Representation of stochastic processes and rational approximation

A. Gombani
LADSEB
Padova, Italy

Summer School on
Harmonic Analysis and Rational Approximation ,
September 2003

OUTLINE

- Stationary gaussian procesess $\{y(t)\}_{\mathbb{R}}$ as elements of a separable Hilbert space $\mathbf{H}_y$ endowed with a shift U (See talk of Deistler)
- Map (non unique!) into square integrable functions L^2.
- Relate different models in terms of inner functions
- Use Hardy spece tools for approximation (See talk of Baratchart. Other example: Hankel norm) in strong sense
- Try to explain why all this is interesting

OUTLINE

- Stationary gaussian processess $\{y(t)\}_{\mathbb{R}}$ as elements of a separable Hilbert space $\mathbf{H}_y$ endowed with a shift U (See talk of Deistler)
- Map (non unique!) into square integrable functions L^2.
- Relate different models in terms of inner functions
- Use Hardy spece tools for approximation (See talk of Baratchart. Other example: Hankel norm) in strong sense
- Try to explain why all this is interesting

OUTLINE

- Stationary gaussian procesess $\{y(t)\}_{\mathbb{R}}$ as elements of a separable Hilbert space $\mathbf{H}_y$ endowed with a shift U (See talk of Deistler)
- Map (non unique!) into square integrable functions L^2.
- Relate different models in terms of inner functions
- Use Hardy spece tools for approximation (See talk of Baratchart. Other example: Hankel norm) in strong sense
- Try to explain why all this is interesting

OUTLINE

- **Stationary gaussian procesess $\{y(t)\}_{\mathbb{R}}$ as elements of a separable Hilbert space $\mathbf{H}_y$ endowed with a shift U (See talk of Deistler)**
- **Map (non unique!) into square integrable functions L^2.**
- **Relate different models in terms of inner functions**
- **Use Hardy spece tools for approximation (See talk of Baratchart. Other example: Hankel norm) in strong sense**
- **Try to explain why all this is interesting**

OUTLINE

- Stationary gaussian procesess $\{y(t)\}_{\mathbb{R}}$ as elements of a separable Hilbert space $\mathbf{H}_y$ endowed with a shift U (See talk of Deistler)
- Map (non unique!) into square integrable functions L^2.
- Relate different models in terms of inner functions
- Use Hardy spece tools for approximation (See talk of Baratchart. Other example: Hankel norm) in strong sense
- Try to explain why all this is interesting

Aim of the talk

- To show that the study of stationary gaussian processes can be carried out in the function space H_2 and thus to the use of H_2 tools.
- Show that this leads to the study of different representations aand thus a quite rich structure in H_2.
- Give an example (Hankel norm approximation) where this choice of representation matters.

Aim of the talk

- To show that the study of stationary gaussian processes can be carried out in the function space H_2 and thus to the use of H_2 tools.
- Show that this leads to the study of different representations aand thus a quite rich structure in H_2.
- Give an example (Hankel norm approximation) where this choice of representation matters.

Aim of the talk

- **To show that the study of stationary gaussian processes can be carried out in the function space H_2 and thus to the use of H_2 tools.**
- **Show that this leads to the study of different representations aand thus a quite rich structure in H_2.**
- **Give an example (Hankel norm approximation) where this choice of representation matters.**

GAUSSIAN STATIONARY PROCESS

A bit of history:

Kolmogorov (1939) Wiener (1947) gave a Hilbert space (infinite dimensional) representation (main reference is Rozanov).

Kalman (1960) gave a finite dimensional representation.

Anderson (1971) Ruckebusch (1974), Lindquist -Picci (1975-85) and others gave a complete description of state space representations.

m-dimensional continuous time gaussian process: sequence of gaussian m-dimensional random vectors

$$y(t, \omega)$$

on $(\Omega, \mathcal{F}, \mathcal{P})$ indexed by $t \in \mathbb{Z}$. It's stationary if its mean

$$\mathrm{E}y(t) = \int_{\Omega} y(t, \omega) d\omega$$

is independent of t and its correlation function

$$C(t, s) = \int_{\Omega} y(t, \omega) y(s, \omega)^* d\omega$$

only depends on the difference $t - s$ We assume w.l.o.g. that $\mathrm{E}y(t) = 0$.

