$t>0$. Now

$$L_1(t)=\frac{dL}{dt}=-\alpha e^{-t}\ln\frac{e^{-t}}{1-\alpha+\alpha e^{-t}}+\alpha\left(\ln\frac{1}{\alpha}\right)(e^{-t}-1),$$

and

$$\frac{d}{dt}(e^tL_1)=\frac{\alpha(1-\alpha)}{1-\alpha+\alpha e^{-t}}-\alpha\left(\ln\frac{1}{\alpha}\right)e^t$$
$$=e^t\alpha\left[\frac{1-\alpha}{(1-\alpha)e^t+\alpha}-\ln\frac{1}{\alpha}\right]$$
$$\leqslant e^t\alpha\left[1-\alpha-\ln\frac{1}{\alpha}\right]<0.$$

Since $L_1(0+)=0$, $L_1(t)<0$ for $t>0$. Hence for $t>0$, $L(t)<L(0+)=0$. This proves the lemma.

REFERENCES

[1] J. R. Pierce, "Optical channels: Practical limits with photon counting," *IEEE Trans. Commun.*, vol. COM-26, pp. 1819–1821, Dec. 1978.
[2] R. J. McEliece, and L. R. Welch, "Coding for optical channels with photon counting," *IEEE Trans. Inform. Theory*, to be published, 1981.
[3] J. R. Pierce, and E. C. Posner, *Introduction to Communication Science and Systems*. New York: Plenum, 1980.
[4] M. Abramovitz, and I. A. Stegun, Eds., *Handbook of Mathematical Functions* (Applied Mathematics Series 55). Washington, DC: U.S. Nat. Bur. Stand. 1964.
[5] S. Karp, E. L. O'Neill, and R. M. Gagliardi, "Communication theory for the free-space optical channel," *Proc. IEEE*, vol. 58, pp. 1611–1626, Oct. 1970.
[6] L. B. Levitin, "Photon channels with small occupation numbers," *Problems of Information Transmission* (translated from Russian), vol. 2, pp. 48–56, 1966.

Recursive Linear Smoothing of Two-Dimensional Random Fields

RICHARD G. OGIER, STUDENT MEMBER, IEEE, AND EUGENE WONG, FELLOW, IEEE

***Abstract*—In an earlier paper, recursive formulas for the causal filtering of two-dimensional random fields were developed. "Causality" in two dimensions is not a physical constraint but rather an artifact introduced to generate recursion, which in turn is motivated by computational efficiency. The earlier results are extended here in order to derive some recursive formulas for "smoothing" estimators which use all the data rather than just the data in the "past".**

I. INTRODUCTION

LET $T=[a_1,b_1]\times[a_2,b_2]$ be a rectangle in the plane. A random field on T is a collection of real-valued random variables indexed by points in T. Suppose that one observes the random field $\{\xi(t), t\in T\}$ which has the form

$$\xi(t)=x(t)+\eta(t), \tag{1.1}$$

where x is a zero-mean random field to be estimated, and η is a two-dimensional white noise uncorrelated with x and having spectral density N_0. The object of this paper is to find an effective means for computing the linear least-squares estimate of $x(t)$ given $\{\xi(s), s\in T\}$, i.e., the smoothed estimate.

The two-dimensional filtering solution in [1] was developed primarily to serve as a foundation on which a recursive solution to the smoothing problem could be obtained. The recursion structure of the filtering solution is shown in Fig. 1. For two points $s=(s_1,s_2)$ and $t=(t_1,t_2)$ in T, $s\leqslant t$ will denote $s_1\leqslant t_1$ and $s_2\leqslant t_2$. A_t will denote the rectangle $\{s\in T\colon s\leqslant t\}$, and δA_t will denote the upper and right borders of A_t. We shall call A_t the *past* at t, and $T-A_t$ the *future* at t. We denote by $\hat{x}(\tau|t)$ the linear least-squares estimate of $x(\tau)$ given $\{\xi(s), s\in A_t\}$. Thus $\hat{x}(\tau|b)$ is the smoothed estimate of $x(\tau)$. By an *increasing path* in T we shall mean a continuous function $\Gamma\colon[0,1]\to T$ such that $\Gamma(0)=a$, $\Gamma(1)=b$, and $\alpha\leqslant\beta\Rightarrow\Gamma(\alpha)\leqslant\Gamma(\beta)$. In [1] it was shown that if the process x is modeled by a class of partial differential equations often so used in the image processing literature [2], then the *filtered estimates* $\{\hat{x}(\tau|t), \tau\in\delta A_t, t\in\Gamma\}$ could be computed recursively for any increasing path Γ using the state

$$\hat{x}_t=\{\hat{x}(\tau|t), \tau\in\delta A_t\}.$$

Note that $\hat{x}_t$ can be considered a random process on δA_t defined by $\hat{x}_t(\tau)=\hat{x}(\tau|t)$. In this paper we shall show that if $\hat{x}_t$ is computed for every t on a forward pass on Γ, then

Manuscript received May 29, 1979; revised March 17, 1980. This work was supported by the Army Research Office under Grant DAAG29-78-G-0186.
R. G. Ogier was with the Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA. He is now with the Coordinated Science Laboratory and the Department of Electrical Engineering, University of Illinois, Urbana, IL 61801
E. Wong is with the Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720.

