Control Theory: From Classical to Quantum Optimal, Stochastic, and Robust Control

Notes for Quantum Control Summer School, Caltech, August 2005

M.R. James*
Department of Engineering
Australian National University
Matthew.James@anu.edu.au

Contents

*This work was supported by the Australian Research Council.

1 Introduction

The purpose of these notes is to provide an overview of some aspects of optimal and robust control theory considered relevant to quantum control. The notes begin with classical deterministic optimal control, move through classical stochastic and robust control, and conclude with quantum feedback control. Optimal control theory is a systematic approach to controller design whereby the desired performance objectives are encoded in a cost function, which is subsequently optimized to determine the desired controller. Robust control theory aims to enhance the robustness (ability to withstand, to some extent, uncertainty, errors, etc) of controller designs by explicitly including uncertainty models in the design process. Some of the material is in continuous time, while other material is written in discrete time. There are two underlying and universal themes in the notes: *dynamic programming* and *filtering*.

Dynamic programming is one of the two fundamental tools of optimal control, the other being Pontryagin's principle, [24]. Dynamic programming is a means by which candidate optimal controls can be verified optimal. The procedure is to find a suitable solution to a *dynamic programming equation* (DPE), which encodes the optimal performance, and to use it to compare the performance of a candidate optimal control. Candidate controls may be determined from Pontryagin's principle, or directly from the solution to the DPE. In general it is difficult to solve DPEs. Explicit solutions exist in cases like the linear quadratic regulator, but in general approximations must usually be sought. In addition, there are some technical complications regarding the DPE. In continuous time, the DPE is a nonlinear PDE, commonly called the *Hamilton-Jacobi-Bellman* (HJB) equation. The complications concern differentiability, or lackthereof, and occur even in "simple" classical deterministic problems, section 2. This is one reason it can be helpful to work in discrete time, where such regularity issues are much simpler (another reason for working in discrete time is to facilitate digital implementation).

Filtering concerns the processing of measurement information. In optimal control, filters are used to represent information about the system and control problem of interest. In general, this information is incomplete, i.e. the state is typically not fully accessible, and may be corrupted by noise. To solve optimal control problems in these situations, the cost function is expressed in terms of the state of a suitably chosen filter, which is often called an *information state*. Dynamic programming can then be applied using the information state dynamics. The nature of the measurements and the purpose for which the data is to be used determine the architecture of the filter. In stochastic situations, this is closely linked to the probabilistic concept of conditional expectation. The famous *Kalman filter* computes dynamically conditional expectations (of states given measurements in linear gaussian models), which are also optimal estimates in the mean square error sense. The quantum *Belavkin filter*, or stochastic master equation, also computes a quantum version of conditional expectation. In linear gaussian cases, the information states are gaussian, a fact which considerably simplifies matters due to the finite number of parameters. Filters such as these based on computing conditional expectations of states or system variables do not include any information about the cost or performance objective. While this is not an issue for many problems such as LQG, where the task of estimation can be completely decoupled from that of control [17], there are important problems where the filter dynamics must be modified to take into account the control objective. These problems include LEQG [48, 49] or risk-sensitive control [8, 37], and H^∞ robust control [19, 54].

Figure 1 shows a physical system being controlled in a feedback loop. The so-called *separation structure* of the controller is shown. The control values are computed in the box marked "control", using a function of the information state determined using dynamic programming. The information state, as has been mentioned, is the state of the filter whose dynamics are built into the box marked "filter". This structure embodies the two themes of these notes.

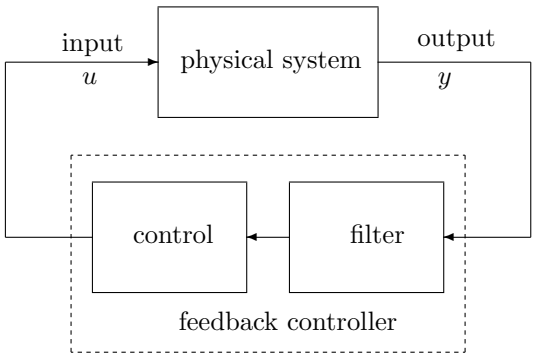

Figure 1: Feedback controller showing the separation structure.

These notes were assembled from various lecture notes and research papers, and so we apologize for the inevitable inconsistencies that resulted.

2 Deterministic Dynamic Programming and Viscosity Solutions

References for this section include [24], [25], [3], [15].

2.1 Introduction

2.1.1 Preamble

Hamilton-Jacobi (HJ) equations are nonlinear first-order partial differential equations of the form

$$F(x, V(x), \nabla V(x)) = 0 \tag{1}$$

(one can also consider second-order equations but we do not do so here). $V(x)$ ($x \in \Omega \subset \mathbf{R}^n$) is the unknown function to be solved for, and $\nabla V(x) = (\frac{\partial V(x)}{\partial x_1}, \ldots, \frac{\partial V(x)}{\partial x_n})$ denotes the gradient. $F(x, v, \lambda)$ is a nonlinear function.

HJ equations have a long history, dating back at least to the calculus of variations of the 19th century, and HJ equations find wide application in science, engineering, etc. Perhaps surprisingly, it was only relatively recently that a satisfactory general notion of solutions for (1) became available, with the introduction of the concept of *viscosity solution* (Crandall-Lions, c. 1980). The difficulty, of course, is that solutions are not in general globally smooth (e.g. C^1). Solutions are often smooth in certain regions, in which the famous *method of characteristics* may be used to construct solutions. There are a number of other notions of solution available, such as encountered in non-smooth analysis (e.g. *proximal solution*), though we will not discuss them here.

In Engineering our principal interest in HJ equations lies in their connection with optimal control (and games) via the *dynamic programming* methodology. The *value function* is a solution to an HJ equation, and solutions of HJ equations can be used to test a controller for optimality, or perhaps to construct a feedback controller. In these notes we discuss dynamic programming and viscosity solutions in the context of two examples, and make some mention of the general theory.

2.1.2 Optimal Control

As a first and perhaps familiar example (e.g. LQR), let's consider a finite time horizon optimal control problem defined on a time interval $[t_0, t_1]$:

$$J^*(t_0, x_0) = \inf_{u(\cdot)} J(t_0, x_0, u(\cdot)) \tag{2}$$

Here, x_0 is the initial state at time t_0, and $u(\cdot)$ is the control; $J(t_0, x_0, u(\cdot))$ represent the associated cost.

