold when $\{E_\nu\}$ are Borel sets in $X_1, \ldots, X_n$.

Remark. Consider the special case when each μ_ν is non-negative Then the product measure μ_* is non-negative. Let $\{\gamma_\nu\}$ be another n-tuple of non-negative measures whose product measure $\gamma_* = \mu_*$. For each fixed $1 \leq \nu \leq n$ we take $\phi \in C^b(X_\nu)$ and get

$$\mu_*(\Phi_\nu) = \prod_{j \neq \nu} \mu_(X_j) \cdot \mu_\nu(\phi)$$

A similar formula holds for γ_* and we conclude that an equality $\mu_* = \gamma_*$ gives for each ν a constant c_ν such that

$$\gamma_\nu = c_\nu \cdot \mu_\nu$$

We obtain a unique n-tuple of components representing μ_* when we choose $\{\mu_\nu\}$ so that each has total mass given by the n:th root of $\mu_*(Y)$.

The operator $\mathcal{K}$. Consider some $k(x) \in C^b(Y)$ where $a \leq k(x) \leq b$ hold for some pair $0 < a < b$. To each $\mu \in \mathfrak{M}(Y)$ we get the measure $\mathcal{K}_\mu$ on Y which satisfies

$$\mathcal{K}_\mu(\prod \phi_\nu(x_\nu)) = \prod \mu(k(x) \cdot \Phi_\nu(x))$$

for every n-tuple $\{\phi_\nu \in C^b(X_\nu)\}$. Consider in particular the case when $\mu \in P^+(Y)$ and

$$\int_Y k \cdot d\mu = 1 \tag{*}$$

Then $\mathcal{K}_\mu$ has total mass one and if $\gamma_1, \ldots, \gamma_n$ are its normalised factors we have

$$\gamma_\nu(\phi) = \mu(\Phi_\nu \cdot k)$$

when $\phi \in C^b(X_\nu)$.

Denote by $P^+_k(Y)$ the set of non-negative measures μ on Y for which (*) above holds. With these notations one has:

1. Theorem. *For each function k as above the operator $\mathcal{K}$ yields a homeomorphism from from $P^+_k(Y)$ onto $P^+(Y)$ where each of these sets are equipped with the strong topology.*

For the proof of Theorem 1 we refer to [Beurling]. At the end of the article a more involved case is studied.

A singular case. Here we restrict the attention to the case $n = 2$ and let $k(x_1, x_2)$ be a bounded and strictly positive continuous function on $Y = X_1 \times X_2$. Let $\mu \in P^+(Y)$ be such that

$$\int_Y \log k \cdot d\mu > -\infty \tag{1}$$

Under this integrability condition one has

2. Theorem. *There exists a unique non-negative measure γ on Y such that $\mathcal{K}(\gamma) = \mu$.*

Remark. In contrast to Theorem 1 the measure γ need not have finite mass but the proof shows that k belongs to $L^1(\gamma)$. Concerning the integrability condition in Theorem 2 it can be relaxed a bit, i.e. it suffices to assume that

$$\min_{s>0} \int (ke^s - s) \cdot d\mu > -\infty \tag{2}$$

As pointed out by Beurling the result in Theorem 2 can be applied to the case $X_1 = X_2 = \mathbf{R}$ both are copies of the real line and

$$k(x_1, x_2) = g(x_1 - x_2)$$

where g is the density of a Gaussian distribution which after a normalisation of the variance is taken to be

$$\frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2}$$

So the integrability condition for μ in Theorem 2 becomes

$$\iint (x_1 - x_2)^2 \cdot d\mu(x_1, x_2) < \infty$$

The proof of Theorem 2 is given on page 218-220 in [loc.cit] and relies upon the method and various estimates from the proof of Theorem 1. For higher dimensional cases, i,e, with $n \geq 3$ Beurling gives the following comments

Theorem 1 relies heavily on the condition that $k \geq a$ for some $a > 0$. If this lower bound condition is dropped the individual equation $\mathcal{K}(\gamma) = \mu$ may still be meaningful, but serious complications will arise concerning the global uniqueness if $n \geq 3$ and the proof of Theorem 2 for the case $nn \geq 3$ cannot be duplicated.