0018-9448/81/0100-0077$00.75

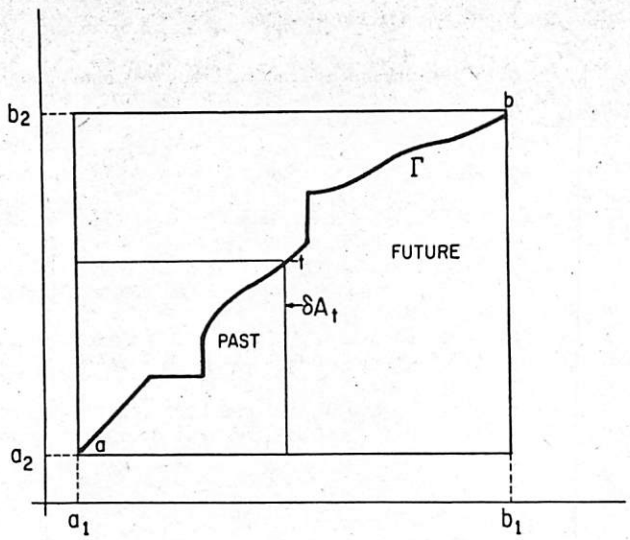

Fig. 1. Recursion structure.

the smoothed estimate $\{\hat{x}(\tau|b),\ t\in T\}$ can be computed recursively on a backward pass on Γ using a new state $\lambda_t=\{\lambda_t(\tau),\ \tau\in\delta A_t\}$ for recursion on the backward pass. The algorithm is a generalization of the corresponding result in the one-dimensional case [3].

In the case where Γ passes through the point (a_1, b_2), the smoothing algorithm becomes one for a function-valued random process with a one-dimensional time parameter. The filtering formula for this case was derived in [4]. An alternate derivation for the smoothing formula in this case could make use of the smoothing formula ([7], Theorem 6.20) for random processes taking values in a separate Hilbert space.

We note that the desire for a recursive solution to the smoothing problem is due, not to any requirement for real-time computation, but to the need to reduce the dimensionality of computation that a "straightforward" solution of the problem would entail. For example, consider a situation where the observation $\xi(t)$ is sampled on an $n\times n$ grid to obtain n^2 samples. The smoothed estimate can be obtained by inverting an n^2 by n^2 covariance matrix that requires roughly n^6 multiplications. In contrast, for a diagonal path, the recursive solution requires roughly $4n^3$ multiplications. Thus there is a potential reduction of dramatic proportion.

II. Innovations Representation

For any rectangle $\Delta=[s_1,t_1)\times[s_2,t_2)$ and any random field $\{V(t),\ t\in T\}$, denote the corresponding increment of V by

$$V(\Delta)=V(t_1,t_2)-V(t_1,s_2)-V(s_1,t_2)+V(s_1,s_2).$$

We say that V is a *standard orthogonal increments* (SOI) process if $E[V(\Delta)]=0$, and

$$E[V(\Delta)V(\Delta')]=\text{area }(\Delta\cap\Delta'),$$

for all rectangles Δ, Δ'. The observation equation (1.1) may be rewritten in differential form as

$$Z(dt)=x(t)\,dt+\sqrt{N_0}\,W(dt) \tag{2.1}$$

where W is an SOI process uncorrelated with x. The precise interpretation of (2.1) is that it defines an integral

$$\int_T g(s)Z(ds)=\int_T g(s)x(s)\,ds+\sqrt{N_0}\int_T g(s)W(ds), \tag{2.2}$$

for all $g\in L^2(T)$, the space of all real-valued, square-integrable functions on T. The second integral on the right side of (2.2) is interpreted in the Wiener sense (see [1]).

Let $\mathcal{H}$ denote the Hilbert space, with inner product $\langle x,y\rangle=Exy$, of all real-valued zero-mean finite-variance random variables on a fixed probability space $(\Omega,\mathcal{F},\mathcal{P})$. We shall assume that for each t, both $Z(t)$ and $x(t)$ belong to $\mathcal{H}$. Let $\mathcal{H}_t^Z$ denote the closed linear subspace spanned by $\{Z(s):\ s\in A_t)\}$. Thus $\hat{x}(\tau|t)$ is the orthogonal projection of $x(\tau)$ onto $\mathcal{H}_t^Z$.

For an SOI process V, the subspace $\mathcal{H}_t^V$ spanned by $\{V(s):\ s\in A_t\}$ is equal to the space of Wiener integrals $\int_{A_t} g(s)V(ds)$. Similarly, for the observation process Z, we have

$$\mathcal{H}_t^Z=\left\{\int_{A_t} g(s)Z(ds):\ g\in L^2(T)\right\}. \tag{2.3}$$

This result is easily proved as a generalization to a similar result in the one-dimensional case.

