The separation principle in stochastic control, revisited

Workshop in honor of Eduardo Sontag

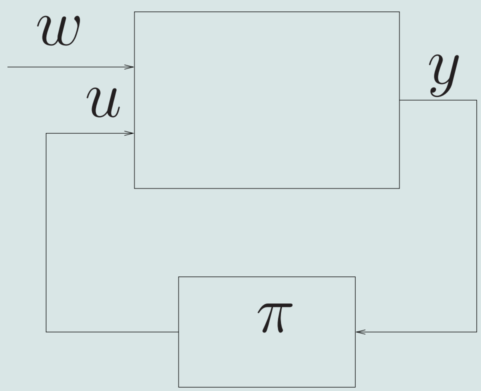
on the occasion of his 60th birthday

Tryphon T. Georgiou

joint work with

Anders Lindquist

linear stochastic system



$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C(t)x(t)dt + D(t)dw \end{cases}$$

$w(t)$ is a vector-valued Wiener process
$x(0)$ is a Gaussian random vector independent of $w(t)$, $y(0) = 0$
A, B_1, B_2, C, D are matrix-valued functions

Goal: Design nonanticipatory control

$$\pi \ : \ y \mapsto u$$

that minimizes

$$J(u) = E\left\{ \int_0^T x(t)'Q(t)x(t)dt \ + \int_0^T u(t)'R(t)u(t)dt + x(T)'Sx(T) \right\}$$

separation priniciple

under suitable assumptions on the class of admissible control $\pi : y \mapsto u$,

the “optimal control” is

$$u(t) = K(t)\hat{x}(t)$$

where $\hat{x}(t) = E\{x(t) \mid \mathcal{Y}_t\}$,

$$\begin{aligned} d\hat{x} &= A(t)\hat{x}(t)dt + B_1(t)u(t)dt \\ &\quad +L(t)(dy - C(t)\hat{x}(t)dt) \\ \hat{x}(0) &= 0. \end{aligned}$$

with $K(t)$ and $L(t)$ computed via a pair of dual Riccati equations

NB:

— attempts to prove separation for $u(t)$ is $\mathcal{Y}_t$ measurable (a.s.). . .

— too big a class; we know no proof which is correct (strong solutions)

historical remarks

Wonham, Kushner, Lindquist, Fleming & Rishel

- treatment overburdened with technicalities
- folk accounts not supported by existing proofs
- non-Gaussian nature due to an *a-priori* nonlinear π is often overlooked

---

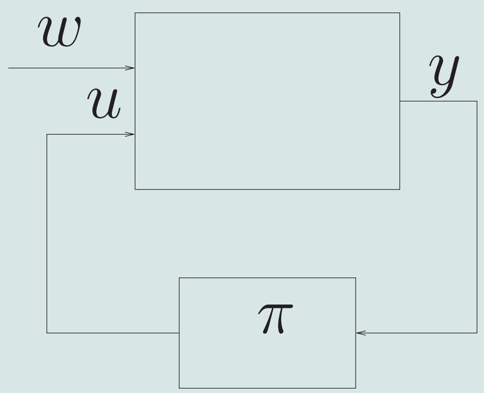
- *herein, separation principle for:*

– the most natural class of controls
all linear/nonlinear and even discontinuous
such that feedback loop makes "engineering" sense

– engineering view point: signals = sample functions

– general semimartingale driving noise, with jumps

– delay-differential linear systems, etc.



the standard "completion of squares"

$$J(u) = E\left\{ x(0)'P(0)x(0) + \int_0^T (u - Kx)'R(u - Kx)dt \right\} + \int_0^T \mathrm{tr}(B_2'PB_2)dt$$

where

$$\begin{cases} \dot{P} = -A'P - PA + PB_1R^{-1}B_1'P - Q \\ P(T) = S \end{cases}$$

$$K(t) := -R(t)^{-1}B_1(t)'P(t).$$

using Itô's rule:

$$\begin{aligned} d(x'Px) &= x'\dot{P}xdt + 2x'Pdx + \mathrm{tr}(B_2'PB_2)dt \\ &= [-x'Qx - u'Ru + (u - Kx)'R(u - Kx) + \mathrm{tr}(B_2'PB_2)]dt + 2x'PB_2dv \end{aligned}$$

with "complete state-information":

$$u_{\mathrm{optimal}}(t) = K(t)x(t)$$

incomplete state information

$u(t)$ *needs to be a function of* $\{y(s); 0 \leq s \leq t\}$

Standard recipe:

$$u(t) = K(t)\hat{x}(t)$$

where

$$\hat{x}(t) = E\{x(t) \mid \mathcal{Y}_t\}$$

justification $\Leftrightarrow$ *separation theorem*

where is the potential problem?