To be specific, and to prepare us for dynamic programming, suppose one wants to minimize the cost functional

$$J(t, x; u(\cdot)) \;=\; \int_t^{t_1} L(x(s), u(s))\, ds + \psi(x(t_1)), \tag{3}$$

where $x(\cdot)$ is the solution of the initial value problem

$$\begin{cases} \dot{x}(s) \;=\; f(x(s), u(s)), \quad t \le s \le t_1, \\ \\ x(t) \;=\; x. \end{cases} \tag{4}$$

Here, $t \in [t_0, t_1]$ is a "variable" initial time, $u(\cdot)$ is a control defined on $[t, t_1]$ taking values in, say, $U \subset \mathbf{R}^m$ (U closed), and $x(\cdot)$ is the state trajectory in $\mathbf{R}^n$. We denote by $\mathcal{U}_{t,t_1}$ a space of admissible controls, containing at least the piecewise continuous controls.

The value function is defined by

$$V(t,x) \;=\; \inf_{u(\cdot)\in\mathcal{U}_{t,t_1}} J(t,x;u(\cdot)) \tag{5}$$

for $(t,x) \in [t_0,t_1] \times \mathbf{R}^n$. The *dynamic programming principle* states that for every $r \in [t,t_1]$,

$$V(t,x) \;=\; \inf_{u(\cdot)\in\mathcal{U}_{t,r}} \left[\int_t^r L(x(s),u(s))\,ds + V(r,x(r)) \right] \tag{6}$$

(we will prove this later on). From this, one can derive formally the equation

$$\frac{\partial}{\partial t} V(t,x) + H(x, \nabla_x V(t,x)) \;=\; 0 \;\text{ in } (t_0,t_1) \times \mathbf{R}^n, \tag{7}$$

with terminal data

$$V(t_1,x) \;=\; \psi(x) \;\text{ in } \mathbf{R}^n. \tag{8}$$

Here, the *Hamiltonian* is given by

$$H(x,\lambda) \;=\; \inf_{v\in U} \{\lambda \cdot f(x,v) + L(x,v)\} \tag{9}$$

The nonlinear first order PDE (7) is the *dynamic programming PDE* or Hamilton-Jacobi-Bellman (HJB) equation. The pair (7), (8) specify what is called a *Cauchy problem*, and can be viewed as a special case of (1) together with suitable boundary conditions, using $\Omega = (t_0,t_1) \times \mathbf{R}^n$. Notice that the Hamiltonian (9) is *concave* in the variable λ (since it is the infimum of linear functions).

Let us see how (7) is obtained. Set $r = t + h$, $h > 0$, and rearrange (6) to yield

$$\inf_{u(\cdot)} \left[\frac{1}{h}(V(t+h, x(t+h)) - V(t,x)) + \frac{1}{h}\int_t^{t+h} L(x(s),u(s))\,ds \right] \;=\; 0.$$

If V and $u(\cdot)$ are sufficiently smooth, then

$$\frac{1}{h}(V(t+h,x(t+h)) - V(t,x)) \to \frac{\partial}{\partial t}V(x,t) + \nabla_x V(x,t)\cdot f(x,u(t)) \;\text{ as }\; h \to 0$$

and

$$\frac{1}{h}\int_t^{t+h} L(x(s),u(s))\,ds \to L(x,u(t)) \;\text{ as }\; h \to 0.$$

Combining these displays one is formally led to (7). A proof of (7) when V is sufficiently smooth requires a careful derivation of two inequalities which combine to give (7). Below we will prove that V is a viscosity solution of (7); in fact, the unique one satisfying the terminal condition (8).

Verification. Let $\tilde{V}(t,x)$ be a C^1 solution of (7), (8). Let $u(\cdot) \in \mathcal{U}_{t_1,t_1}$ be any control. Then using (7)

$$\begin{aligned} \tfrac{d}{dt}\tilde{V}(t,x(t)) &= \tfrac{\partial}{\partial t}\tilde{V}(t,x(t)) + \nabla\tilde{V}(t,x(t))\dot{x}(t) \\ &= \tfrac{\partial}{\partial t}\tilde{V}(t,x(t)) + \nabla\tilde{V}(t,x(t))f(x(t),u(t)) \\ &\geq -L(x(t),u(t)) \end{aligned}$$

Integrating, we get

$$\tilde{V}(t_1,x(t_1)) - \tilde{V}(t_0,x_0) \geq -\int_{t_0}^{t_1} L(x(t),u(t))dt$$

or

$$\begin{aligned} \tilde{V}(t_0,x_0) &\leq \textstyle\int_{t_0}^{t_1} L(x(t),u(t))dt + \tilde{V}(t_1,x(t_1)) \\ &= \textstyle\int_{t_0}^{t_1} L(x(t),u(t))dt + \psi(x(t_1)) \end{aligned}$$

using (8). This shows that $\tilde{V}(t_0,x_0) \leq V(t_0,x_0)$ (V is the value function defined by (5)). Now this same calculation for the control $u(\cdot) = u^*(\cdot) \in \mathcal{U}_{t_0,t_1}$ satisfying

$$u^*(t) \in \operatorname*{argmin}_{v\in U}\left\{\nabla_x\tilde{V}(t,x^*(t))\cdot f(x^*(t),v) + L(x^*(t),v)\right\}, \tag{10}$$

for $t \in [t_1,t_1]$, where $x^*(\cdot)$ is the corresponding state trajectory, gives

$$\tilde{V}(t_0,x_0) = \int_{t_0}^{t_1} L(x^*(t),u^*(t))dt + \psi(x^*(t_1))$$

showing that in fact u^* is optimal and $\tilde{V}(t_0,x_0) = V(t_0,x_0)$. Indeed we have $\tilde{V} = V$ in $[t_0,t_1]\times\mathbf{R}^n$ by this arguement, and so we have shown that any *smooth* solution to (7), (8) must equal the value function - this is a uniqueness result. Unfortuneatly, in general there may be no such smooth solutions.

Optimal feedback. The above calculations suggest how one might obtain an optimal *feedback* controller. To simplify a bit, suppose that

$$U = \mathbf{R}^m, \quad f(x,u) = f(x) + g(x)u, \quad L(x,u) = \ell(x) + \tfrac{1}{2}|u|^2.$$

Then evaluating the infimum in (9) gives

$$u^* = -g(x)'\lambda'$$

and

$$H(x,\lambda) = \lambda f(x) - \tfrac{1}{2}\lambda g(x)g(x)'\lambda' + \ell(x).$$

Hence the HJB equation can be written as

$$\frac{\partial}{\partial t}V + \nabla V f - \tfrac{1}{2}\nabla V g g' \nabla V' + \ell = 0 \tag{11}$$

with optimal feedback controller

$$u^*(t,x) = -g(x)'\nabla V(t,x)'. \tag{12}$$

This means that the optimal control $u^*(\cdot) \in \mathcal{U}$ is given by

$$u^*(t) = u^*(t, x^*(t)), \quad t_0 \le t \le t_1.$$

Of course, this makes sense only when V is sufficiently smooth.

The equation (11) is sometimes refered to as a *nonlinear Riccati equation*.