For a fixed increasing path Γ and any $s\in T$, denote

$$s\Gamma=\max\{t\in\Gamma:\ s\in\delta A_t\}, \tag{2.4}$$

and define

$$\hat{Z}_\Gamma(dt)=Z(dt)-\hat{x}(t|t\Gamma)\,dt. \tag{2.5}$$

Defining the error

$$\mathcal{E}(\tau|t)=x(\tau)-\hat{x}(\tau|t), \tag{2.6}$$

we can use (2.1) to rewrite (2.5) as

$$\hat{Z}_\Gamma(dt)=\mathcal{E}(t|t\Gamma)\,dt+\sqrt{N_0}\,W(dt). \tag{2.7}$$

The process $\hat{Z}_\Gamma$ will be called the *innovations process* for the path Γ.

In one-dimensional estimation, the usefulness of the innovations process is due to the following properties. 1) Any element in the linear space spanned by past observations can be expressed as an integral with respect to the past innovations. 2) The innovations process has orthogonal increments. We shall see that $\hat{Z}_\Gamma$ possesses only restricted versions of these properties which nevertheless allow a useful innovations representation for $\hat{x}(\tau|t)$.

Using (2.5), (2.7), and the fact that $\epsilon(\tau|t)\perp\mathcal{H}_t^Z$, we find that

$$E\hat{Z}_\Gamma(ds)\hat{Z}_\Gamma(ds')=\delta_{s,s}N_0\,ds+\delta_{s\Gamma,s'\Gamma}\rho(s,s'|s\Gamma)\,ds\,ds', \tag{2.8}$$

where $\delta_{s,s'}=1$, if $s=s'$ and zero otherwise, and ρ is the

(3.4). For $\tau\in\delta A_t$,

$$d_{t_1}\epsilon(\tau|t)=\left[\delta_{\tau_1 t_1}\alpha_1(\tau)\epsilon(\tau|t)+\int_{a_2}^{t_2}g_1(\tau;t_1,u_2|t)\epsilon(t_1,u_2|t)\,du_2\right]dt_1+\int_{a_2}^{\tau_2}h_1(t_1,\tau_2;u_2)V(dt_1,du_2)-\frac{1}{\sqrt{N_0}}\int_{a_2}^{t_2}\rho(\tau;t_1,u_2|t)W(dt_1,du_2),\tag{3.6}$$

and $d_{t_2}\epsilon(\tau|t)$ is given by the dual equation. The error equations will play a large role in the derivation of the smoothing formula.

The filtering error covariances $\{\rho(\tau;u|t),\ \tau,u\in\delta A_t,\ t\in\Gamma\}$ required by the filtering equations are given recursively by the following Riccati equations. For $\tau,u\in\delta A_t$ with $\tau_1=t_1$,

$$\frac{\partial}{\partial t_1}\rho(\tau;u|t)=\alpha_1(\tau)\rho(\tau;u|t)+\int_{a_2}^{\tau_2}f_1(\tau;u_2')\rho(t_1,u_2';u|t)\,du_2'+\delta_{u_1t_1}\left[\alpha_1(u)\rho(\tau;u|t)+\int_{a_2}^{u_2}f_1(u;u_2')\rho(\tau;t_1,u_2'|t)\,du_2'+\int_{a_2}^{\min(\tau_2,u_2)}h_1(\tau;u_2')h_1(u;u_2')\,du_2'\right]-\frac{1}{N_0}\int_{a_2}^{t_2}(\rho;t_1,u_2'|t)\rho(t_1,u_2';u|t)\,du_2'.\tag{3.7}$$

For $\tau,u\in\delta A_t$ with $\tau_2=t_2$, $(\partial\rho(\tau;u|t)/\partial t_2)$ is given by the dual of (3.7). Once the filtering error covariances are known, (3.4) can be used to recursively compute the filtered estimates $\{\hat{x}(\tau|t),\ \tau\in\delta A_t,\ t\in\Gamma\}$.

IV. Smoothing Formula

In this section it is assumed that we have applied the filtering formula along some fixed increasing path Γ and have obtained the filtered estimates $\{\hat{x}(\tau|t),\ \tau\in\delta A_t,\ t\in\Gamma\}$. The following smoothing formula uses these estimates and the observations $\{Z(t),\ t\in T\}$ to recursively compute the smoothed estimates $\{\hat{x}(t|b),\ t\in T\}$.