set

$$\tilde{x}(t) := x(t) - \hat{x}(t)$$

then

$$E\int_0^T (u - Kx)' R (u - Kx) dt = E\int_0^T [(u - K\hat{x})' R (u - K\hat{x})] dt + \mathrm{tr}(K' R K \Sigma)$$

since $E\{[u(t) - K(t)\hat{x}(t)]\tilde{x}(t)'\} = 0$,

and where $\Sigma(t) := E\{\tilde{x}(t)\tilde{x}(t)'\}$

why isn't obvious that $u = K\hat{x}$ is optimal?

subtlety: in general, Σ may depend on the control

source of fallacy (?)

due to linearity

$$x(t) = x_0(t) + \int_0^t \Phi(t,s)B_1(s)u(s)ds$$

the control term cancels out:

$$\tilde{x}(t) = \tilde{x}_0(t) := x_0(t) - \hat{x}_0(t),$$

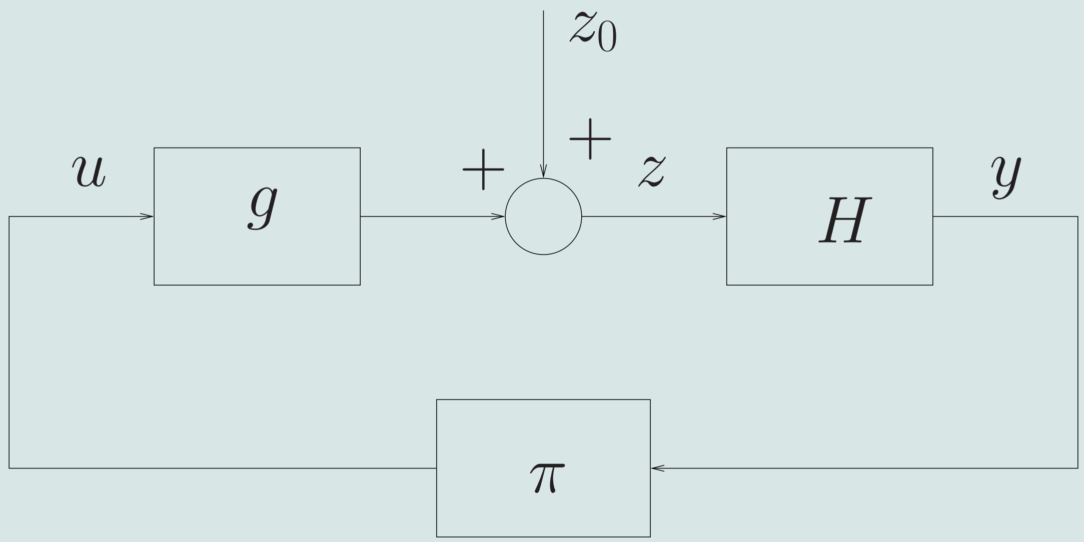
where $\hat{x}_0(t) := E\{x_0(t) \mid \mathcal{Y}_t\}$

how could $E\{\tilde{x}_0(t)\tilde{x}_0(t)'\}$ depend on the control?

because the filtration $\mathcal{Y}_t$, and hence $\hat{x}_0$, might depend on u!

— u is in general a nonlinear function of y
— hence, y may not be Gaussian
— despite the fact that x_0 is Gaussian,
$\hat{x}_0(t) = E\{x_0(t) \mid \mathcal{Y}_t\}$ may not be linear in the data $\{y(\tau);\ \tau \in [0,t]\}$
— $\hat{x}_0(t)$ may not be given by a Kalman filter.