LQR. Take

$$U = \mathbf{R}^m, \quad f(x,u) = Ax + Bu, \quad L(x,u) = \tfrac{1}{2}|x|^2 + \tfrac{1}{2}|u|^2, \quad \psi(x) = \tfrac{1}{2}x'\Psi x.$$

As a trial solution of (11) we use

$$\tilde{V}(t,x) = \tfrac{1}{2}x'P(t)x,$$

where $P(t) \ge 0$ (symmetric) is to be determined. Now

$$\frac{\partial}{\partial t}\tilde{V}(t,x) = \tfrac{1}{2}x'\dot{P}(t)x, \quad \text{and} \quad \nabla V(t,x) = x'P(t).$$

Plugging these into (11) gives

$$\tfrac{1}{2}x'\dot{P}(t)x + x'P(t)Ax - \tfrac{1}{2}x'P(t)BB'P(t)x + \tfrac{1}{2}x'x = 0.$$

Since this holds for all $x \in \mathbf{R}^n$ we must have

$$\dot{P}(t) + A'P(t) + P(t)A - P(t)BB'P(t) + I = 0. \tag{13}$$

At time $t = t_1$ we have $\tilde{V}(t_1, x) = \frac{1}{2}x'\Psi x$, and so

$$P(t_1) = \Psi. \tag{14}$$

Therefore if there exists a C^1 solution $P(t)$ to the Riccati differential equation (13) on $[t_0, t_1]$ with terminal condition (14) we obtain a smooth solution $\frac{1}{2}x'P(t)x$ to (7), (8), and as argued above the value function for the LQR problem is given by

$$V(t,x) = \tfrac{1}{2}x'P(t)x. \tag{15}$$

The optimal feedback controller is given by

$$u^*(t,x) = -B'P(t)x. \tag{16}$$

This gives the optimal control $u^*(\cdot) \in \mathcal{U}$:

$$u^*(t) = -B'P(t)x^*(t), \quad t_0 \le t \le t_1. \tag{17}$$

2.1.3 Distance Function

As another example, we consider the distance function $d(x, \partial\Omega)$ to the boundary $\partial\Omega$ of an open, bounded set $\Omega \subset \mathbf{R}^n$. In some ways the HJ equation for this function is simpler than that of the optimal control problem described above, and we can more easily explain viscosity solutions and issues of uniqueness, etc, in this context.

The distance function is defined by

$$d(x, \partial\Omega) = \inf_{y \in \partial\Omega} |x - y|. \tag{18}$$

Note that the infimum here is always attained, not necessarily uniquely, since $\partial\Omega$ is compact and $y \mapsto |x - y|$ is continuous; denote by $\pi(x) \subset \partial\Omega$ the set of minimizing y.

We write

$$V(x) = d(x, \partial\Omega) \tag{19}$$

for simplicity, and consider $V(x)$ as a function on the closed set $\overline{\Omega}$. It can be verified that $V(x)$ is a non-negative Lipschitz continuous function. In fact, we shall see that V is the unique continuous viscosity solution of

$$|\nabla V| - 1 = 0 \ \text{ in } \ \Omega \tag{20}$$

satisfying the boundary condition

$$V = 0 \ \text{ on } \ \partial\Omega. \tag{21}$$

Equations (20) and (21) constitute a *Dirichlet problem*.

Example 2.1 $\Omega = (-1, 1) \subset \mathbf{R}^1$. Here, $\partial\Omega = \{-1, 1\}$ and $\overline{\Omega} = [-1, 1]$. Then

$$V(x) = \begin{cases} 1 + x & \text{if } -1 \le x \le 0 \\ 1 - x & \text{if } 0 \le x \le 1 \end{cases}$$

which is Lipschitz continuous, and differentiable except at $x = 0$. At each point $x \neq 0$ V solves the HJ equation (20), and V satisfies the boundary condition (21) ($V(-1) = v(1) = 0$), see Figure 2. Note that $\pi(x) = -1$ for $-1 \le x < 0$, $\pi(x) = 1$ for $0 < x \le 1$, and $\pi(0) = \{-1, 1\}$. The Lipschitz function $V_1(x) = |x| - 1$ also satisfies (20) a.e. and (21); there are many other such functions.

Dynamic programming. The distance function satisfies a simple version of the dynamic programming principle: for any $r > 0$ we have

$$V(x) = \inf_{|x-z|<r} \{|x - z| + V(z)\}. \tag{22}$$

We will use this later to show that V is a viscosity solution of (20), but for now we discuss and derive (22).

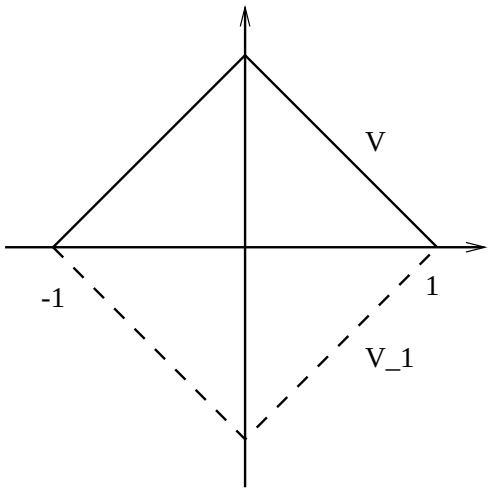

Figure 2: Distance function V and another Lipschitz solution V_1.

Fix $x \in \Omega$ and $r > 0$, and let $|x - z| < r$. Choose $y^*(z) \in \pi(z)$, so that $V(z) = |z - y^*(z)|$. Then

$$\begin{aligned} V(x) &\leq |x - y^*(z)| \\ &\leq |x - z| + |z - y^*(z)| \\ &= |x - z| + V(z). \end{aligned}$$

Since this holds for all $|x - z| < r$ we have

$$V(x) \leq \inf_{|x-z|<r} \{|x - z| + V(z)\}.$$

To see that equality holds, simply take $z = x$. Thus establishes (22). Note that there are many minimizers z^* for the RHS of (22), viz. segments of the lines joining x to points in $\pi(x)$.

2.1.4 Viscosity Solutions

We turn now to the concept of viscosity solution for the HJ equation (1). The terminology comes from the *vanishing viscosity method*, which finds a solution V of (1) as a limit $V^\varepsilon \to V$ of solutions to

$$-\frac{\varepsilon}{2}\Delta V^\varepsilon(x) + F(x, V^\varepsilon(x), \nabla V^\varepsilon(x)) = 0 \tag{23}$$

The Laplacian term $\frac{\varepsilon}{2}\Delta V^\varepsilon = \frac{\varepsilon}{2}\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} V^\varepsilon$ can be used to model fluid viscosity. The definition below is quite independent of this limiting construction, and is closely related to dynamic programming; however, the definition applies also to equations that do not necessarily correspond to optimal control.