Theorem: For $t\in\Gamma$, $\tau\in\delta A_t$, the smoothed estimates are given by

$$\hat{x}(\tau|b)=\hat{x}(\tau|t)+\frac{1}{N_0}\left[\rho(\tau;t|t)\lambda_t(t)+\int_{a_2}^{t_2}\rho(\tau;t_1,u_2)\lambda_t(t_1,u_2)\,du_2+\int_{a_1}^{t_1}\rho(\tau;u_1,t_2|t)\lambda_t(u_1,t_2)\,du_1\right],\tag{4.1}$$

where the processes $\lambda_t:\ \delta A_t\to\mathbb{R}$ are given recursively by $\lambda_b=0$ and for $t\in\Gamma$, $u\in\delta A_t$,

$$\frac{\partial}{\partial t_1}\lambda_t(u)=-\delta_{u_1,t_1}\alpha_1(u)\lambda_t(u)-\delta_{u,t}\lambda_t(^-t)-I(u_2<t_2)\left[g_1(t;u|t)\lambda_t(t)+\int_{a_2}^{t_2}g_1(t_1,u_2';u|t)\lambda_t(t_1,u_2')\,du_2'+\int_{a_1}^{t_1}g_1(u_1',t_2;u|t)\lambda_t(u_1',t_2)\,du_1'+\nu_\Gamma(u)\right],\tag{4.2a}$$

$$\frac{\partial}{\partial t_2}\lambda_t(u)=-\delta_{u_1,t_1}\alpha_1(u)\lambda_t(u)-\delta_{u,t}\lambda_t(t_-)-I(u_1<t_1)\left[g_2(t;u|t)\lambda_t(t)+\int_{a_1}^{t_1}g_2(u_1',t_2;u|t)\lambda_t(u_1',t_2)\,du_1'+\int_{a_2}^{t_2}g_2(t_1,u_2';u|t)\lambda_t(t_1,u_2')\,du_2'+\nu_\Gamma(u)\right],\tag{4.2b}$$

where

$$\lambda_t(^-t)=\lim_{u_1\uparrow t_1}\lambda_t(u_1,t_2),$$
$$\lambda_t(t_-)=\lim_{u_2\uparrow t_2}\lambda_t(t_1,u_2),$$

and ν_Γ is the white noise derivative of $\hat{Z}_\Gamma$ defined by

$$\nu_\Gamma(u)\,du=\hat{Z}_\Gamma(du).\tag{4.3}$$

Observe that the state $\lambda_t=\{\lambda_t(\tau),\ \tau\in\delta A_t\}$ is computed recursively by (4.3) as t moves *backward* along Γ. If the direction in which t moves is parallel to one leg of δA_t, then $\lambda_t(u)$ is unchanged for u along that leg. In order to compute $\{\hat{x}(\tau|b),\ \tau\in T\}$ from (4.1) we do not require all the computed filtered estimates: only $\{\hat{x}(\tau|\tau\Gamma),\ \tau\in T\}$. Since $\nu_\Gamma(u)\,du=Z(du)-\hat{x}(u|u\Gamma)\,du$, we see that the filtering formula requires only the data $\{\hat{x}(\tau|\tau\Gamma),\ Z(\tau),\ \tau\in T\}$.

We begin the derivation by using the innovations representation of Section II to express each smoothed estimate $\hat{x}(\tau|b)$ as a filtered estimate plus an innovations integral.

Lemma 1: For $t\in\Gamma$, $\tau\in\delta A_t$,

$$\hat{x}(\tau|b)=\hat{x}(\tau|t)+\frac{1}{N_0}\int_{A_b-A_t}E[\epsilon(\tau|t)\epsilon(s|s\Gamma)]\hat{Z}_\Gamma(ds).\tag{4.4}$$

Proof: Using (2.13) and (2.14), the smoothed estimate can be expressed as

$$\hat{x}(\tau|b)=\int_{A_b}k(\tau;s)\hat{Z}_\Gamma(ds)=\hat{x}(\tau|t)+\int_{A_b-A_t}k(\tau;s)\hat{Z}_\Gamma(ds),\tag{4.5}$$

for *any* $\tau\in T$, $t\in\Gamma$, where

$$k(\tau;s)\,ds=\frac{1}{N_0}Ex(\tau)\hat{Z}_\Gamma(ds).$$

If $t\in\Gamma$, $\tau\in\delta A_t$, and $s\in A_b-A_t$, we have

$$\begin{aligned}k(\tau;s)\,ds&=\frac{1}{N_0}E\Big[x(\tau)\big\{\epsilon(s|s\Gamma)\,ds+\sqrt{N_0}\,W(ds)\big\}\Big]\\&=\frac{1}{N_0}E[x(\tau)\epsilon(s|s\Gamma)]\,ds\\&=\frac{1}{N_0}E[\epsilon(\tau|t)\epsilon(s|s\Gamma)]\,ds,\end{aligned}\tag{4.6}$$

where the orthogonality of $\epsilon(s|s\Gamma)$ and $\hat{x}(\tau|t)$ was used in the third equality. Substituting (4.6) in (4.5) yields (4.4). Q.E.D.