Can define inner product

$$\langle y(t), y(s) \rangle := \mathrm{E}y(s)^* y(t) \tag{1}$$

(notice that this is not the correlation function, although similar!)

Define linear span

$$\mathbf{H}_y^0 := \{\xi; \xi = \sum_{k=1}^{n} c_k y(t_k), n < +\infty, t_1, ..., t_n \in \mathbb{Z}\}$$

Can extend (1) to $\mathbf{H}_y^0$ and take closure

$$\mathbf{H}_y := \overline{\mathbf{H}}_y^0$$

It's Hilbert!

Define $\mathbf{U}$ as

$$\mathbf{U}y(t) := y(t+1)$$

and extend it easily to $\mathbf{H}_y^0$. Stationarity implies that $\mathbf{U}$ is unitary. In fact,

$$\|\mathbf{U}y(t)\|^2 = \mathbf{E}y(t+1)^* y(t+1) = \mathbf{E}y(t)^* y(t) = \|y(t)\|^2$$

so, $\mathbf{U}$ is in fact unitary on $\mathbf{H}_y^0$ and thus on $\mathbf{H}_y$.

Spectral representation of stationary gaussian processes

- **It's a very nice story...**
- **...but a bit too long!**
- **See written notes (and Deistler's talk)!**

Spectral representation of stationary gaussian processes

- It's a very nice story...
- ...but a bit too long!
- See written notes (and Deistler's talk)!

Spectral representation of stationary gaussian processes

- It's a very nice story...
- ...but a bit too long!
- See written notes (and Deistler's talk)!

Easy way...

(Use Wold decomposition for p.n.d process)

Define past of y as

$$\mathbf{H}_y^-(t) := \overline{span}\{y(s); s \leq t\}$$

$$u_0(t) := y(t) - \mathbf{E}_{\mathbf{H}_y^-(t-1)} y(t)$$

$$u_-(t) := \frac{u_0(t)}{\|u_0(t)\|}$$

u_- **is a white noise since** $u_-(t)u_-(s)^* = \delta_{t,s}$.

Have a basis in H_y

In fact, $u_-(n)$ is an orthonormal family (p.n.d does the rest). Then, setting

$$w_n := \langle y(t), u_-(t-n)\rangle \qquad n \in \mathbb{N}$$

we obviously have

$$\{w_n\}_{n\in\mathbb{N}} \in l^2(0,\infty)$$

and

$$y(t) = \sum_{n=0}^{\infty} w_n u_-(t-n)$$

Hardy space

Define

$$W_-(z) := \sum_{n=0}^{\infty} w_n z^{-n}$$

It's clearly analytic in the complement $\mathbb{E}$ of $\mathbb{D}$.

If we define (see Baratchart's talk),

$$H_2 := \overline{span}\{z^{-n}; n \geq 0\} \qquad (2)$$

we have

$$W_- \in H_2$$

In fact, setting

$$I_{u_-} : \mathrm{H}_y \mapsto L^2$$
$$I_{u_-} u_-(n) := z^n$$

can write

$$I_{u_-} y(0) = \sum_{n=0}^{\infty} w_n I_{u_n} u_-(-n) = \sum_{n=0}^{\infty} w_n z^{-n} = W_-(z)$$

Are we happy?

Not really:

- **The Fourier Transform is quite artificial**
- **Is this representation unique?**
- **If not, what is the structure?**

Are we happy?

Not really:

- **The Fourier Transform is quite artificial**
- **Is this representation unique?**
- **If not, what is the structure?**

Are we happy?

Not really:

- **The Fourier Transform is quite artificial**
- **Is this representation unique?**
- **If not, what is the structure?**

Fourier Transform

It's intrinsic in the shift U

$$U = \int_{-\pi}^{\pi} e^{i\lambda} E(d\lambda)$$

and

$$U^n = \int_{-\pi}^{\pi} e^{i\lambda n} E(d\lambda)$$

See notes... (and Manfred's talk)

(Can use same techinque in continuous time, using the Spectral Theorem for self-adjoint operators)

Uniqueness

There is none (there are many basis in a Hilbert space).