generalization - notation



$$z(t) = z_0(t) + \int_0^t G(t,\tau)u(\tau)d\tau$$
$$y(t) = Hz(t)$$

where

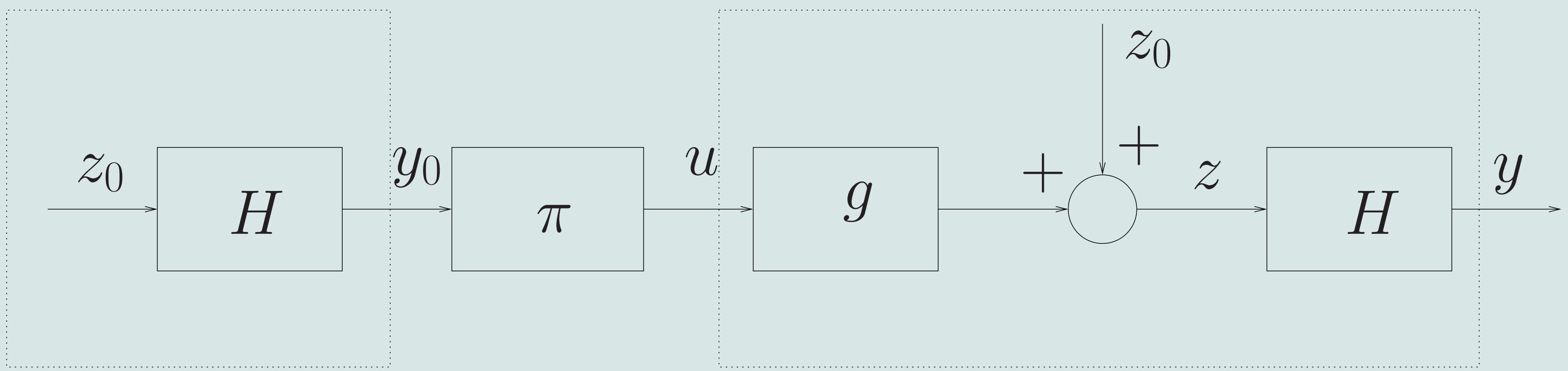
$$g \ : \ (t,u) \mapsto \int_0^t G(t,\tau)u(\tau)d\tau$$

E.g., $z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ and $H = [0, I]$

ways out (?)

SOL: stochastic open loop (Lindquist)

limit control so as to be adapted to $\{\mathcal{Y}_t^0\}$

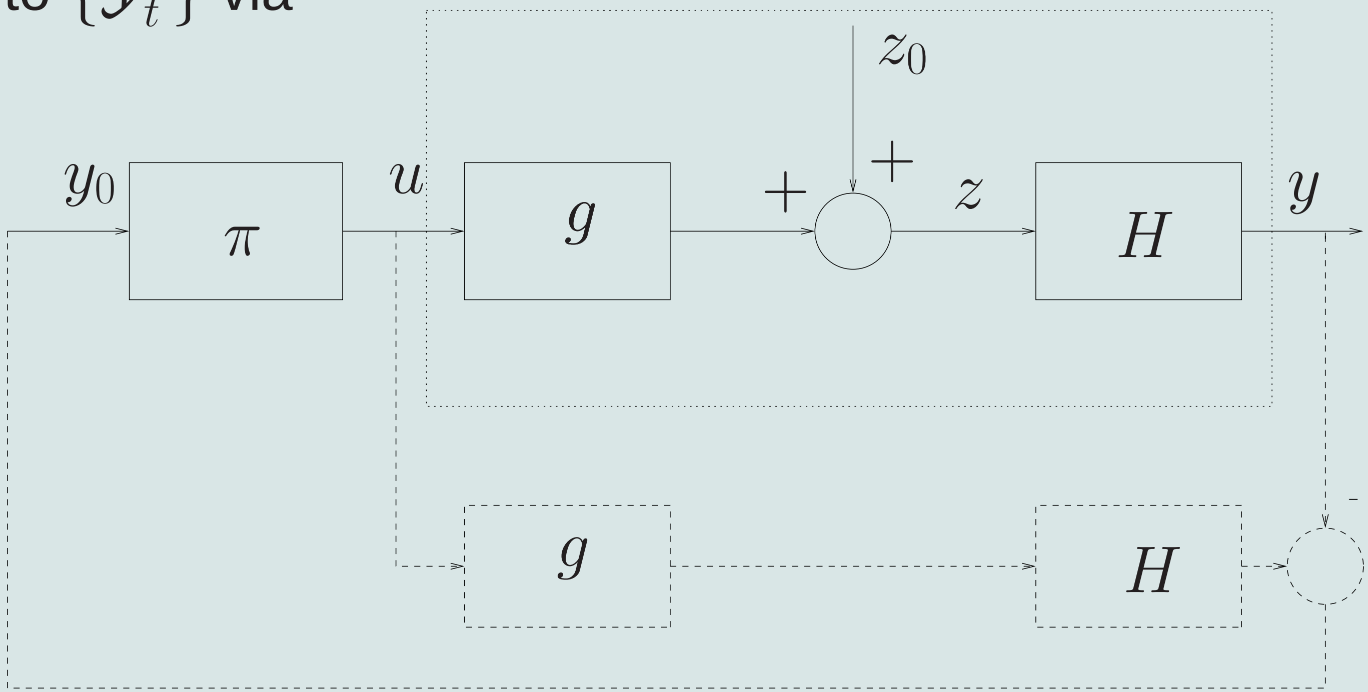


examples

— linear control

— Lipschitz feedback

e.g., control adapted to $\{\mathcal{Y}_t^0\}$ via



example: linear feedback

$$u(t) = u_{\mathrm{deterministic}} + \int_0^t F(t,\tau)dy$$

then the Gaussian character is preserved.

