An automorphism on product measures

Introduction. The main result is Theorem xx below which was proved in [Beurling]. First we insert comments from Beurling's ariticle about the significance of Theorem XX below.

The article *Théorie relativiste de l'electron et l'interprétation de la mécanique quantique* published by Schrödinger in 1932 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 classically 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

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

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 and concluded was that the the density function which substitutes the heat-solution $u(x,t)$ should belong to a non-linear class of functions formed by products

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

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. Namely, let f_0, f_1 be a pair of non-negative functions such that

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

Then one asks if there exists a unique pair (u_1, u_2) as above such that the product function w satisfies

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

Remark. The solvability of this non-linear boundary value problem was left open by Schrödinger. When Ω is a bounded set and has a smooth boundary one can use the Poisson-Greens function for the classical equation (*) and rewrite Schrödinger's equation to a system of non-linear integral equations. An account about (eventual) mathematical solutions to Schrödinger equations was presented by I.N. Bernstein in a plenary lecture at the IMU-congress at Zürich 1932. A first example appears on the product of two copies of the real line where Schrödinger's equations lead to non-linear equation for measures which goes as follows: Consider the Gaussian density function

$$g(t) = \frac{1}{\sqrt{2\pi}} \cdot e^{-t^2/2}$$

Denote by $\mathcal{S}^*$ the class of non-negative product measures $\gamma_1 \times \gamma_2$ on $\mathbf{R}^2$ for which

$$(*) \qquad \iint g(x_1 - x_2) \cdot d\gamma_1(x_1) \cdot d\gamma_2(x_2) = 1$$

The product measure gives another product measure

$$\mathcal{T}_g(\gamma_1 \times \gamma_2) = \mu_1 \times \mu_2$$

where

$$\mu_1(E_1) \cdot \mu_2(E_2) = \iint_{E_1 \times E_2} g(x_1 - x_2) \cdot d\gamma_1(x_1) \cdot d\gamma_2(x_2)$$

hold for all pairs of bounded Borel sets. Notice that $\mu_1 \times \mu_2$ becomes a probability measure since (*) above holds. With these notations one has

Theorem. *For every product measure $\mu_1 \times \mu_2$ which in addition is a probability measure there exists a unique $\gamma_1 \times \gamma_2$ in $\mathcal{S}^*$ such that*

$$\mathcal{T}_g(\gamma_1 \times \gamma_2) = \mu_1 \times \mu_2$$

This theorem is a special case of Theorem § XX where the g-function is replaced by an arbitrary non-negative and bounded function $k(x_1, x_2)$ such that

$$\iint_{\mathbf{R}^2} \log k \, dx_1 dx_2 > -\infty$$

0. 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 its Boolean σ-algebra is generated by $\{\mathcal{B}_\nu\}$.

0.1 Product measures. Let $\{\gamma_\nu\}$ be an n-tuple of signed measures on $X_1, \ldots, X_n$. 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 and refer to γ^* as the product measure. It is uniquely determined because the Boolean σ-algebra $\mathcal{B}$ on Y by definition 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

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

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. 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 that there exist N-tuples $\{\alpha_j$ and $\{\beta_k\}$ such that $\sum \alpha_\nu = \sum \beta_k = 1$ and $a_{jk} = \alpha_j \cdot \beta_k$.

0.2 The operator T_k. With a fixed $1 \leq \nu \leq n$ we consider $\mathcal{B}_\nu$-measurable functions g_ν. Every such g_ν yields the function on Y defined by

$$g^*_\nu(x_1, \ldots, x_n) = g_\nu(x_n)$$

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

(i) $$\int_Y k \cdot d\mu = 1$$

For each ν we find a unqiue non-negative masrue on X_ν denoted by $(k\mu)_\nu$ where

(ii) $$\int_{X_\nu} g_\nu \cdot (k\mu)_\nu = \int_Y g^*_\nu \cdot k \cdot d\mu$$

hold for all bounded measure functions on X_ν. Now we construct the product measure

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

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) and let $\mathcal{S}_1^*$ be the family of product measures which at the same time are probability measures.

Main Theorem. *T_k 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 measures 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

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

Exercise. Verify that T_k is a continuous and injective map from $\mathcal{S}_k^*$ to $\mathcal{S}_1^*$. So the main theorem anounts to prove that T_k is surjective which will be achieved in § 0.4 below.