A function $V \in C(\Omega)$ is a *viscosity subsolution* of (1) if, for any $\phi \in C^1(\Omega)$ and any local maximum $x_0 \in \Omega$ of $V - \phi$ we have

$$F(x_0, V(x_0), \nabla\phi(x_0)) \leq 0 \tag{24}$$

A function $V \in C(\Omega)$ is a *viscosity supersolution* of (1) if, for any $\phi \in C^1(\Omega)$ and any local minimum $x_0 \in \Omega$ of $V - \phi$ we have

$$F(x_0, V(x_0), \nabla\phi(x_0)) \geq 0 \tag{25}$$

A function $V \in C(\Omega)$ is a *viscosity solution* of (1) if it is both a subsolution and a supersolution.

This definition may at first sight appear strange, though in practice it is often easy to use. Note that derivatives in (24) and (25) appear only on the smooth function ϕ. There are a number of equivalent formulations, and the key point is that the definitions relate sub- or superdifferentials (of functions which need not be differentiable) to inequalities associated with the HJ equation.

The *superdifferential* of a function $V \in C(\Omega)$ is defined by

$$D^+V(x) = \{\lambda \in \mathbf{R}^n \ : \ \limsup_{y \to x,\, y \in \Omega} \frac{V(y) - V(x) - \lambda(y-x)}{|x-y|} \leq 0\} \tag{26}$$

The *subdifferential* of a function $V \in C(\Omega)$ is defined by

$$D^-V(x) = \{\lambda \in \mathbf{R}^n \ : \ \liminf_{y \to x,\, y \in \Omega} \frac{V(y) - V(x) - \lambda(y-x)}{|x-y|} \geq 0\} \tag{27}$$

If $V \in C^1(\Omega)$ then $D^+V(x) = D^-V(x) = \{\nabla V(x)\}$. In general, $\lambda \in D^+V(x)$ iff there exists $\phi \in C^1(\Omega)$ such that $\nabla\phi(x) = \lambda$ and $V - \phi$ has a local maximum at x; and $\lambda \in D^-V(x)$ iff there exists $\phi \in C^1(\Omega)$ such that $\nabla\phi(x) = \lambda$ and $V - \phi$ has a local minimum at x.

Therefore the viscosity definition is equivalently characterized by

$$F(x, V(x), \lambda) \leq 0 \ \ \forall \ \ \lambda \in D^+V(x) \tag{28}$$

and

$$F(x, V(x), \lambda) \geq 0 \ \ \forall \ \ \lambda \in D^-V(x) \tag{29}$$

Example 2.2 Continuing with Example 2.1, we see that

$$\begin{aligned} D^+V(x) &= D^-V(x) = \begin{cases} \{1\} & \text{if } \ -1 < x < 0 \\ \{-1\} & \text{if } \ 0 < x < 1 \end{cases} \\ D^+V(0) &= [-1, 1], \\ D^-V(0) &= \emptyset. \end{aligned} \tag{30}$$

Consequently V is a viscosity solution of (20).

However, the function V_1 is *not* is viscosity solution, since $0 \in D^-V_1(0) = [-1, 1]$, and

$$|0| - 1 \not\geq 0.$$

Properties. Some properties of viscosity solutions:

1. (Consistency.) If $V \in C(\Omega)$ is a viscosity solution of (1), then for any point $x \in \Omega$ at which V is differentiable we have

$$F(x, V(x), \nabla V(x)) = 0.$$

2. If V is locally Lipschitz continuous in Ω, then

$$F(x, V(x), \nabla V(x)) = 0 \ \text{ a.e. in } \ \Omega.$$

3. (Stability.) Let $V^N \in C(\Omega)$ $(N \geq 0)$ be viscosity solutions of

$$F^N(x, V^N(x), \nabla V^N(x)) = 0 \ \text{ in } \ \Omega,$$

and assume $V^N \to V$ locally uniformly in Ω, and $F^N \to F$ locally uniformly in $\Omega \times \mathbf{R} \times \mathbf{R}^n$, as $N \to \infty$. Then $V \in C(\Omega)$ is a viscosity solution of (1).

4. (Monotonic change of variable.) Let $V \in C(\Omega)$ be a viscosity solution of (1) and $\Psi \in C^1(\mathbf{R})$ be such that $\Phi'(t) > 0$. Then $W = \Phi(V)$ is a viscosity solution of

$$F(x, \Psi(W(x)), \Psi'(W(x))\nabla W(x)) = 0 \tag{31}$$

where $\Psi = \Phi^{-1}$.

2.2 Value Functions are Viscosity Solutions

2.2.1 The Distance Function is a Viscosity Solution

We showed in Example 2.2 that in the specific case at hand the distance function is a viscosity solution. Let's now consider the general case. We use the dynamic programming principle (22) to illustrate a general methodology.

Subsolution property. Let $\phi \in C^1(\Omega)$ and suppose that $V - \phi$ attains a local maximum at $x_0 \in \Omega$; so there exists $r > 0$ such that the ball $B(x_0, r) \subset \Omega$ and

$$V(x) - \phi(x) \leq V(x_0) - \phi(x_0) \ \ \forall \ x \in B(x_0, r). \tag{32}$$

We want to show that

$$|\nabla \phi(x_0)| - 1 \leq 0. \tag{33}$$

Let $h \in \mathbf{R}^n$, and set $x = x_0 + th$. Then for $t > 0$ sufficiently small $x \in B(x_0, r)$, and so from (32),

$$-(\phi(x_0 + th) - \phi(x_0)) \leq -(V(x_0 + th) - V(x_0)) \tag{34}$$

Now from the dynamic programming principle (22) we have

$$V(x_0) \leq t|h| + V(x_0 + th) \tag{35}$$

Combining (34) and (35) we find that

$$-(\phi(x_0 + th) - \phi(x_0)) \leq t|h|, \tag{36}$$

and so

$$-(\frac{\phi(x_0 + th) - \phi(x_0)}{t|h|}) - 1 \leq 0 \tag{37}$$

Send $t \downarrow 0$ to obtain

$$-\frac{1}{|h|}\nabla\phi(x_0)h - 1 \leq 0. \tag{38}$$

Since $h \in \mathbf{R}^n$ is arbitrary we obtain (32) as required. This proves that V is a viscosity subsolution.