It can be shown that $\mathcal{Y}_t = \mathcal{Y}_t^0$.

Hence,

$$d\tilde{x} = (A - LC)\tilde{x}dt + (B_2 - LD)dw$$
$$\tilde{x}(0) = x(0)$$

$\Sigma(t) := E\{\tilde{x}(t)\tilde{x}(t)'\}$ is independent of u

$$u(t) = \int_0^t F(t,\tau)dy(\tau) \Rightarrow dy = dy_0 + \int_0^t M(t,s)u(s)dsdt$$

$$\Rightarrow dy = dy_0 + \int_0^t N(t,\tau)dy(\tau)dt$$

where $\quad N(t,\tau) = \int_\tau^t M(t,s)F(s,\tau)ds$

Volterra resolvent $\quad R(t,\tau) = \int_\tau^t R(t,s)N(s,\tau)ds + N(t,s)$

Then

$$\int_0^t N(t,\tau)dy(\tau) = \int_0^t R(t,\tau)dy_0(\tau)$$

$$\Rightarrow \quad dy = dy_0 + \int_0^t R(t,\tau)dy_0(\tau)dt$$

$$\Rightarrow \quad \sigma\{y(\tau); 0 \leq \tau \leq t\} = \sigma\{y_0(\tau); 0 \leq \tau \leq t\}$$

example: Lipschitz continuous control

[Wonham] Assuming that

$$dy(t) = x(t)dt + D(t)dw(t)$$

i.e., $C(t) = I$ is invertible!
Then among control laws of the form

$$u(t) = \psi(t, \hat{x}(t))$$

the choice $u(t) = K(t)\hat{x}(t)$ is optimal.

[Fleming & Rishel]
removed the assumption on $C(t)$; Lipschitz on y; simpler proof.

example: Lipschitz (cont.)

[Kushner] $$\hat{\xi}_0(t) := E\{x_0(t) \mid \mathcal{Y}_t^0\}$$

given by the Kalman filter

$$d\hat{\xi}_0 = A\hat{\xi}_0(t)dt + L(t)dv_0, \; \hat{\xi}_0(0) = 0$$
$$dv_0 = dy_0 - C\hat{\xi}_0(t)dt, \qquad v_0(0) = 0$$

define

$$\hat{\xi}(t) := \hat{\xi}_0(t) + \int_0^t \Phi(t,s)B_1(s)u(s)ds$$

and assume

$u(t) = \psi(t, \hat{\xi}(t))$ *is Lipschitz*

Then $\hat{\xi}$ is the unique strong solution of

$$d\hat{\xi} = \big(A\hat{\xi}(t) + B_1\psi(t, \hat{\xi}(t))\big)dt + L(t)dv_0, \; \hat{\xi}(0) = 0.$$

This choice force u to be adapted to $\{\mathcal{Y}_t^0\}$ $\Rightarrow \; \{\mathcal{Y}_t^0\} = \{\mathcal{Y}_t\} \; \Rightarrow \; \hat{\xi} = \hat{x}$

example: delay in the loop

when $u(t)$ is a function of $y(\tau);\ 0 \le \tau \le t - \varepsilon$,

$$\mathcal{Y}_t = \mathcal{Y}_t^0$$

the possibility of a control-dependent σ-field
does not arise in the usual (predictive) discrete-time formulation

— Taking $\epsilon \longrightarrow 0$ and general nonlinear feedback
there is no guarantee that $\mathcal{Y}_t$ is left-continuous

— “Proofs” of separation using such limits are circular,
misleading accounts in textbooks.

signals and systems

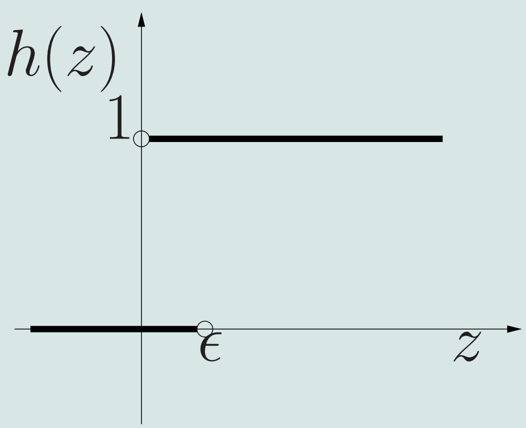
signals :
sample paths; possibly having bounded discontinuities in D (càdlàg – Skorokhod space)