We proceed to derive a recursive formula for the integral in (4.4). The integrand is the covariance between two filtering errors: one on δA_t and one in A_b-A_t. With the immediate goal of expressing this integrand in terms of the filtering error covariances $\rho(\tau;u|t)$, $\tau,u\in\delta A_t$, we introduce the evolution operator $\mathcal{E}$ associated with the error equation (3.6) and its dual. For $t\in T$, let C_t denote the space of continuous functions mapping δA_t to $\mathbb{R}$. Let Λ be an increasing path, $t\in\Lambda$, and $y\in C_t$. Then for $r\in\Lambda$, $\mathcal{E}_{r;t}$: $C_t\to C_r$ is defined recursively by

$$\mathcal{E}_{t;t}[y]=y,\tag{4.7}$$

and for $s\in\delta A_r$,

$$\frac{\partial}{\partial r_1}\mathcal{E}_{r;t}[y](s)=\delta_{s_1,r_1}\alpha_1(s)\mathcal{E}_{r;t}[y](s)+\int_{a_2}^{r_2}g_1(s;r_1,s_2'|r)\mathcal{E}_{r;t}[y](r_1,s_2')\,ds_2',\tag{4.8}$$

where

$$\frac{\partial}{\partial r_2}\mathcal{E}_{r;t}[y](s)$$

is given by the dual of (4.8). Hence $\mathcal{E}_{r;t}[y]$ gives the state of the zero-noise error equations at r given that the state at t is y.

Lemma 2: Let $t,r\in T$ with $t\leqslant r$. Then $\mathcal{E}_{r;t}$ has the form

$$\begin{aligned}\mathcal{E}_{r;t}[y](s)=\;&\phi_{r;t}(s)y(s\wedge t)+\psi_{r;t}(s)y(t)\\&+\int_{a_2}^{t_2}\Phi^1_{r;t}(s;u_2)y(t_1,u_2)\,du_2\\&+\int_{a_1}^{t_1}\Phi^2_{r;t}(s;u_1)y(u_1t_2)\,du_1,\end{aligned}\tag{4.9}$$

where $s\wedge t=(\min\{s_1,t_1\},\min\{s_2,t_2\})$, and $\phi_{r;t},\psi_{r;t},\Phi^1_{r;t},\Phi^2_{r;t}$ have the following properties:

$$\begin{aligned}&\text{a)}\;\; s\in\delta A_t\Rightarrow\phi_{r;t}(s)=1,\\&\text{b)}\;\; r_1=t_1\text{ or }r_2=t_2\Rightarrow\psi_{r;t}(s)=0,\\&\text{c)}\;\; r_1=t_1\Rightarrow\Phi^1_{r;t}(s;u_2)=0,\\&\text{d)}\;\; r_2=t_2\Rightarrow\Phi^2_{r;t}(s;u_1)=0.\end{aligned}\tag{4.10}$$

The backward evolution operator $\mathcal{E}_{t;r}$ has the form

$$\begin{aligned}\mathcal{E}_{t;r}[y](u)=\;&\phi_{t;r}(u)y(u^*r)\\&+\int_{a_2}^{r_2}\Phi^1_{t;r}(u;s_2)y(r_1,s_2)\,ds_2\\&+\int_{a_1}^{r_1}\Phi^2_{t;r}(u;s_1)y(s_1,r_2)\,ds_1,\end{aligned}\tag{4.11}$$

where

$$u^*r=\begin{cases}r, u=t,\\(r_1,u_2), & u_2<t_2,\\(u_1,r_2), & u_1<t_1.\end{cases}$$

Proof: We first find forms for $\mathcal{E}_{r;t}$ when $r_2=t_2$ and when $r_1=t_1$, i.e., for the "horizontal" operator $\mathcal{E}^{1,t_2}_{r_1,t_1}=\mathcal{E}_{r_1,t_2;t}$ and for the "vertical" operator $\mathcal{E}^{2,t_1}_{r_2,t_2}=\mathcal{E}_{t_1,r_2;t}$. From (4.7) and (4.8), we have for $r_1'\in[t_1,r_1]$ and $s\in\delta A_{r_1',t_2}$,

$$\frac{\partial}{\partial r_1'}\mathcal{E}^{1,t_2}_{r_1',t_1}[y](s)=\delta_{s_1,r_1'}\alpha_1(s)\mathcal{E}^{1,t_2}_{r_1',t_1}[y](s)+\int_{a_2}^{t_2}g_1(s;r_1',s_2'|r_1',t_2)\mathcal{E}^{1,t_2}_{r_1',t_1}[y](r_1',s_2')\,ds_2',\tag{4.12}$$

where $\mathcal{E}^{1,t_2}_{t_1,t_1}[y]=y$. Using the method of successive approximations [6], it can be shown that (4.12) has a solution of the form

$$\mathcal{E}^{1,t_2}_{r_1,t_1}[y](s)=\phi^{1,t_2}_{r_1,t_1}(s)y(s_1\wedge t_1,s_2)+\int_{a_2}^{t_2}\Phi^{1,t_2}_{r_1,t_1}(s;u_2)y(t_1,u_2)\,du_2,\tag{4.13}$$

where

$$\begin{aligned}&\text{a)}\;\; s\in\delta A_t\Rightarrow\phi^{1,t_2}_{r_1,t_1}(s)=1,\\&\text{b)}\;\; r_1=t_1\Rightarrow\Phi^{1,t_2}_{r_1,t_1}(s;u_2)=0.\end{aligned}\tag{4.14}$$