Is the representation unique if we impose the basis to be a white noise? The answer is again no, as we can see in the following

Example

Let y be scalar and set

$$u'(t) := u_-(t+1)$$

Then

$$\begin{aligned} y(t) &= \sum_{n=0}^{\infty} w_n u_-(t-n) \\ &= \sum_{n=0}^{\infty} w_n u'(t-n-1) \\ &= \sum_{n=1}^{\infty} w_{n-1} u'(t-n) \end{aligned}$$

Setting

$$I_{u'}u'(n) = z^n$$

we can see now that

$$I_{u'}y(0) = W'(z) = \sum_{n=1}^{\infty} w_{n-1} z^{-n} = \sum_{n=0}^{\infty} w_n z^{-n-1} = z^{-1} W_-(z) \tag{3}$$

i.e. W_- and W' are related by an inner function, that is a function $Q \in H_2$ such that

$$Q(e^{i\omega})\overline{Q(e^{i\omega})} = 1 \qquad \omega \in [0, 2\pi)$$

That's equivalent to say that

$$W_-(z)\overline{W_-(1/\overline{z})} = W'(z)\overline{W'(1/\overline{z})}$$

General fact: set

$$c_n := Ey(t)y^*(t-n)$$

$$\Phi := \sum_{n=0}^{\infty} c_n z^n$$

Then W is a spectral factor of Φ, i.e.

$$WW^* = \Phi$$

(W^* denotes $\overline{W(1/\overline{z})}^T$) iff $\exists u(t)$ s.t.

$$I_u y(0) = W$$

Corollary 1 *if* W_1, W_2 *are spectral factors, then*

$$Q := W_1 W_2^{-1}$$

is a unitary function, i.e.

$$Q(e^{i\omega})Q(e^{i\omega})^* = I$$

We are interested, in applications, in stable representations, i.e. functions W analytic in the complement $\mathbb{E}$ of $\mathbb{D}$.

Theorem 2 *A stable spectral factor of y has an essentailly unique factorization*

$$W = W_- Q$$

with Q inner and W_- outer (i.e. it essentially generates H_2; if it is rational, and it has no zeros on $i\mathbb{R}$, it is invertible in H_2).

This induces a partial ordering on factors $W_i = W_- Q_i$ for $i = 1, 2$, given by:

$$W_1 < W_2 \Leftrightarrow Q_1^{-1} Q_2 \in H_2$$

i.e. ordering of factors is equivalent to ordering of inner functions.

Assume now that Φ is rational. We say that W is minimal if there is no representation of smaller degree.

The above example (3) was not minimal.

Is there a unique minimal representation?

No!

The set of minimal representations

The outer factor is certainly minimal. Suppose now $W_- = \frac{p(z)}{q(z)}$, with p, q coprime. Then the function

$$Q := \frac{\overline{p(1/\overline{z})}}{p(z)} = \frac{p^*(z)}{p(z)}$$

is inner (in $\mathbb{E}$) and

$$W_+ := W_- Q = \frac{p^*(z)}{q(z)}$$

is also minimal!

To complete the picture, introduce another inner function K

$$K := \frac{q^*(z)}{q(z)}$$

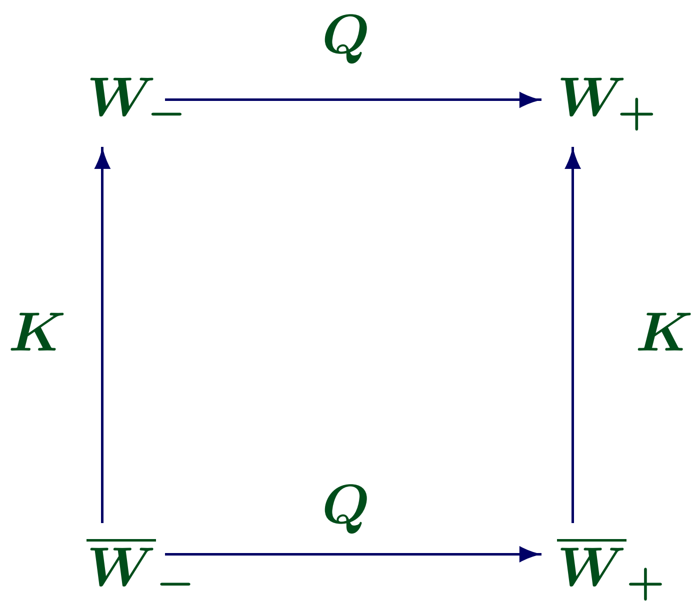
which will flip poles of W_- and W_+. That is, the functions $\overline{W}_-$ and $\overline{W}_+$