Supersolution property. Let $\phi \in C^1(\Omega)$ and suppose that $V - \phi$ attains a local minimum at $x_0 \in \Omega$; so there exists $r > 0$ such that the ball $B(x_0, r) \subset \Omega$ and

$$V(x) - \phi(x) \geq V(x_0) - \phi(x_0) \ \ \forall\ x \in B(x_0, r). \tag{39}$$

We want to show that

$$|\nabla\phi(x_0)| - 1 \geq 0. \tag{40}$$

Suppose that (40) is false, so that

$$|\nabla\phi(x_0)| - 1 \leq -\alpha < 0 \tag{41}$$

for some $1 > \alpha > 0$. By making $r > 0$ smaller if necessary, we may assume

$$|\nabla\phi(x)| - 1 \leq -\alpha/2 < 0 \ \ \forall \ \ x \in B(x_0, r). \tag{42}$$

By the fundamental theorem of calculus, we have

$$\phi(x) = \phi(x_0) + \int_0^1 \nabla\phi(\gamma x + (1 - \gamma)x_0)(x - x_0)d\gamma \tag{43}$$

Now from the dynamic programming relation (22), pick $z^* \in B(x_0, r)$, $z^* \neq x_0$, such that

$$V(x_0) = |x_0 - z^*| + V(z^*). \tag{44}$$

Using this and (39) we have

$$-(\phi(z^*) - \phi(x_0)) \geq |z^* - x_0|. \tag{45}$$

However, from (42) and (43) we must have

$$-(\phi(z^*) - \phi(x_0)) \leq (1 - \alpha/2)|z^* - x_0|. \tag{46}$$

Inequalities (45) and (46) are in contradiction, so in fact (40) holds. This proves that V is a supersolution.

It can be seen here that the dynamic programming principle provided the key inequalities to derive the sub- and supersolution relations.

2.2.2 The Optimal Control Value Function is a Viscosity Solution

Dynamic programming. The *dynamic programming principle* states that for every $r \in [t, t_1]$,

$$V(t,x) \;=\; \inf_{u(\cdot)\in\mathcal{U}_{t,r}} \left[\int_t^r L(x(s),u(s))\,ds + V(r,x(r)) \right]. \tag{6}$$

We now prove this (by a standard technique in optimal control).

Fix $r \in [t,t_1]$, and $u(\cdot) \in \mathcal{U}_{t,t_1}$. Let $x(\cdot)$ denote the corresponding trajectory with initial state $x(t) = x$, and consider $(r, x(r))$. Let $\varepsilon > 0$ and choose $u_1(\cdot) \in \mathcal{U}_{r,t_1}$, with trajectory $x_1(\cdot)$ on $[r,t_1]$ with $x_1(r) = x(r)$ be such that

$$V(r,x(r)) \geq J(r,x(r);u_1(\cdot)) - \varepsilon. \tag{47}$$

Define

$$u_2(s) = \begin{cases} u(s) & t \leq s < r \\ u_1(s) & r \leq s \leq t_1 \end{cases} \tag{48}$$

with trajectory $x_2(\cdot)$, $x_2(t) = x$. Now $x_2(s) = x(s)$, $s \in [t,r]$, and $x_2(s) = x_1(s)$, $s \in [r,t_1]$. Next,

$$\begin{aligned} V(r,x) &\leq J(t,x;u_2(\cdot)) \\ &= \textstyle\int_t^{t_1} L(x(s),u(s))\,ds + \psi(x(t_1)) \\ &= \textstyle\int_r^t L(x(s),u(s))\,ds + \int_t^r L(x_1(s),u_1(s))\,ds + \psi(x(t_1)) \\ &= \textstyle\int_r^t L(x(s),u(s))\,ds + V(r,x(r)) + \varepsilon \end{aligned} \tag{49}$$

using (47). Therefore

$$V(t,x) \leq \inf_{u(\cdot)\in\mathcal{U}_{t,r}} \left[\int_t^r L(x(s),u(s))\,ds + V(r,x(r)) \right] + \varepsilon. \tag{50}$$

Since $\varepsilon > 0$ was arbitrary, we have

$$V(t,x) \leq \inf_{u(\cdot)\in\mathcal{U}_{t,r}} \left[\int_t^r L(x(s),u(s))\,ds + V(r,x(r)) \right]. \tag{51}$$

This proves one half of (6).

For the second half of (6), let $u(\cdot) \in \mathcal{U}_{t,t_1}$, and let $x(\cdot)$ be the corresponding trajectory with $x(t) = x$. Then

$$\begin{aligned} J(t,x;u(\cdot)) &= \textstyle\int_t^{t_1} L(x(s),u(s))\,ds + \psi(x(t_1)) \\ &= \textstyle\int_t^r L(x(s),u(s))\,ds + \int_r^{t_1} L(x(s),u(s))\,ds + \psi(x(t_1)) \\ &\geq \textstyle\int_t^r L(x(s),u(s))\,ds + V(r,x(r)). \end{aligned} \tag{52}$$

Now minimizing, we obtain

$$\begin{aligned} V(t,x) &= \inf_{u(\cdot)\in\mathcal{U}_{t,t_1}} J(t,x;u(\cdot)) \\ &\geq \inf_{u(\cdot)\in\mathcal{U}_{t,t_1}} \{ \textstyle\int_t^r L(x(s),u(s))\,ds + V(r,x(r)) \} \end{aligned} \tag{53}$$

which is the desired second half of (6). This establishes the dynamic programming principle (6).

Regularity. By regularity we mean the degree of continuity or differentiability; i.e. of smoothness. The regularity of value functions is determined by both the regularity of the data defining it (e.g. f, L, ψ), and on the nature of the optimization problem. In many applications, the value function can readily be shown to be continuous, even Lipschitz, but not C^1 in general. The finite horizon value function $V(t,x)$ defined by (5) can be shown to be bounded and Lipschitz continuous under the following (rather strong) assumptions on the problem data: f, L, ψ are bounded with bounded first order derivatives. We shall assume this.

It should be noted that in general it can happen that value functions fail to be continuous. In fact, the viscosity theory is capable of dealing with semicontinuous or even only locally bounded functions.

Viscosity solution. Let us re-write the HJ equation (7) as follows:

$$-\frac{\partial}{\partial t}V(t,x) + H(x, \nabla_x V(t,x)) \;=\; 0 \;\text{ in } (t_0,t_1)\times \mathbf{R}^n, \tag{7}'$$

with a new definition of the Hamiltonian

$$H(x,\lambda) \;=\; \sup_{v\in\mathbf{R}^m} \{-\lambda\cdot f(x,v) - L(x,v)\}\,. \tag{9}'$$

The sign convention used in (7)' relates to the maximum principle in PDE, and is compatible with the convention used for the general HJ equation (1). Note that the Hamiltonian is now *convex* in λ.

A function $\tilde{V}\in C([t_0,t_1]\times\mathbf{R}^n)$ is a viscosity subsolution (resp. supersolution) of (7)' if for all $\phi\in C^1((t_0,t_1)\times\mathbf{R}^n)$,

$$-\frac{\partial}{\partial t}\phi(s_0,x_0) + H(x_0, \nabla\phi(s_0,x_0)) \;\le\; 0 \;\;(\text{resp. } \ge 0) \tag{54}$$

at every point (s_0,x_0) where $\tilde{V}-\phi$ attains a local maximum (resp. minimum). $\tilde{V}$ is a viscosity solution if it is both a subsolution and a supersolution.

We now show that the value function V defined by (5) is a viscosity solution of (7)'.