A dual argument yields the form for $\mathcal{E}^{2,t_1}_{r_2,t_2}$, which is just the dual of (4.13) and (4.14). We can now obtain the form for $\mathcal{E}_{r;t}$ by composing the horizontal and vertical forms, i.e.,

$$\mathcal{E}_{r;t}[y]=\mathcal{E}^{2,r_1}_{r_2,t_2}\Big[\mathcal{E}^{1,t_2}_{r_1,t_1}[y]\Big].\tag{4.15}$$

Substituting (4.13) and its dual in (4.15) gives the form (4.9). Properties (4.10) follow from knowing that (4.10) must reduce to (4.13) or its dual under each given condition. The backward form is obtained similarly. Q.E.D.

We now use $\mathcal{E}$ to rewrite the integral in (4.4) in terms of the filtering error covariances $\rho(\tau,u|t)$, $\tau,u\in\delta A_t$.

Lemma 3: For $t\in\Gamma$, $\tau\in\delta A_t$

$$\begin{aligned}\hat{x}(\tau|b)=\hat{x}(\tau|t)+\frac{1}{N_0}\Big[&\rho(\tau;t|t)\lambda_c(t)\\&+\int_{a_2}^{t_2}\rho(\tau;t_1,u_2)\lambda_1(t;du_2)\\&+\int_{a_1}^{t_1}\rho(\tau;u_1,t_2)\lambda_2(t;du_1)\Big],\end{aligned}\tag{4.16}$$

where, for $u \in \delta A_t$,

$$\lambda_c(t)=\int_{t_1}^{b_1}\int_{t_2}^{b_2}\phi_{s\Gamma;t}(s)\hat{Z}_\Gamma(ds_1,ds_2)+\int_{A_b-A_t}\psi_{s\Gamma;t}(s)\hat{Z}_\Gamma(ds), \quad (4.17a)$$

$$\lambda_1(t;du_2)=\int_{t_1}^{b_1}\phi_{s\Gamma;t}(s_1,u_2)\hat{Z}_\Gamma(ds_1,du_2)+\int_{A_b-A_t}\Phi^1_{s\Gamma;t}(s;u_2)\hat{Z}_\Gamma(ds)\,du_2, \quad (4.17b)$$

$$\lambda_2(t;du_1)=\int_{t_2}^{b_2}\phi_{s\Gamma;t}(u_1,s_2)\hat{Z}_\Gamma(du_1,ds_2)+\int_{A_b-A_t}\Phi^2_{s\Gamma;t}(s;u_1)\hat{Z}_\Gamma(ds)\,du_1. \quad (4.17c)$$

Proof: Let $r > t$. By linearity, the state $\epsilon(\cdot|r)$ of the error equations at r, given that the state at t is $\epsilon(\cdot|t)$, is

$$\epsilon(\cdot|r)=\mathcal{E}_{r;t}[\epsilon(\cdot|t)]+\mathcal{N}_{r;t}[V(dv),W(dv),v\in A_r-A_t],$$

where $\mathcal{N}_{r;t}$ is a linear operator. Therefore, for $\tau\in\delta A_t$ and $s\in A_b-A_t$,

$$\begin{aligned}E\epsilon(\tau|t)\epsilon(s|s\Gamma)&=E\epsilon(\tau|t)\{\mathcal{E}_{s\Gamma;t}[\epsilon(\cdot|t)](s)\\&\quad+\mathcal{N}_{s\Gamma;t}[V(dv),W(dv),v\in A_{s\Gamma}-A_t](s)\}\\&=\mathcal{E}_{s\Gamma;t}[E\epsilon(\tau|t)\epsilon(\cdot|t)](s)\\&=\mathcal{E}_{s\Gamma;t}[\rho(\tau;\cdot|t)](s),\end{aligned} \quad (4.18)$$

where the orthogonality of $\epsilon(\tau|t)$ and $V(dv),W(dv)$, $v\in A_b-A_b$ was used in the second equality. Hence

$$\begin{aligned}E\epsilon(\tau|t)\epsilon(s|s\Gamma)&=\phi_{s\Gamma;t}(s)\rho(\tau;s\wedge t|t)\\&\quad+\psi_{s\Gamma;t}(s)\rho(\tau;t|t)\\&\quad+\int_{a_2}^{t_2}\Phi^1_{s\Gamma;t}(s;u_2)\rho(\tau;t_1,u_2|t)\,du_2\\&\quad+\int_{a_1}^{t_1}\Phi^2_{s\Gamma;t}(s;u_1)\rho(\tau;u_1,t_2|t)\,du_1.\end{aligned} \quad (4.19)$$