$$\overline{W}_- := W_- K^* = \frac{p(z)}{q^*(z)}$$

$$\overline{W}_+ := W_- K^* = \frac{p^*(z)}{q^*(z)}$$

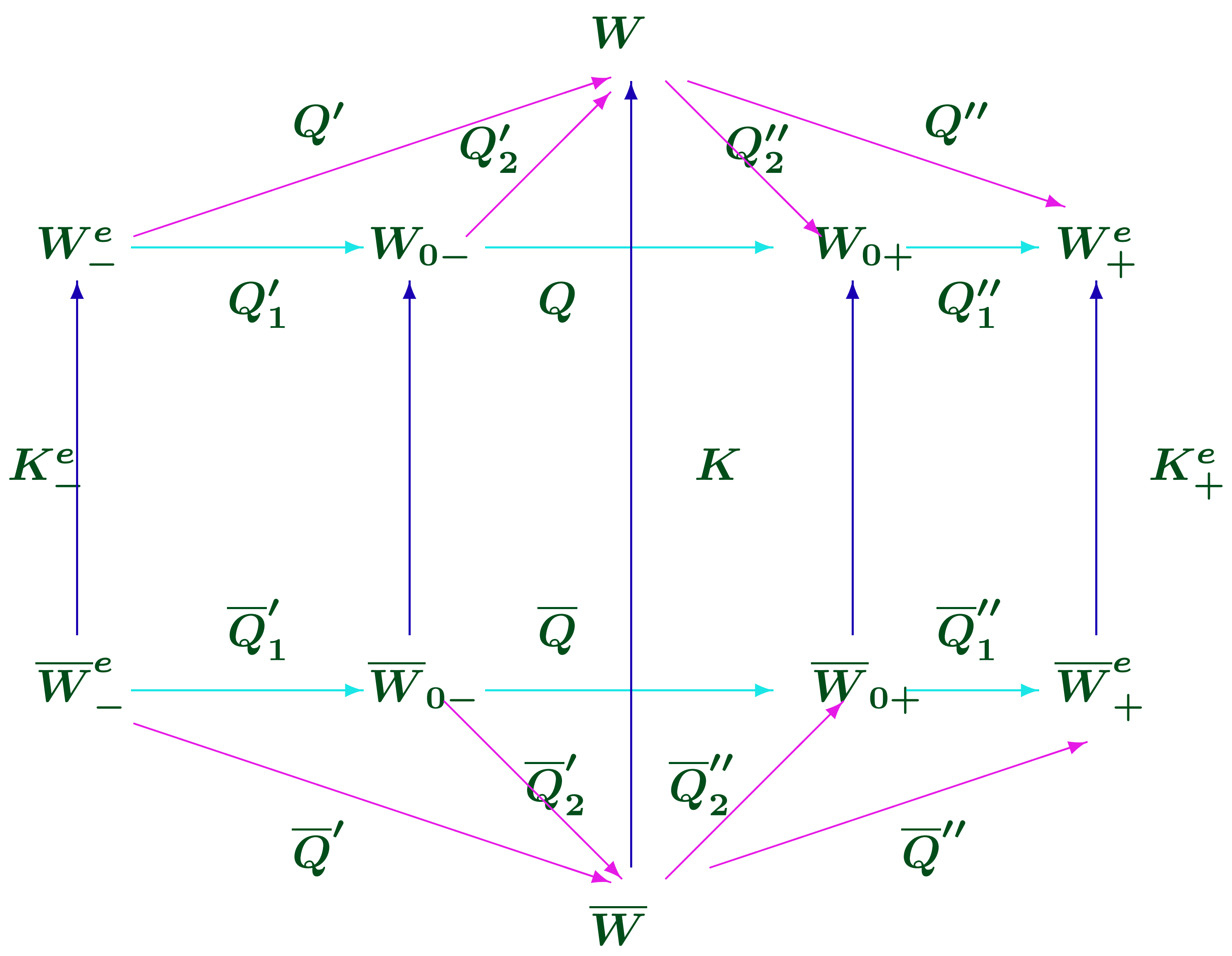
are antistable (their poles are in $\mathbb{E}$). The above is the Douglas-Shapiro-Shields factorization of W_- and W_+.