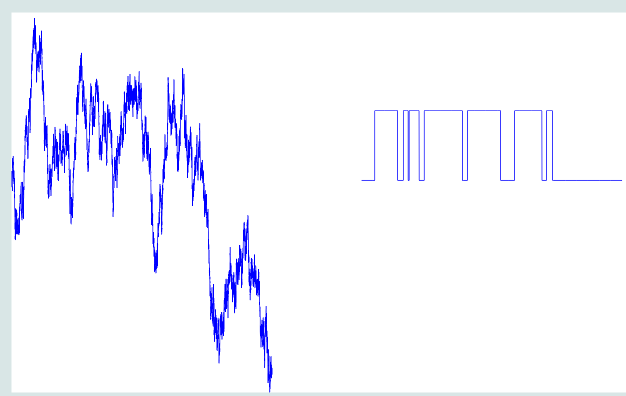
systems: measurable nonanticipatory maps

examples:

i) SDE's that have strong solutions

ii) nonlinearities, hysteresis ($C \longrightarrow D$), etc.

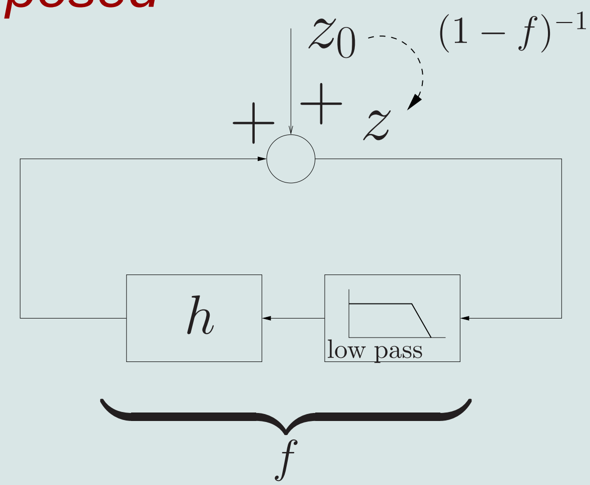


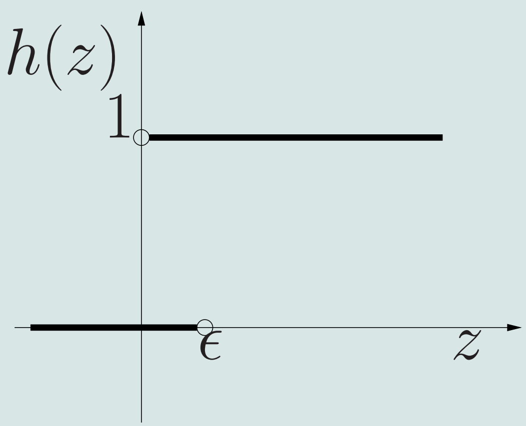


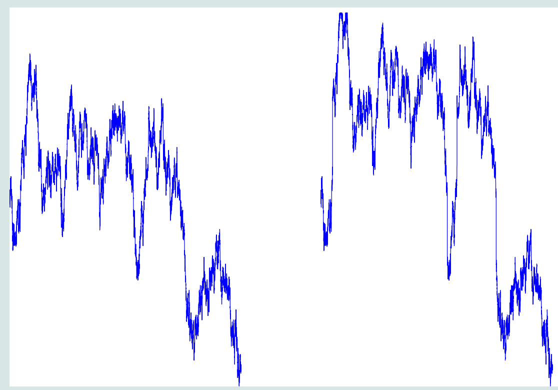
$z(t) \longrightarrow h(z(t))$

well-posedness of feedback

Defn. a feedback loop, that is $z = z_0 + f(z)$ is *well-posed* if it has a unique solution in D for all $z_0 \in D$ and $(1-f)^{-1}$ is a system.





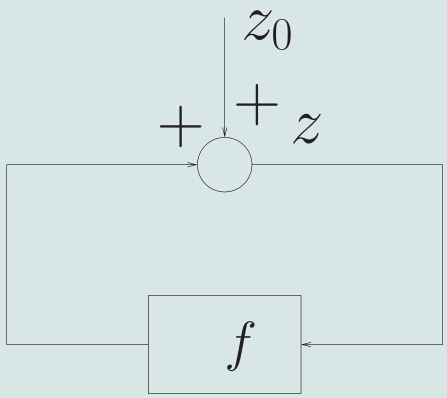


$z_0(t)$ $\qquad$ $z(t) = (1-f)^{-1} z_0(t)$

well-posedness (cont.)