Substituting (4.19) in (4.4) and rearranging, we obtain

$$\begin{aligned}\hat{x}(\tau|b)=\hat{x}(\tau|t)+\frac{1}{N_0}\Bigg[&\int_{a_2}^{t_2}\rho(\tau;t_1,s_2|t)\int_{t_1}^{b_1}\phi_{s\Gamma;t}(s)\hat{Z}_\Gamma(ds_1,ds_2)+\int_{a_1}^{t_1}\rho(\tau;s_1,t_2|t)\int_{t_2}^{b_2}\phi_{s\Gamma;t}(s)\hat{Z}_\Gamma(ds_1,ds_2)\\&+\rho(\tau;t|t)\int_{t_1}^{b_1}\int_{t_2}^{b_2}\phi_{s\Gamma;t}(s)\hat{Z}_\Gamma(ds_1,ds_2)+\rho(\tau;t|t)\int_{A_b-A_t}\psi_{s\Gamma;t}(s)\hat{Z}_\Gamma(ds)\\&+\int_{a_2}^{t_2}\rho(\tau;t_1,u_2|t)\int_{A_b-A_t}\Phi^1_{s\Gamma;t}(s;u_2)\hat{Z}_\Gamma(ds)\,du_2+\int_{a_1}^{t_1}\rho(\tau;u_1,t_2|t)\int_{A_b-A_t}\Phi^2_{s\Gamma;t}(s;u_1)\hat{Z}_\Gamma(ds)\,du_1\Bigg].\end{aligned} \quad (4.20)$$

Defining $\lambda_c,\lambda_1,\lambda_2$ as in (4.17), we can rewrite (4.20) as (4.16). Q.E.D.

With the goal of finding recursive formulas for $\lambda_c,\lambda_1,\lambda_2$, we first find the partial derivatives with respect to t_1 and t_2 of $\phi_{r;t}$, $\psi_{r;t}$, $\Phi^1_{r;t}$, and $\Phi^2_{r;t}$.

Lemma 4: Let $t,r\in T$ with $t\leq r$, and let $s\in\delta A_r$. Then

$$\frac{\partial}{\partial t_1}\phi_{r;t}(s)=-I(s_1\geq t_1)\alpha_1(s\wedge t)\phi_{r;t}(s), \quad (4.21a)$$

$$\frac{\partial}{\partial t_1}\psi_{r;t}(s)=-\alpha_1(t)\psi_{r;t}(s)-\Phi^2_{r;t}(s;t_1), \quad (4.21b)$$

$$\begin{aligned}\frac{\partial}{\partial t_1}\Phi^1_{r,t}(s;u_2)=-\Bigg[&g_1(s\wedge t;t_1,u_2|t)\phi_{r;t}(s)\\&+g_1(t;t_1,u_2|t)\psi_{r;t}(s)\\&+\alpha_1(t_1,u_2)\Phi^1_{r;t}(s;u_2)\\&+\int_{a_2}^{t_2}g_1(t_1,u_2';t_1,u_2|t)\Phi^1_{r;t}(s;u_2')\,du_2'\\&+\int_{a_1}^{t_1}g_1(u_1',t_2;t_1,u_2|t)\Phi^2_{r;t}(s;u_1')\,du_1'\Bigg],\end{aligned} \quad (4.21c)$$

$$\frac{\partial}{\partial t_1}\Phi^2_{r;t}(s;u_1)=0.$$

The dual equations give the partial derivatives with respect to t_2.

Proof: Let $\epsilon>0$ and let $\{y_{\tau_1,t_2}\in C_{\tau_1,t_2},\ \tau_1\in(t_1-\epsilon,t_1+\epsilon)\cap[a_1,b_1]\}$ be a family of functions with $d_{t_1}y_{\tau_1,t_2}=0$. It is clear that for $r>t$ and $u\in\delta A_t$,

$$y_t(u)=\mathcal{E}_{t;r}[\mathcal{E}_{r;t}[y_t]](u). \quad (4.22)$$

Using (4.9) and (4.11) to write out (4.22), taking $\partial/\partial t_1$, using (4.8), and using (4.22) again, we obtain

$$0=\left[\delta_{u_1t_1}\alpha_1(u)y_t(u)+\int_{a_2}^{t_2}g_1(u;t_1,u_2'|t)y_t(t_1,u_2')\,du_2'\right]+\mathcal{E}_{t;r}\left[\frac{\partial}{\partial t_1}\mathcal{E}_{r;t}[y_t]\right](u). \quad (4.23)$$