To summarize, in this case, the situation is very simple



What happens if you consider a general minimal spectral factor of arbitrary width, for a multivariable non full-rank process? (Fuhrmann, G.)

Things get easily out of hand!



W
Q'
Q'_2
Q''_2
Q''
W^e_-
W_{0-}
W_{0+}
W^e_+
Q'_1
Q
Q''_1
K^e_-
K
K^e_+
$\overline{Q}'_1$
$\overline{Q}$
$\overline{Q}''_1$
$\overline{W}^e_-$
$\overline{W}_{0-}$
$\overline{W}_{0+}$
$\overline{W}^e_+$
$\overline{Q}'_2$
$\overline{Q}''_2$
$\overline{Q}'$
$\overline{Q}''$
$\overline{W}$

This mess can be sorted out by exploiting (heavily) the Hilbert space structure.

The idea is to associate to each factor an inner function and thus a coinvariant subspace in H^2 and study the partial ordering of projection operators on these spaces.

Why would you care?

Important tool for designing filters; an intuitive example is the smoothing problem, that is to find the estimate of y in a certain time interval $[T_1, T_2]$, given the observations outside that interval.

More refined applications are in error in the variables models, telecommunications and finance.

In the dynamic errors in the variables problem (cfr. Deistler's talk), we want to decompose a square density $\mathbf{\Phi}$ as

$$\mathbf{\Phi}(z) = \hat{\mathbf{\Phi}}(z) + \mathbf{\Delta}$$

where $\hat{\mathbf{\Phi}}$ is low rank and $\mathbf{\Delta}$ is constant. Can show that this amount to find a suitable factor of the form:

$$\boldsymbol{W} = [\hat{\boldsymbol{W}}, \boldsymbol{D}]$$

(is also has a very interesting approximate version!)

A simple but nice example is in the use of Hankel norm approximation (see Nikolski's and Baratchart's talks).

In fact, it might happen that a W constructed from the data is rational, but of very high degree.

Hankel norm approximation

We defined H_2 in (2).

We define the orthogonal complement

$$\overline{H}_2 := L^2 \ominus H_2$$

and

$$H_\infty := \{f \in L_2; \sup_{0<\rho<1} \operatorname{ess\,sup}_{\omega\in[0,2\pi)} |f(\rho e^{i\omega})| < \infty\}$$

The Hankel operator with symbol W (for $W \in H_\infty$) is defined as

$$\mathcal{H}_W : \overline{H}_2 \mapsto H_2$$
$$\mathcal{H}_W f := P_{H_2} W f \qquad\qquad f \in \overline{H}_2$$

The singular values $\sigma_1 > \sigma_2 > \ldots > \sigma_n$ of $\mathcal{H}_W$ are the square roots of the eigenvalues of $\mathcal{H}_W \mathcal{H}_W^*$ in decreasing order.

We define the Hankel norm of $W \in H_\infty$ as

$$\|W\|_H := \|H_W\|$$

Theorem 3 (AAK) *Let H_W have a (simple) singular value σ_k. Then there exists a unique rational function W_k of degree k, such that*

$$\|W - W_k\|_H = \sigma_k$$

So, Hankel norm approximation provides a unique minimum (and a bound on the L^2 norm). But which factor W should be used?

Lemma 4 *If $W_1 = W_2 Q$ with Q inner, then*

$$\sigma_k^1 \geq \sigma_k^2$$

Corollary 5 (G., Pavon) *The best Hankel-norm approximant of a process $y(t)$ is the one obtained by the outer factor W_-.*