by defn z, z_0 stochastic processes
well-posedness implies that

$$\mathcal{Z}_t^0 = \mathcal{Z}_t, \quad t \in [0, T].$$



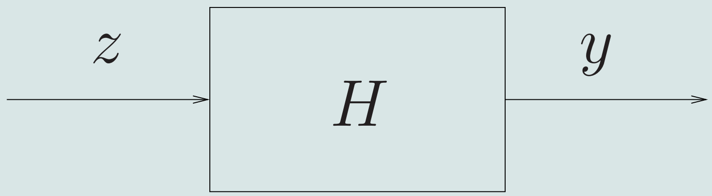
$(1-f)$ *and* $(1-f)^{-1}$ *are systems*

$\Rightarrow$ $z_0 = z - f(z)$ *and* $z = (1-f)^{-1} z_0$

NB.

— no more information other than what is contained in $\mathcal{Z}_t^0$

how about incomplete state-information?



$$z_1 = \begin{pmatrix} w \\ 0 \end{pmatrix}, \quad z_2 = \begin{pmatrix} 0 \\ w \end{pmatrix}$$

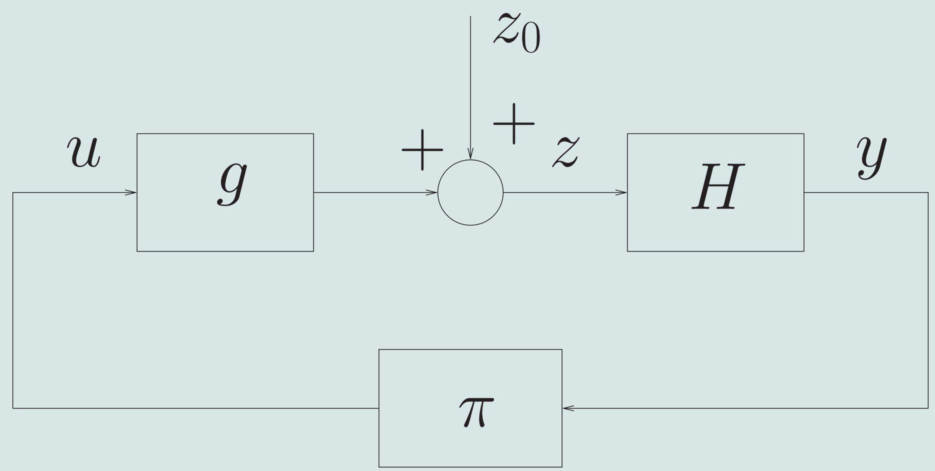
generate the same filtrations, i.e., $\mathcal{Z}_t^1 = \mathcal{Z}_t^2$

while for $H = \begin{pmatrix} 1 & 0 \end{pmatrix}$,

$$y_1 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} w \\ 0 \end{pmatrix}, \quad y_2 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ w \end{pmatrix}$$

do not, i.e., $\mathcal{Y}_t^1 \neq \mathcal{Y}_t^2$.

linear read-out map



Assume

$$z(t) = z_0(t) + g \circ \pi(y(t))$$
$$y(t) = Hz(t)$$

is well-posed with H linear,

it follows that

$$\mathcal{Y}_t = \mathcal{Y}_t^0, \quad t \in [0, T].$$

Proof:

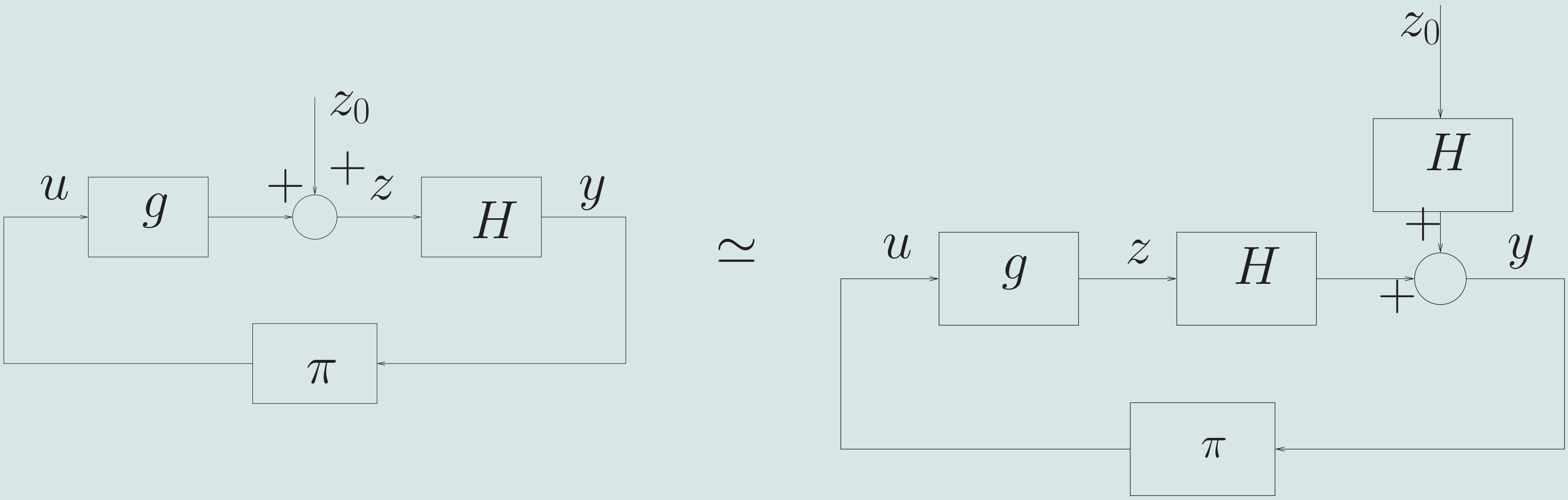
$$\begin{aligned}(1 - Hg\pi)H &= H - Hg\pi H \\ &= H(1 - g\pi H)\end{aligned}$$

$$H(1 - g\pi H)^{-1} = (1 - Hg\pi)^{-1}H$$

$\Rightarrow y = (1 - Hg\pi)^{-1}y_0$, and $y_0 = (1 - Hg\pi)y$.

essence of the lemma

well-posedness resolves the issue of circular control dependence



the separation principle

thm: assuming

$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C(t)x(t)dt + D(t)dw \end{cases}$$

$w(t)$ is a vector-valued Wiener process
$x(0)$ is a Gaussian random vector independent of $w(t)$, $y(0) = 0$
A, B_1, B_2, C, D are matrix-valued functions
there is a unique control law $\pi \; : \; y \mapsto u$ minimizing

$$J(u) = E\left\{ \int_0^T x(t)'Q(t)x(t)dt \; + \int_0^T u(t)'R(t)u(t)dt + x(T)'Sx(T) \right\}$$

in the class of well-posed control laws, and has the form

$$u(t) = K(t)\hat{x}(t)$$

the separation principle (general)

thm: for the same linear system, assuming
w is a *semimartingale* and $x(0)$ an independent random vector
the unique optimal control
in the class of well-posed controllers
is given by

$$u(t) = K(t)\hat{x}(t)$$

where $\hat{x}$ is the conditional mean.

remarks: no need for Lipschitz continuity
allows jump processes
$K(t)$ is still given by a Riccati equation
in general, the difficult part is constructing $\hat{x}(t) = E\{x(t)|\mathcal{Y}_t\}$.

Proof: i) $\mathcal{Y}_t = \mathcal{Y}_t^0, \quad t \in [0,T]$.

ii) completion-of-squares using Itô's rule:

$$x(T)'Px(T) - x(0)'Px(0) = f_\Delta + \\ + \int_0^T \{x'\dot{P}x dt + 2x'P dx + d\,\mathrm{tr}([x,x']P)\}$$

iii) $x(t) = \int_0^t \Phi(t,s)\big(A(s)x(s) + B_1(s)u(s)\big)ds + v(t)$

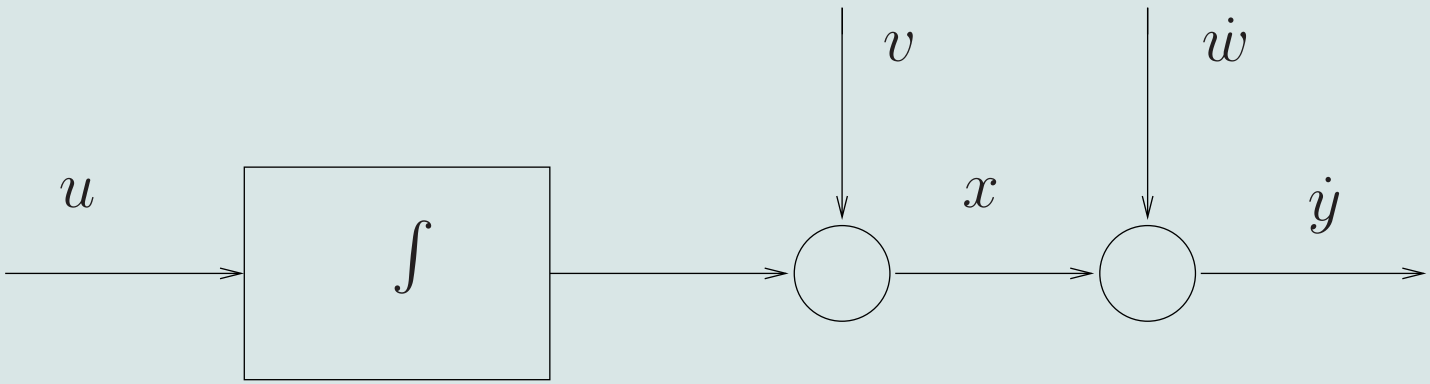
i.e., continuous/BV $+v(t)$ where $dv = B_2 dw$