Subsolution property. Let $\phi\in C^1((t_0,t_1)\times\mathbf{R}^n)$ and suppose that $V-\phi$ attains a local maximum at (s_0,x_0); so there exists $r>0$ such that

$$V(t,x)-\phi(t,x) \le V(s_0,x_0)-\phi(s_0,x_0) \;\;\forall\; |x-x_0|<r, \;\; |t-s_0|<r. \tag{55}$$

Fix $u(t)=u\in U$ for all t (constant control) and let $\xi(\cdot)$ denote the corresponding state trajectory with $\xi(s_0)=x_0$. By standard ODE estimates, we have

$$|\xi(s_0+h)-x_0|<r \tag{56}$$

for all $0\le h\le h_0$ (some $h_0>0$) - since U and f are bounded. Then by (55)

$$V(s_0+h,\xi(s_0+h))-\phi(s_0+h,\xi(s_0+h)) \le V(s_0,x_0)-\phi(s_0,x_0) \tag{57}$$

for all $0 < h \leq h_0$. Now from the dynamic programming principle (6) we have

$$V(s_0, x_0) \leq \int_{s_0}^{s_0+h} L(\xi(s), u(s))\, ds + V(s_0 + h, \xi(s_0 + h)) \tag{58}$$

which with (57) implies

$$-(\frac{\phi(s_0 + h, \xi(s_0 + h)) - \phi(s_0, x_0))}{h}) - \frac{1}{h}\int_{s_0}^{s_0+h} L(x(s), u(s))\, ds \leq 0. \tag{59}$$

Send $h \to 0$ to obtain

$$-\frac{\partial}{\partial t}\phi(s_0, x_0) - \nabla\phi(s_0, x_0) f(x_0, u) - L(x_0, u) \leq 0. \tag{60}$$

Now maximize over u to obtain

$$-\frac{\partial}{\partial t}\phi(s_0, x_0) + \sup_{u \in U}\{-\nabla\phi(s_0, x_0) f(x_0, u) - L(x_0, u)\} \leq 0. \tag{61}$$

This proves that V is a viscosity subsolution.

Supersolution property. Let $\phi \in C^1((t_0, t_1) \times \mathbf{R}^n)$ and suppose that $V - \phi$ attains a local minimum at (s_0, x_0); so there exists $r > 0$ such that

$$V(t, x) - \phi(t, x) \geq V(s_0, x_0) - \phi(s_0, x_0) \;\; \forall\; |x - x_0| < r, \quad |t - s_0| < r. \tag{62}$$

Again by ODE estimates, there exists $h_0 > 0$ such that

$$|\xi(s_0 + h) - x_0| < r \tag{63}$$

for all $0 \leq h \leq h_0$ and all $u(\cdot) \in \mathcal{U}_{s_0,t_1}$, where $\xi(\cdot)$ denotes the corresponding state trajectory with $\xi(s_0) = x_0$.

Assume the supersolution property is false, i.e. there exists $\alpha > 0$ such that

$$-\frac{\partial}{\partial t}\phi(s_0, x_0) + \sup_{u \in U}\{-\nabla\phi(s_0, x_0) f(x_0, u) - L(x_0, u)\} \leq -3\alpha h < 0, \tag{64}$$

where $0 < h < h_0$. Now (64) implies

$$-\frac{\partial}{\partial t}\phi(s, \xi(s)) - \nabla\phi(s, \xi(s)) f(\xi(s), u(s)) - L(\xi(s), u(s)) \leq -2\alpha h < 0, \tag{65}$$

for all $s \in [s_0, s_0 + h]$ and all $u(\cdot) \in \mathcal{U}_{s_0,t_1}$, for $h > 0$ sufficiently small.

By the dynamic programming formula (6), there exists $u_0(\cdot) \in \mathcal{U}_{s_0,t_1}$ such that

$$V(s_0, x_0) \;\geq\; \int_{s_0}^{s_0+h} L(\xi_0(s), u_0(s))\, ds + V(s_0 + h, \xi_0(s_0 + h)) - \alpha h \tag{66}$$

where $\xi_0(\cdot)$ denotes the corresponding trajectory with $\xi(s_0) = x_0$, and combining this with (62) we have

$$-(\frac{\phi(s_0+h,\xi_0(s_0+h)) - \phi(s_0,x_0))}{h}) - \frac{1}{h}\int_{s_0}^{s_0+h} L(\xi_0(s),u_0(s))\,ds \geq -\alpha. \tag{67}$$

However, integration of (65) implies

$$-(\frac{\phi(s_0+h,\xi_0(s_0+h)) - \phi(s_0,x_0))}{h}) - \frac{1}{h}\int_{s_0}^{s_0+h} L(\xi_0(s),u_0(s))\,ds \leq -2\alpha. \tag{68}$$

which contradicts (67) since $\alpha > 0$. This proves the supersolution property.

2.3 Comparison and Uniqueness

The most important features of the theory of viscosity solutions are the powerful comparison and uniqueness theorems. Comparison theorems assert that inequalities holding on the boundary and/or terminal time also hold in the entire domain. Uniqueness follows from this. Such results are important, since they guarantee unique characterization of viscosity solutions, and ensure that convergent approximations converge to the correct limit. In the context of optimal control problems, value functions are the unique viscosity solutions.

In this section we give a detailed proof of the comparison and uniqueness results for a class of Dirichlet problems, and apply this to equation (20) for the distance function. We also present without proof results for Cauchy problems of the type (7), (8).

2.3.1 Dirichlet Problem

Here we follow [3, Chapter II] and consider the HJ equation

$$V(x) + H(x,\nabla V(x)) = 0 \ \text{ in } \ \Omega, \tag{69}$$

a special case of (1).

To help get a feel for the ideas, suppose $V_1, V_2 \in C(\overline{\Omega}) \cap C^1(\Omega)$ (i.e. are smooth) satisfy

$$\begin{array}{ll} V_1(x) + H(x,\nabla V_1(x)) \leq 0 & \text{(subsolution)} \\ V_2(x) + H(x,\nabla V_2(x)) \geq 0 & \text{(supersolution)} \end{array} \tag{70}$$

in Ω and

$$V_1 \leq V_2 \ \text{ on } \ \partial\Omega \quad \text{(boundary condition)}. \tag{71}$$

Let $x_0 \in \overline{\Omega}$ be a maximum point of $V_1 - V_2$. Now if $x_0 \in \Omega$ (interior, not on boundary) then $\nabla V_1(x_0) = \nabla V_2(x_0)$ and subtracting the first second line of (70) from the first gives

$$V_1(x_0) - V_2(x_0) \leq 0$$

which implies

$$V_1(x) - V_2(x) \leq V_1(x_0) - V_2(x_0) \leq 0 \ \ \forall \ \ x \in \overline{\Omega}. \tag{72}$$

If it happened that $x_0 \in \partial\Omega$, then using (71)

$$V_1(x) - V_2(x) \leq V_1(x_0) - V_2(x_0) \leq 0 \ \ \forall \ \ x \in \overline{\Omega}. \tag{73}$$

Therefore $V_1 \leq V_2$ in $\overline{\Omega}$, a *comparison* result.

