23. *Transmutation, Filtering, and Scattering*

By Robert CARROLL
Department of Mathematics, University of Illinois

(Communicated by Kôsaku YOSIDA, M. J. A., March 12, 1984)

Abstract. It is shown that in suitable circumstances the characterization of transmutation kernels via minimization can be achieved via stochastic information and accomplishes the same thing in stochastic geometry as linear least squares estimation.

1. Introduction. We will show here how the two areas of transmutation and linear filtering theory are directly connected by a minimization principle. Thus we refer first to transmutation theory as developed in [1], [2] for example where the basic theme is to study operators $B: P \to Q$ intertwining P and Q ($BP=QB$ acting on suitable functions); P and Q are two second order ordinary differential operators and B is generally an integral operator with a distribution kernel. Such transmutations B are often characterized by their action on suitable eigenfunctions φ_λ^P of P ($P\varphi_\lambda^P = -\lambda^2 \varphi_\lambda^P$) and $\varphi_\lambda^Q = B\varphi_\lambda^P$ satisfies (*) $Q\varphi_\lambda^Q = -\lambda^2\varphi_\lambda^Q$; they play an important role in the study of special functions, eigenfunction integral transforms, inverse problems, etc. In particular in classical quantum scattering theory with $P=D^2$ and $Q = D^2 - q(x)$ extensive use of transmutation methods appears in the physics literature (cf. [8]). We take this as our basic situation here also, in establishing links with estimation theory, and take $\varphi_\lambda^P(x) = \operatorname{Cos} \lambda x$ with $\varphi_\lambda^Q(x)$ defined to satisfy (*) with $\varphi_\lambda^Q(0)=1$ and $D_x\varphi_\lambda^Q(0)=h \neq 0$. Then

$$\varphi_\lambda^Q(y) = (B\varphi_\lambda^P)(y) = \operatorname{Cos} \lambda y + \int_0^y K(y,x) \operatorname{Cos} \lambda x dx$$

and $K(y,x)$ is a function with smoothness depending on q. Strictly speaking one should index with h, i.e. B_h, K_h, etc. but we omit the index h for simplicity. Also assume the spectral theory for Q ($=Q_h$) is based on a measure $d\omega(\lambda) = \omega d\lambda$ (no bound states). Now recently in [3]–[6] it was shown that various transmutations can be characterized by minimization with Gelfand-Levitan (G-L) or Marčenko (M) equations arising as Euler equations (cf. also [10]). In the same spirit K ($=K_h$) above will arise from minimizing

$$\Xi = \int_0^T \int_0^\infty \left\{ \varphi_\lambda^Q(y) - \operatorname{Cos} \lambda y - \int_0^y \mathfrak{K}(y,x) \operatorname{Cos} \lambda x dx \right\}^2 d\omega dy \tag{1.1}$$

($T < \infty$ fixed) over a suitable class of kernels $\mathfrak{K}$ having the same properties as K above. For questions of linear estimation, prediction, smoothing, and filtering there is a staggering literature (we cite here only [9], [11], [13], [14]). Various connections between inverse scattering techniques and linear estimation are known (cf. [7], [12]–[14]) so it is no surprise to have many relations between transmutation and linear filtering theory. We will spell out here what seems however to be the basic connection, namely, the minimization indicated in (1.1) can be accomplished with stochastic information and signifies the same thing in stochastic geometry as linear least squares estimation, when in fact there is an underlying stochastic problem related to Q. We note that minimizations such as (1.1) can be associated to various transmutations for more general operators as in [3]–[6] so (1.1) represents a more general situation to which there may not always be associated a stochastic model.

2. Stochastic framework. Let us sketch the framework of [13]. Thus one takes $y(t)=z(t)+v(t)$ for $-T\leq t\leq T$ to be observation of a (wide sense) stationary zero mean Gaussian signal $z(t)$ with covariance $Ez(t)z(s)=k(t-s)$ and assume $v(t)$ represents a Gaussian white noise with $Ev(t)v(s)=\delta(t-s)$ and $Ez(t)v(s)=0$. Define the even and odd processes by $y_{\pm}(t)=(y(t)\pm y(-t))/2$, etc. so that $y_{\pm}(t)=z_{\pm}(t)+v_{\pm}(t)$ with $0\leq t\leq T$. Then $Ez_{+}(t)z_{-}(s)=Ev_{+}(t)v_{-}(s)=0$ with $Ev_{\pm}(t)v_{\pm}(s)=\delta(t-s)/2$ and although $z_{\pm}$ are not stationary one has $Ez_{\pm}(t)z_{\pm}(s)=k_{\pm}(t,s)$ $=(k(t-s)\pm k(t+s))/2$. One denotes by $Y_T^{\pm}$ the Hilbert space spanned by $y_{\pm}(t)$ for $0\leq t\leq T$ (with scalar product $(f,g)=Efg$). One asks then for the best (least squares) linear estimation of $z_{\pm}$ given $Y_T^{\pm}$ in the form of a filtering estimate

