An automorphism on product measures

Introduction. The results is expose material from the article [Beurling]. Before the measure theoretic study starts we insert comments from [Beurling] about the significance of the main theorem in 0.§§ below.

Schrödinger equations. The article *Théorie relativiste de l'electron et l'interprétation de la mécanique quantique* was published 1932. Here Schrödinger 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. Classically the density at a later time $t > 0$ is 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

$$(1) \qquad u(x,0) = f_0(x) \quad \text{and} \quad \frac{\partial u}{\partial \mathbf{n}}(x,t) = 0 \quad \text{when} \quad x \in \partial\Omega \quad \text{and} \quad t > 0$$

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. He concluded 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

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

where u_0 is a solution to (1) while $u_1(x,t)$ is a solution to an adjoint equation

$$(2) \qquad \frac{\partial u_1}{\partial t} = -\Delta(u) \quad : \quad \frac{\partial u_1}{\partial \mathbf{n}}(x,t) = 0 \quad \text{on} \quad \partial\Omega$$

defined when $t < t_1$. This leads to a new type of Cauchy problems where one asks if there exists a w-function given by (*) 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 has remained as an active field in mathematical physics. 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. The interested reader should consult the talk by I.N. Bernstein a the IMU-congress at Zürich 1932 for a first account about mathematical solutions to Schrödinger equations. Examples occur already on the product of two copies of the real line where Schrödinger's equations lead to certain non-linear equation for measures which goes as follows: Consider the Gaussian density function

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

Next, consider the family $\mathcal{S}_g^*$ of all non-negative product measures $\gamma_1 \times \gamma_2$ for which

$$\text{(i)} \qquad \iint g(x_1 - x_2) \cdot d\gamma_1(x_1) \cdot d\gamma_2(x_2) = 1$$

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 (i) above holds. With these notations one has

0.1 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 S_g^* such that*

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

In [Beurling] a more general result is established where the g-function can be replaced by an arbitrary non-negative and bounded function $k(x_1, x_2)$ such that

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

1. The $\mathcal{T}$-operator and 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. There exists 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

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

The family of all such product measures is denoted by $\text{prod}(\mathcal{M}_B)$.

0.2 Remark. The set of product measures is a proper non-linear subset of the space $\mathcal{M}_B$ of all signed 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 $a_{jk} = \alpha_j \cdot \beta_k$.

0.3 The space $\mathcal{A}$. We have the linear space of functions on Y whose elements are of the form

$$a = g_1^* + \ldots + g_n^* \tag{i}$$

where $\{g_\nu\}$ are functions on the separate product factors $\{X_\nu\}$. It is clear that a pait of product measures γ and μ on Y are equal if and only if

$$\int_Y a \cdot d\gamma = \int_Y a \cdot d\mu$$

hold for every $a \in \mathcal{A}$.

0.4 The measure $e^a \cdot \gamma^*$ Let $a = \sum g_\nu^*$ be as above. Then we get the exponential function

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

If $\gamma^* = \prod \gamma_\nu$ is some product measure we get a new product measure defined by

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

0.5 The $\mathcal{T}$-operators. To every bounded $\mathcal{B}$-measurable function k we shall construct a map $\mathcal{T}_k$ from the space of product measures into itself. First, let $1 \leq \nu \leq n$ be given and g_ν is some $\mathcal{B}_\nu$-measurable function. Then there exists the function g_ν^* on the product space Y defined by

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

Let us now consider a product measure γ. If $1 \leq \nu \leq n$ we find a unique measure on X_ν denoted by $(k \cdot \gamma)_\nu$ such that

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

hold for every bounded $\mathcal{B}_\nu$-measurable function g_ν on X_ν. Now we get the product measure

$$(*) \qquad \mathcal{T}_k(\gamma) = \prod (k\gamma)_\nu$$

Remark. In the the case when

$$k(x_1, \ldots, x_n) = g_1^* \cdots g_n^*$$

we see that

$$\mathcal{T}_k(\gamma) = \prod g_\nu \cdot \gamma\nu$$

Exercise. Consider the case $n = 2$ where X_1 and X_2 both consist of two points, say a_1, a_2 and b_1, b_2 respectively. A measure $\gamma \in S_1^*$ is given by $\gamma_1 \times \gamma_2$ and we can identify this product measure by a 2×2-matrix

$$XXX$$

where $\alpha_i \cdot \beta_\nu$ is the mass of γ at the point (a_i, b_ν). Next, let k be a positive function on the product space which means that we assign four positive numbers

$$k_{i,\nu} = k(a_i, b_\nu)$$

Find the measure $\mathcal{T}_k(\gamma)$ and express it as above by a 2×2-matrix.

Now we are prepared to announce the main result in this section. Consider a positive $\mathcal{B}$-measurable function k such that k and k^{-1} both are bounded functions. Denote by $\mathcal{S}_k^*$ the family of non-negative product measures γ on Y such that

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

We have also the set $\mathcal{S}_1^*$ of product measures μ which are non-negative and have total mass one, i.e.

$$\int_Y d\mu = 1$$

It is easily seen that $\mathcal{T}_k$ yields an injective map from S_k^* into S_1^*. It turns out that the map also is surjective, i.e. the following hold:

Main Theorem. *$\mathcal{T}_k$ yields a homeomorphism between S_k^* and S_1^*.*

0.6 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 $\gamma_1 - \gamma_2$. The reader may verify that S_k^* and S_1^* both appear as closed subsets in the normed space of all signed measures on Y. Recall also 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

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

This completeness property will be used in the subsequent proof. We shall also need some inequalities which are announced below.