Comparison implies uniqueness for the Dirichlet problem

$$\begin{aligned} V(x) + H(x, \nabla V(x)) &= 0 \ \text{ in } \ \Omega, \\ V &= \psi \ \text{ on } \ \partial\Omega. \end{aligned} \tag{74}$$

To see this, suppose V_1 and V_2 are two solutions. Now $V_1 = V_2 = \psi$ on $\partial\Omega$. Then by the comparison result, we get $V_1 \leq V_2$ in $\overline{\Omega}$. Similarly, interchanging V_1, V_2 we again apply comparison to obtain $V_2 \leq V_1$ in $\overline{\Omega}$. Hence $V_1 = V_2$ in $\overline{\Omega}$.

This illustrates the role of sub- and supersolutions and boundary conditions in the comparison and uniqueness theory. We now give a precise theorem and proof ([3, Theorem II.3.1]). This result does not use convexity or any connection to optimal control, and applies generally.

Theorem 2.3 *Let Ω be a bounded open subset of $\mathbf{R}^n$. Assume $V_1, V_2 \in C(\overline{\Omega})$ are, respectively, viscosity sub- and supersolution of (69), and satisfy the inequality (71) on the boundary. Assume that H satisfies*

$$|H(x, \lambda) - H(y, \lambda)| \leq \omega_1(|x - y|(1 + |\lambda|)), \tag{75}$$

for $x, y \in \Omega$, $\lambda \in \mathbf{R}^n$, where $\omega_1 : [0, +\infty) \to [0, +\infty)$ is continuous, nondecreasing with $\omega_1(0) = 0$ (ω_1 is called a modulus*). Then*

$$V_1 \leq V_2 \ \text{ in } \ \overline{\Omega}. \tag{76}$$

PROOF. For $\varepsilon > 0$ define the continuous function on $\overline{\Omega} \times \overline{\Omega}$ by

$$\Phi_\varepsilon(x, y) = V_1(x) - V_2(y) - \frac{|x - y|^2}{2\varepsilon}.$$

Let $(x_\varepsilon, y_\varepsilon) \in \overline{\Omega} \times \overline{\Omega}$ be a maximum point for Φ_ε over $\overline{\Omega} \times \overline{\Omega}$. Then

$$\max_{x \in \overline{\Omega}} (V_1 - V_2)(x) \leq \max_{x \in \overline{\Omega}} \Phi_\varepsilon(x, x) \leq \max_{x, y \in \overline{\Omega}} \Phi_\varepsilon(x, y) = \Phi_\varepsilon(x_\varepsilon, y_\varepsilon). \tag{77}$$

We claim that

$$\limsup_{\varepsilon \to 0} \Phi_\varepsilon(x_\varepsilon, y_\varepsilon) = 0, \tag{78}$$

which together with (77) proves the theorem.

Let us prove (78). Now the fact that

$$\Phi_\varepsilon(x_\varepsilon, x_\varepsilon) \leq \Phi_\varepsilon(x_\varepsilon, y_\varepsilon)$$

implies

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \leq V_2(x_\varepsilon) - V_2(y_\varepsilon) \leq C \tag{79}$$

for suitable $C > 0$ (recall V_2 is bounded on $\overline{\Omega}$), and so

$$|x_\varepsilon - y_\varepsilon| \leq (C\varepsilon)^{1/2}. \tag{80}$$

Therefore

$$|x_\varepsilon - y_\varepsilon| \to 0 \ \text{ as } \ \varepsilon \to 0, \tag{81}$$

and by continuity, $V_2(x_\varepsilon) - V_2(y_\varepsilon) \to 0$ as $\varepsilon \to 0$; hence (79) gives

$$\frac{|x_\varepsilon - y_\varepsilon|^2}{2\varepsilon} \to 0 \ \text{ as } \ \varepsilon \to 0. \tag{82}$$

We now need to consider where the points $x_\varepsilon, y_\varepsilon$ lie.

Case (i). Suppose $x_\varepsilon, y_\varepsilon \in \Omega$ (both interior points), for all sufficiently small $\varepsilon > 0$. Let

$$\phi_1(y) = V_1(x_\varepsilon) - \frac{|x_\varepsilon - y|^2}{2\varepsilon}, \quad \phi_2(x) = V_2(y_\varepsilon) + \frac{|x - y_\varepsilon|^2}{2\varepsilon}, \tag{83}$$

Now $\phi_1, \phi_2 \in C^1(\Omega)$, x_ε is a local maximum for $V_1 - \phi_2$, and y_ε is a local minimum for $V_2 - \phi_1$. Also,

$$\nabla\phi_1(y_\varepsilon) = \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} = \nabla\phi_2(x_\varepsilon). \tag{84}$$

The viscosity sub- and supersolution definition implies

$$\begin{array}{l} V_1(x_\varepsilon) + H(x_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon}) \leq 0, \\ V_2(y_\varepsilon) + H(y_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon}) \geq 0. \end{array} \tag{85}$$

Subtracting we have

$$V_1(x_\varepsilon) - V_2(y_\varepsilon) + H(x_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon}) - H(y_\varepsilon, \frac{x_\varepsilon - y_\varepsilon}{\varepsilon}) \leq 0 \tag{86}$$

and using the assumption on H

$$V_1(x_\varepsilon) - V_2(y_\varepsilon) \leq \omega_1(|x_\varepsilon - y_\varepsilon|(1 + \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon})). \tag{87}$$

This implies

$$\Phi_\varepsilon(x_\varepsilon, y_\varepsilon) \leq \omega_1(|x_\varepsilon - y_\varepsilon|(1 + \frac{|x_\varepsilon - y_\varepsilon|}{\varepsilon})), \tag{88}$$

and hence (78) follows using (81) and (82).

Case (ii). Now suppose there exists a subsequence $\varepsilon_i \to 0$ as $i \to \infty$ such that $x_{\varepsilon_i} \in \partial\Omega$ or $y_{\varepsilon_i} \in \partial\Omega$. If $x_{\varepsilon_i} \in \partial\Omega$,

$$V_1(x_{\varepsilon_i}) - V_2(y_{\varepsilon_i}) \leq V_2(x_{\varepsilon_i}) - V_2(y_{\varepsilon_i}) \to 0 \tag{89}$$

as $i \to \infty$, or if $y_{\varepsilon_i} \in \partial\Omega$,

$$V_1(x_{\varepsilon_i}) - V_2(y_{\varepsilon_i}) \leq V_1(x_{\varepsilon_i}) - V_1(y_{\varepsilon_i}) \to 0 \tag{90}$$

as $i \to \infty$; hence (78).