$$\hat{z}_{\pm}(T\,|\,T)=E(z_{\pm}(T)\,|\,Y_T)=\int_0^T g_{\pm}(T,t)y_{\pm}(t)dt. \tag{2.1}$$

Thus for $\tilde{z}_{\pm}(T,T)=z_{\pm}(T)-\hat{z}_{\pm}(T\,|\,T)$, $E(\tilde{z}_{\pm}(T,T)^2)$ is to be minimal. Hence $\hat{z}_{\pm}(T\,|\,T)\in Y_T^{\pm}$ is the Hilbert space orthogonal projection of $z_{\pm}(t)$ on $Y_T^{\pm}$ so that $\tilde{z}_{\pm}(T,T)\perp Y_T^{\pm}$ and $g_{\pm}(T,t)$ serves to locate $\hat{z}_{\pm}(T\,|\,T)$ in $Y_T^{\pm}$. This orthogonality condition can be expressed via $E\tilde{z}_{\pm}(T,T)y_{\pm}(s)=0$ for $0\leq s\leq T$ and, changing T to t, one obtains

$$k_{\pm}(t,s)=g_{\pm}(t,s)/2+\int_0^t g_{\pm}(t,\tau)k_{\pm}(\tau,s)d\tau \qquad (0\leq s\leq t). \tag{$\blacklozenge$}$$

From properties of $k_{\pm}$ one has then (see [13]): Assume $k\in C^2$ and set $V_{\pm}(t)=-2D_t g_{\pm}(t,t)$. Then $g_{\pm}$ satisfies $(D_t^2-D_s^2)g_{\pm}(t,s)=V_{\pm}(t)g_{\pm}(t,s)$ with $D_s g_{+}(t,0)=0$ and $g_{-}(t,0)=0$. Next one defines the innovations processes $\nu_{\pm}(t)=y_{\pm}(t)-\hat{z}_{\pm}(t\,|\,t)$ which are Gaussian white noise with $E\nu_{\pm}(t)\nu_{\pm}(s)=\delta(t-s)/2$. Thus $\nu_{\pm}(t)=(I-G_{\pm})y_{\pm}$ relates $\nu_{\pm}$ to $y_{\pm}$ by a causal and causally invertible filter. Now go to the spectral domain and write

$$\hat{r}(\lambda)=(1/2\pi)\left\{1+\int_{-\infty}^{\infty} k(t)e^{-i\lambda t}dt\right\}=(1/2\pi)(1+\hat{k}(\lambda)).$$

Then $y(t)\sim e^{i\lambda t}$ with

$$Ey(t)y(s)=\int_{-\infty}^{\infty} e^{i\lambda t}e^{-i\lambda s}\hat{r}(\lambda)d\lambda=\delta(t-s)+k(t-s)$$

and $y_+\sim \mathrm{Cos}\,\lambda t$ with $y_-\sim i\,\mathrm{Sin}\,\lambda t$. Evidently

$$(\spadesuit)\qquad \nu_+\sim \gamma_+(t,\lambda)=\mathrm{Cos}\,\lambda t-\int_0^t g_+(t,s)\,\mathrm{Cos}\,\lambda s ds$$

and
$$\nu_-\sim \gamma_-(t,\lambda)=i\left\{\mathrm{Sin}\,\lambda t-\int_0^t g_-(t,s)\,\mathrm{Sin}\,\lambda s ds\right\}.$$

From $E\nu_\pm(t)\nu_\pm(s)=\delta(t-s)/2$ one obtains

$$(\dagger)\qquad \int_{-\infty}^{\infty}\gamma_\pm(t,\lambda)\gamma_\pm(s,\lambda)\hat{r}(\lambda)d\lambda=\delta(t-s)/2$$

and then (see [13]): The functions $\gamma_\pm$ satisfy $D_t^2\gamma_\pm(t,\lambda)+(\lambda^2-V_\pm(t))\cdot\gamma_\pm(t,\lambda)=0$ with $\gamma_+(0,\lambda)=1, D_t\gamma_+(0,\lambda)=-2k(0), \gamma_-(0,\lambda)=0$, and $D_t\gamma_-(0,\lambda)=i\lambda$.