0.4 A variational problem.

The proof that T_k is surjective relies upon a variational problem. Before it is presented we insert a preliminry result which plays an essential role later on.

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

$$\Big|\int_Y g_\nu^* \cdot d\gamma_1 - \int_Y g_\nu^* \cdot d\gamma_2\Big| \leq \epsilon$$

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.

0.6 The linear space $\mathcal{A}$. It denotes the class of functions on Y of the form

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

where each g_ν^* comes from a function g_ν on X_ν as above. 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 $\{e^{g_\nu^*} \cdot \gamma_\nu\}$. For every pair $\gamma \in \mathcal{S}_1^*$ and $a \in \mathcal{A}$ we put

$$W(a, \gamma) = \int_Y (e^a k - a) \cdot d\gamma \tag{0.6.1}$$

Keeping γ fixed we set

$$W_*(\gamma) = \min_{a \in \mathcal{A}} W(a, \gamma) \tag{0.6.2}$$

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

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

It follows that $W(a, \gamma) \geq 1 + \log k_*$ where k_* is the minium of the positive k-function and hence that

$$W_*(\gamma) \geq 1 + \log k_* \tag{*}$$

The requested surjectivity of T_k follows from the following:

0.7 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 $T_k(\gamma) = \mu$.

Before we enter the proof we insert the following:

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

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

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

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

Hence Lemma 0.8 follows with $s = 1$.

Proof of Proposition 0.7

Let $0 < \epsilon < 1$ and consider a pair a, b in $\mathcal{A}$ such that $W(a,\gamma)$ and $W(b,\gamma)$ both are $\leq W_*(\gamma) + \epsilon$. The inclusion $\frac{1}{2}(a+b) \in \mathcal{A}$ gives

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

Notice that

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

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 (i-ii) that

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

Next, we notice the identity

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

The Cauchy-Schwarz inequality gives

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

Lemma 0.8 implies that the last factor is bounded by a fixed constant and hence there exists a constant C such that

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

Replacing C by C/k_* where k_* is the minimum of k we get

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

Since this hold for every ϵ it follows that when $\{a_\nu\}$ is a sequence such that

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

then $\{e^a \cdot \gamma\}$ is a Cauchy sequence and therefore converges to a limit measure μ where

(viii) $$\lim_{\nu \to \infty} ||e^{a_\nu} \cdot \gamma - \mu|| \to 0$$

0.9 The equality $T(\mu) = \gamma$

.

Consider some $\rho \in \mathcal{A}$ whose maximum norm $|\rho|_Y \leq 1$ which enable us to write

(1) $$e^{-\rho} = 1 - \rho + \rho_1 \quad : \; 0 \leq \rho_1 \leq \rho^2$$

Now

$$W(a_\nu - \rho) \geq W(a_\nu) - \epsilon_\nu \quad : \text{where} \quad \epsilon_\nu \to 0 \implies$$

(2) $$\int_Y [ke^{a_\nu} - a_\nu] \cdot d\gamma \leq \int_Y [ke^{a_\nu - \rho} - a_\nu + \rho] \cdot d\gamma + \epsilon_\nu = \int_Y [\, ke^{a_\nu} - a_\nu + e^{ka_\nu}\rho_1 + \rho(1 - e^{ka_\nu})] \cdot d\gamma + \epsilon_\nu$$

where the last equality used (1). Hence we have

(3) $$\int_Y \rho(e^{ka_\nu} - 1) \cdot d\gamma \leq \epsilon_\nu + \int_Y e^{ka_\nu}\rho_1 \cdot d\gamma$$

Next, we have a fixed constant C such that

$$\int_Y e^{ka_\nu} \cdot d\gamma \leq C$$

So (3) entails that

(4) $$\int_Y (e^{ka_\nu} - 1) \cdot \rho] \cdot d\gamma \leq \epsilon_\nu + C \cdot ||\rho||_Y^2$$

The same inequality with ρ replaced by $-\rho$ which entails that

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

At this stage we apply Lemma 0.5 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:

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

Passing to the limit we get the requested equality in § 0.9 and Proposition 0.7 is proved.