0.7 Some useful inequalities. Let γ_1 and γ_2 be a pair of product measures such that

$$|\int_Y g_\nu^* \cdot d\gamma_1 - \int_Y g_\nu^* \cdot d\gamma_2| \leq \epsilon \quad : \quad 1 \leq \nu \leq n$$

hold for some $\epsilon > 0$ and every function g_ν on X_ν with maximum norm ≤ 1. Then the norm

$$\text{(i)} \qquad ||\gamma_1 - \gamma_2|| \leq n \cdot \epsilon$$

The proof of (i) is left to the reader where the hint is to make repeated use of Fubini's theorem. More generally, let k be a bounded measurable function on Y and γ, μ is a pair of product measures. Denote by $\mathcal{A}_*$ the set of $\mathcal{A}$-functions a with maximum norm ≤ 1. Then there exists a constant C which only depends on k and n such that

$$(*) \qquad ||\mathcal{T}_k(\mu) - \gamma|| \leq \max_{a \in A_*} |\int_Y a(kd\mu - d\gamma)\,|$$

Again we leave the proof as an exercise.

0.8 A variational problem. Since we already have observed that $\mathcal{T}_k$ is injective there remains to prove surjectivity. For this we shall study a a variational problem which we begin to describe before the proof is finished in 0.§§ X below. We are given the function k on Y where both k and k^{-1} are bounded and the space $\mathcal{A}$ was defined in 0.3. For every pair $\gamma \in \mathcal{S}_1^*$ and $a \in \mathcal{A}$ we set

$$W(a, \gamma) = \int_Y (e^a k - a) \cdot d\gamma \quad \text{and} \quad W_*(\gamma) = \min_{a \in \mathcal{A}} W(a, \gamma)$$

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

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

Then the sequence $\{e^{a_\nu} \cdot \gamma\}$ converges to a measure $\mu \in S_1^*$ such that $\mathcal{T}_k(\gamma) = \mu$.

Before we enter the proof we insert a preliminary result which will be used later on.

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

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

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

Lemma 0.10 follows if we take $s = 1$.

Proof of Proposition 0.9 Keeping γ fixed we set $W(a) = W(a, \gamma)$. Let $0 < \epsilon < 1$ and consider a pair a, b in $\mathcal{A}$ such that $W(a)$ and $W(b)$ both are $\leq W_*(\gamma) + \epsilon$. Since $\frac{1}{2}(a + b)$ belongs to $\mathcal{A}$ we get

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

Notice that

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

Next, we have 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

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

Next, the identity $|e^a-e^b)|=(e^{a/2}+e^{b/2})\cdot|e^{a/2}-e^{b/2}|$ and the Cauchy-Schwarz inequality give:

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

By Lemma 0.6 the last factor is bounded by a fixed constant and hence (iv) gives a constant C such that

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

Next, let k_* be the minimum value taken by k on Y which by assumption is positive since k^{-1} is bounded. Replacing C by C/k_* where we get

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

Now (v) applies to pairs in the sequence $\{a_\nu\}$ and shows that $\{e^a\cdot d\gamma\}$ is a Cauchy sequence with respect to the norm of measures on Y. So from Remark 0.6 there exists a non-negative measure μ such that

$$\text{(vii)} \qquad \lim_{\nu\to\infty}||e^{a_\nu}\cdot\gamma-\mu||=0$$

The equality $\mathcal{T}_k(\mu)=\gamma$. Consider the 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

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

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

$$\text{(2)} \qquad e^\rho=1+\rho+\rho_1\quad\text{where}\quad 0\leq\rho_1\leq\rho^2$$

Then we see that (1) gives

$$\text{(3)} \qquad \int_Y (\rho-ke^{a_\nu}\cdot\rho)\cdot d\gamma\leq\epsilon_\nu+\int\rho_1\cdot\gamma\leq\epsilon+||\rho||_Y^2$$

where the last inequality follows since γ is a probability measure and the inequality in (2) above. The same inequality holds with ρ replaced by $-\rho$ which entails that

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

Notice that Lemma 0.10 entails that the sequence of functions $\{ke^{a_\nu}\}$ are uniformly bounded. Now we apply the inequality (*) from 0.7 while we use ρ-functions in $\mathcal{A}$ of norm $\leq\sqrt{\epsilon_\nu}$. It follows that there exists a constant C which is independent of ν such that the following inequality for the total variation:

$$||\mathcal{T}_k(e^{a_\nu}\cdot\gamma)-\gamma||\leq C\cdot n\cdot\frac{1}{\sqrt{\epsilon}}\cdot(\epsilon_\nu+\epsilon_\nu)=2\cdot Cn\cdot\sqrt{\epsilon_\nu}$$

Passing to the limit it follows from (vii) that we have the equality

$$\mathcal{T}_k(\mu) = \gamma$$

Since $\gamma \in S_1^*$ was arbitrary we have proved that the $\mathcal{T}_k$ yields a surjective map from S_k^* to S_1^* which finishes the proof of the Main Theorem.

0.11 The singular case.

We restrict to the case $n = 2$ where $k(x_1, x_2)$ is a bounded and strictly positive continuous function on $Y = X_1 \times X_2$. Let $\gamma \in S_1^*$ satisfy:

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

Under this integrability condition the following hold:

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

Remark. In general the measure μ need not have finite mass but the proof shows that k belongs to $L^1(\mu)$, i.e.

$$\int_Y k \cdot d\mu < \infty$$

As pointed out by Beurling Theorem 0.12 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 becomes

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

A proof of Theorem 0.12 is given on page 218-220 in [loc.cit] and relies upon similar but technically more involved methods as in the Main Theorem. Concerning higher dimensional cases, i.e. singular versions of the Main Theorem when $n \geq 3$, Beurling gives the following comments at the end of [ibid] where the citation below has changed numbering of the theorems as compared to [ibid]:

The proof of the Main Theorem 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 0.12 for the case $n \geq 3$ cannot be duplicated.