3. Transmutation and minimization. First we recall from [1] for $P=D^2$ and $Q=D^2-q$ that the transmutation $B:\varphi_\lambda^P(x)=\mathrm{Cos}\,\lambda x\to\varphi_\lambda^Q(y)$ has kernel $\beta(y,x)=\langle\varphi_\lambda^Q(y),\,\mathrm{Cos}\,\lambda x\rangle_\nu$ $(d\nu=(2/\pi)d\lambda)$ and $\delta(x-\lambda)=\langle\varphi_\lambda^Q(y),\varphi_\lambda^Q(x)\rangle_\omega$. $B^{-1}=\mathcal{B}$ has kernel $\gamma(x,y)=\langle\varphi_\lambda^Q(y),\,\mathrm{Cos}\,\lambda x\rangle_\omega$ and $\tilde{B}$ with kernel $\tilde{\beta}(y,x)=\gamma(x,y)$ is a transmutation $P\to Q$. The generalized G-L equation is $\langle\beta(y,t),\,\mathfrak{A}(t,x)\rangle=\tilde{\beta}(y,x)$ where $(\star)$ $\mathfrak{A}(t,x)=\langle\mathrm{Cos}\,\lambda t,\,\mathrm{Cos}\,\lambda x\rangle_\omega$. Such formulas are derived in [1] generally when φ_λ^Q is a "spherical function" (i.e. $h=0$); however the procedures are unchanged for the present situation and we simply indicate the results. We note also that $\tilde{\beta}(y,x)=\delta(x-y)+\tilde{K}(y,x)$ and $\mathfrak{A}(t,x)=\delta(t-x)+\Omega(t,x)$ while for $x<y$, $\tilde{K}(y,x)=0$ so the G-L equation becomes $(s<t,\ s\sim x,\ t\sim y)$

$$(\bullet)\qquad 0=\Omega(t,s)+K(t,s)+\int_0^t K(t,\tau)\Omega(\tau,s)d\tau.$$

Assume now that there is an underlying stochastic model related to Q as in § 2 so $y_+\sim\mathrm{Cos}\,\lambda t$, $\Omega(t,s)\sim 2k_+(t,s)$, $K(t,s)\sim -g_+(t,s)$, $\varphi_\lambda^Q(t)\sim\gamma_+(t,\lambda)$ $(h=-2k(0))$, and $(\spadesuit)$ represents $\varphi_\lambda^Q(t)=\langle\beta(t,s),\,\mathrm{Cos}\,\lambda s\rangle$. Also $\omega(\lambda)\sim 4\hat{r}(\lambda)$ from $(\dagger)$ so that

$$\mathfrak{A}(t,s)\sim 4\int_0^\infty \mathrm{Cos}\,\lambda t\,\mathrm{Cos}\,\lambda s\hat{r}(\lambda)d\lambda.$$

Note
$$r(t)=\int_{-\infty}^{\infty}\hat{r}(\lambda)e^{i\lambda t}d\lambda=2\int_0^\infty \mathrm{Cos}\,\lambda t\hat{r}(\lambda)d\lambda$$

so
$$4\int_0^\infty \mathrm{Cos}\,\lambda t\,\mathrm{Cos}\,\lambda s\hat{r}(\lambda)d\lambda$$
$$=\{r(t-s)+r(t+s)\}=\delta(t-s)+\delta(t+s)+2k_+(t,s)$$

and $\delta(t+s)$ contributes nothing for $s,t>0$. Consider next the orthogonality $\tilde{z}_+(T,T)\perp Y_T^+$ in the spectral domain ($t=T$ again). One knows $E\nu_+(t)y_+(\tau)=\delta(t-\tau)/2$ so for $0\le\tau\le t$

$$(\blacktriangle)\qquad \delta(t-\tau)=2\int_{-\infty}^{\infty}\gamma_+(t,\lambda)\,\mathrm{Cos}\,\lambda\tau\hat{r}(\lambda)d\lambda=\tilde{\beta}(t,\tau)=\langle\gamma_+(t,\lambda),\,\mathrm{Cos}\,\lambda\tau\rangle_\omega.$$

Thus $\tilde{\beta}(t,\tau)=0$ for $\tau<t$ is a consequence of $E\nu_+(t)y_+(\tau)=Ev_+(t)y_+(\tau)$