$$\Rightarrow$$

$$\begin{aligned} \mathrm{iii_a}) \ [x,x'] &= [v,v'] \text{ independent of } u \\ \mathrm{iii_b}) \quad f_\Delta &= \sum_{s \le T} \Big[(x(s)'P(s)(x(s) - x(s_-)'P(s)x(s_-) \\ &\qquad -2x(s_-)'P(s)\Delta_s - \Delta_s'P(s)\Delta_s\Big] \\ &= 0 \end{aligned}$$

where $\Delta_s := x(s) - x(s_-)$.

example: step change in white noise



$$v(t) = \begin{cases} 1 & t \geq \tau \\ 0 & t < \tau \end{cases}$$

with τ exponentially distributed

minimize $E\left\{\int_0^T (x^2 + u^2)dt\right\}$

$$\begin{cases} dx = u(t)dt + dv, \; x(0) = 0, \\ dy = x(t)dt + dw \end{cases}$$

i) Wonham-Shiryaev filter:

$$d\hat{x} = (1-\hat{x})dt + udt + \hat{x}(1-\hat{x})(dy - \hat{x}dt)$$

ii) optimal feedback:

$$u(t) = -p(t)\hat{x}(t)$$

where $\dot{p} = p^2 - 1 \Rightarrow p(t) = \tanh(T-t)$.

iii) cost: since $[v,v](t) = v(t)$,

$$E\left\{\int_0^T p(t)d[v,v](t)\right\} = E\left\{\int_\tau^T p(t)dt\right\}$$

$$= \ln(\cosh T)(1-e^{-T}) - \int_0^T \ln(\cosh t)e^{-t}dt.$$

separation for delay-differential systems

$$\begin{cases} dx = A_1(t)x(t)dt + A_2(t)x(t-h)dt \\ \quad + \int_{t-h}^{t} A_0(t,s)x(s)dsdt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C_1(t)x(t)dt + C_2(t)x(t-h)dt + D(t)dw \end{cases}$$

more generally

$$\begin{cases} dx = \int_{t-h}^{t} d_s A(t,s)x(s)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = \int_{t-h}^{t} d_s C(t,s)x(s)dt + D(t)dw \end{cases}$$

determine π to minimize

$$E\left\{\int_0^T x(t)'Q(t)x(t)d\alpha(t) + \int_0^T u(t)'R(t)u(t)dt\right\}$$

System can be written in the form:

$$z(t) = z_0(t) + \int_0^t G(t,\tau)u(\tau)d\tau$$

$$y(t) = H(t)z(t)$$

Deterministic optimal control

Deterministic optimal control problem (with $w = 0$) is

$$u_{\text{optimal}}(t) = \int_{t-h}^{t} d_\tau K(t,\tau)x(\tau)$$

separation thm for delay systems

w a Gaussian martingale
over all feedback laws π that are <u>well-posed</u>
the unique optimal control law is given by

$$u(t) = \int_{t-h}^{t} d_s K(t,s)\hat{x}(s|t)$$

with

$$\hat{x}(s|t) := E\{x(s) \mid \mathcal{Y}_t\}$$

is given by a linear (distributed) filter *[Lindquist]*

$$\begin{aligned} d\hat{x}(t|t) &= \int_{t-h}^{t} d_s A(t,s)\hat{x}(s|t)dt + B_1 u dt + X(t,t)dv \\ d_t\hat{x}(s|t) &= X(s,t)dv,\ s \le t \\ dv &= dy - \int_{t-h}^{t} d_s C(t,s)\hat{x}(s|t)dt,\ v(0) = 0 \end{aligned}$$

Key points

— well-posedness $+$ linearity $\Rightarrow$ control-independent σ-field

— separation principle holds over a wide class of nonlinear control:
$u = K\hat{x}$ is optimal

— noise: semi-martingale with possible jumps

Happy birthday Eduardo!!!