Subtracting the last term from both sides, applying $\mathcal{E}_{r;t}$ to each side as a function of u, using (4.22) for the left side, and writing out both sides using (4.9), we obtain

$$-\Big[\frac{\partial}{\partial t_1}\phi_{r;t}(s)y_t(s\wedge t)+\frac{\partial}{\partial t_1}\psi_{r;t}(s)y_t(t)$$
$$+\int_{a_2}^{t_2}\frac{\partial}{\partial t_1}\Phi^1_{r;t}(s;u_2)y_t(t_1,u_2)\,du_2$$
$$+\int_{a_1}^{t_1}\frac{\partial}{\partial t_1}\Phi^2_{r;t}(s;u_1)y_t(u_1,t_2)\,du_1$$
$$+\Phi^2_{r;t}(s;t_1)y_t(t)\Big]$$
$$=\phi_{r;t}(s)\Big[I(s_1>t_1)\alpha_1(s\wedge t)y_t(s\wedge t)$$
$$+\int_{a_2}^{t_2}g_1(s\wedge t;t_1,u_2'|t)y_t(t_1,u_2')\,du_2'\Big]$$
$$+\psi_{r;t}(s)\Big[\alpha_1(t)y_t(t)$$
$$+\int_{a_2}^{t_2}g_1(t;t_1,u_2'|t)y_t(t_1,u_2')\,du_2'\Big]$$
$$+\int_{a_2}^{t_2}\Phi^1_{r;t}(s;u_2)\Big[\alpha_1(t_1,u_2)y_t(t_1,u_2)$$
$$+\int_{a_2}^{t_2}g_1(t_1,u_2;t_1,u_2'|t)y_t(t_1,u_2')\,du_2'\Big]du_2$$
$$+\int_{a_2}^{t_2}\Phi^2_{r;t}(s;u_1)\Big[\int_{a_2}^{t_2}g_1(u_1,t_2;t_1,u_2'|t)$$
$$\cdot y_t(t_1,u_2')\,du_2'\Big]du_1. \tag{4.24}$$

Since y_t is arbitrary, it is easy to show that (4.24) implies (4.21). Q.E.D.

Lemma 5: The processes $\lambda_c,\lambda_1,\lambda_2$ are given recursively by $\lambda_c(b)=\lambda_1(b;du_2)=\lambda_2(b;du_1)=0$, and for $t\in\Gamma$, $u\in\delta A_t$, $u\neq t$,

$$d_{t_1}\lambda_c(t)=-\alpha_1(t)\lambda_c(t)\,dt_1-\lambda_2(t;dt_1), \tag{4.25a}$$

$$d_{t_1}\lambda_1(t;du_2)=-\Big[\alpha_1(t_1,u_2)\lambda_1(t;du_2)$$
$$+g_1(t;t_1,u_2|t)\lambda_c(t)\,du_2$$
$$+\int_{a_2}^{t_2}g_1(t_1,u_2';t_1,u_2|t)\lambda_1(t;du_2')\,du_2$$
$$+\int_{a_1}^{t_1}g_1(u_1',t_2;t_1,u_2|t)\lambda_2(t;du_1')\,du_2\Big]dt_1$$
$$-\hat{Z}_\Gamma(dt_1,du_2), \tag{4.25b}$$
$$d_{t_1}\lambda_2(t;du_1)=0. \tag{4.25c}$$

The dual equations give $d_{t_2}\lambda_c(t)$, $d_{t_2}\lambda_1(t;du_2)$, and $d_{t_2}\lambda_2(t;du_1)$.

Proof: The initial conditions are obvious from (4.17). Equations (4.25) are obtained by differentiating equations (4.17) using (4.21) and properties (4.10).

Proof of Theorem: Define the process $\{\lambda_t(u),u\in\delta A_t\}$ by

$$\lambda_t(t)=\lambda_c(t),$$
$$\lambda_t(t_1,u_2)\,du_2=\lambda_1(t;du_2),$$
$$\lambda_t(u_1,t_2)\,du_1=\lambda_2(t;du_1). \tag{4.26}$$

Then (4.1) is simply (4.16) rewritten in terms of λ_t. Equations (4.2) are simply (4.25) and their duals rewritten in terms of λ_t and ν_Γ as defined in (4.3). Q.E.D.

References

[1] E. Wong, "Recursive causal linear filtering for two-dimensional random fields," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 50–59, Jan. 1978.

[2] S. Attasi, "Stochastic state space representation of images," in *Lecture Notes in Economics and Mathematical Systems*, vol. 107. Berlin: Springer, 1975, pp. 218–230.

[3] T. Kailath, "An innovations approach to least-squares estimation, part II: Linear smoothing in additive white noise," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 655–660, Dec. 1968.

[4] E. Wong and E. T. Tsui, "One-sided recursive filters for two-dimensional random fields," *IEEE Trans. Inform. Theory*, vol. IT-2, pp. 84–86, Jan. 1975.

[5] M. H. A. Davis, *Linear Estimation and Stochastic Control.* London: Chapman and Hall, 1977.

[6] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations.* New York: McGraw-Hill, 1971.

[7] R. F. Curtain and A. J. Pritchard, *Infinite Dimensional Linear Systems.* Berlin: Springer, 1978.