$=\delta(t-\tau)/2$. On the other hand in the present context one wants to minimize

$$\Xi=\int_0^T\int_0^\infty \{\Upsilon_+(t,\lambda)-\operatorname{Cos}\lambda t+\langle g_+(t,\tau),\operatorname{Cos}\lambda\tau\rangle\}^2 d\omega dt$$

over some class of causal kernels g_+; this amounts to minimizing

$$(\ddagger)\qquad \Xi=2E\int_0^T \{\nu_+(t)-y_+(t)+\langle g_+(t,\tau),y_+(\tau)\rangle\}^2 dt$$

$(\mathfrak{K}\sim -g_+)$. The procedure of [3]–[6] for Ξ leads to

$$(\blacksquare)\qquad \Xi=\hat{\Xi}+\operatorname{Tr}\{\mathfrak{K}(1+\Omega)\mathfrak{K}^*+\mathfrak{K}\Omega+\Omega\mathfrak{K}^*\}$$

where
$$\hat{\Xi}=\int_0^T\int_0^\infty (\varphi_\lambda^Q(t)-\operatorname{Cos}\lambda t)^2 d\omega dt.$$

When $\mathfrak{K}_0$ is a minimizing kernel a standard variational argument yields the G-L equation (●) as the minimizing criterion (Euler equation) and hence $\mathfrak{K}_0=K$ is uniquely determined. Now in going from (1.1) to (■) one uses the fact that $\tilde{\beta}(t,\tau)=\delta(t-\tau)+\tilde{K}(t,\tau)$ with $\tilde{K}(t,\tau)=0$ for $\tau<t$ (known from general transmutation theory—cf. [1]). We see here however that (▲), which arises when $\varphi_\lambda^Q\sim\Upsilon_+$, also provides the required information to produce (■). Further $\mathfrak{A}(t,\tau)=\delta(t-\tau)+\Omega(t,\tau)$, in the form $\Omega(t,\tau)=2k_+(t,\tau)$, follows from the stochastic theory. Hence

Theorem 3.1. *Given a correspondence $\varphi_\lambda^Q(t)\sim\Upsilon_+(t,\lambda)$, $\omega\sim 4\hat{r}$, $\Omega(t,\tau)\sim 2k_+(t,\tau)$, etc. one can characterize the (unique) minimizing kernel for Ξ via (■) and the corresponding G-L equation (●) (with $K(t,\tau)=-g_+(t,\tau)$), using only stochastic information, and the corresponding stochastic problem is that of minimizing in (‡) over a suitable class of causal kernels g_+. This serves to locate $\hat{z}_+(t\,|\,t)=y_+(t)-\nu_+(t)$ in Y_t^+, which is exactly what is accomplished in linear estimation theory.*

References

[1] R. Carroll: Transmutation, Scattering Theory, and Special Functions. North-Holland, Amsterdam (1982).
[2] ——: Some topics in transmutation. Proc. Inter. Conf. Diff. Eqs., Univ. Alabama (1983) (to appear).
[3] ——: Least squares, transmutation, and the Marčenko equation (to appear).
[4] R. Carroll and S. Dolzycki: Transmutation as a minimizing procedure. J. Math. Physics (to appear).
[5] ——: Quelques procédés minimisants dans la transmutation. CR Acad. Sc. Paris (to appear).
[6] R. Carroll and F. Santosa: ONR/SRO Tech. Rept. 12, Cornell Univ. (1983).
[7] K. Case: Advances Math. Supp. Stud., **3**, 25–43 (1978).
[8] K. Chadan and P. Sabatier: Inverse Problems in Quantum Scattering Theory. Springer, N.Y. (1977).
[9] H. Dym and H. McKean: Gaussian Processes, Function Theory, and the Inverse Spectral Problem. Academic Press, N.Y. (1976).
[10] F. Dyson: Studies Math. Physics, Princeton Univ. Press, pp. 151–167 (1976).
[11] T. Kailath: Lectures on Wiener and Kalman Filtering. Springer, N.Y. (1981).
[12] T. Kailath, A. Vieira, and M. Morf: SIAM Review, **20**, 106–119 (1978).
[13] B. Levy and J. Tsitsiklis: Linear estimation of stationary stochastic processes, vibrating strings, and inverse scattering. MIT preprint (1982).
[14] ——: An inverse scattering procedure for the linear estimation of two dimensional isotropic random processes. ibid. (1983).