This completes the proof.

The distance function is the unique viscosity solution of (20), (21). At first sight Theorem 2.3 does not apply to (20), (21). This is because equation (20) does not have the additive V-term that (69) has, and this term was used in an essential way in the proof of Theorem 2.3. In fact, in general viscosity solutions to equations of the form

$$H(x, \nabla V) = 0 \tag{91}$$

may not be unique! For instance, in the context of the Bounded Real Lemma both the available storage and required supply are viscosity solutions of equations of the type (91). It turns out that comparison/uniqueness for HJ equation (20) for the distance function can be proved, either directly using additional hypothesis (such as convexity [3, Theorem II.5.9]), or via a transformation as we now show.

We use the Kruskov transformation, a useful trick. Define

$$W = \Phi(V) \stackrel{\triangle}{=} 1 - e^{-V}, \tag{92}$$

where V is the distance function (19). Then W is a viscosity solution of

$$\begin{aligned} W(x) + |\nabla V(x)| - 1 &= 0 \ \text{ in } \ \Omega, \\ W &= 0 \ \text{ on } \ \partial\Omega, \end{aligned} \tag{93}$$

by the general properties of viscosity solutions mentioned above. By Theorem 2.3, we see that W is the unique viscosity solution of (93), and hence

$$V = \Psi(W) \stackrel{\triangle}{=} \Phi^{-1}(W) = -\log(1 - W) \tag{94}$$

is the unique viscosity solution of (20), (21). Comparison also follows in the same way.

2.3.2 Cauchy Problem

In this section we simply state without proof an example of a comparison/uniqueness result, [3, Theorem III.3.15]. There are many results like this available, with various kinds of structural assumptions (e.g. (95), (96)) which must be checked in order to apply them.

Theorem 2.4 *Assume H is continuous and satisfies*

$$|H(x,\lambda_1) - H(x,\lambda_2)| \leq K(1+|x|)|\lambda_1 - \lambda_2| \tag{95}$$

for all $x, \lambda_1, \lambda_2 \in \mathbf{R}^n$, and

$$|H(x_1,\lambda) - H(x_2,\lambda)| \leq \omega(|x_1 - x_2|, R) + \omega(|x_1 - x_2||\lambda|, R) \tag{96}$$

for all $\lambda \in \mathbf{R}^n$, $x_1, x_2 \in B(0,R)$, $R > 0$, where ω is a modulus (depending on R). Let $V_1, V_2 \in C([t_0,t_1] \times \mathbf{R}^n)$ be, respectively, viscosity sub- and supersolution of (7)' satisfying

$$V_1(t_1,x) \leq V_2(t_1,x) \quad \forall \quad x \in \mathbf{R}^n. \tag{97}$$

Then

$$V_1(t,x) \leq V_2(t,x) \quad \forall \quad (t,x) \in [t_0,t_1] \times \mathbf{R}^n. \tag{98}$$

3 Stochastic Control

References on stochastic control and probability theory include [38], [17], [10], [22], [24], [2], [9], [39], [1].

3.1 Some Probability Theory

While probability theory, especially quantum probability, will be covered by other speakers at this workshop, we present in this subsection some basic definitions and ideas.

3.1.1 Basic Definitions

Classical probability theory considers events F, subsets of a sample space Ω, and assigns a numerical value $0 \leq \mathbf{P}(F) \leq 1$ to each event F indicating the probability of occurrence of F. The collection of all allowed events is denoted $\mathcal{F}$. The basic construct of classical probability is the triple $(\Omega, \mathcal{F}, \mathbf{P})$, called a *classical probability space*.

To facilitate an adequate framework for integration (expectation), convergence, etc, there are a number of technical requirements placed on probability spaces. While the set Ω of outcomes can be arbitrary (e.g. colors of balls in an urn, the set of real numbers, etc), the collection of allowed events $\mathcal{F}$ is required to be a σ-algebra. A σ-algebra $\mathcal{F}$ contains the empty set ($\emptyset \in \mathcal{F}$), is closed under complements ($F \in \mathcal{F}$ implies $F^c = \{\omega \in \Omega \, : \, \omega \notin F\} \in \mathcal{F}$), and is closed under countable unions ($\{F_i\}_{i=1}^{\infty} \subset \mathcal{F}$ implies $\bigcup_{i=1}^{\infty} F_i \in \mathcal{F}$). A pair $(\Omega, \mathcal{F})$ is called a *measurable space* (on which one or more probability measures may be defined).

A *probability measure* is a function $\mathbf{P} : \mathcal{F} \to [0, 1]$ such that (i) $0 \leq \mathbf{P}(F) \leq 1$ for all $F \in \mathcal{F}$, (ii) $\mathbf{P}(\Omega) = 1$, and (iii) if $F_1, F_2, \ldots$ is a disjoint sequence of events in $\mathcal{F}$, then $\mathbf{P}(\bigcup_{i=1}^{\infty} F_i) = \sum_{i=1}^{\infty} \mathbf{P}(F_i)$.

In many cases, the set of outcomes Ω will be a topological space, i.e. a set Ω together with a collection τ of subsets (the open sets), called a *topology*. A topology τ contains the empty set, and is closed under arbitrary unions and intersections. If Ω is discrete, then the set of all subsets defines the standard topology. If $\Omega = \mathbf{R}$ or $\mathbf{C}$ (real or complex numbers), the standard topology can be defined by considering all open intervals or discs (and their arbitrary unions and intersections). Given a topological space (Ω, τ), the Borel σ-algebra $\mathcal{B}(\Omega)$ is the σ-algebra generated by the open sets (it is the smallest σ-algebra containing all open sets). Events in a Borel σ-algebra are called *Borel sets*. Often, when the topology or σ-algebra is clear, explicit mention of them is dropped from the notation.

Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces. A *random variable* or *measurable function* is a function $X : \Omega_1 \to \Omega_2$ such that $X^{-1}(F_2) = \{\omega_1 \in \Omega_1 \, : \, X(\omega_1) \in F_2\} \in \mathcal{F}_1$ for all $F_2 \in \mathcal{F}_2$. In particular, a *real-valued random variable* X defined on $(\Omega, \mathcal{F})$ is a function $X : \Omega \to \mathbf{R}$ such that $X^{-1}(B) = \{\omega \in \Omega \, : \, X(\omega) \in B\} \in \mathcal{F}$ for any Borel set $B \subset \mathbf{R}$. Similarly, we can consider complex-valued random variables. If $\mathbf{P}$ is a probability measure on $(\Omega, \mathcal{F})$, the probability distribution induced by X is

$$\mathbf{P}_X(B) = \mathbf{P}(X^{-1}(B)),$$

so that $(\mathbf{R}, \mathcal{B}(\mathbf{R}), \mathbf{P}_X)$ is a probability space. If X has a density $p_X(x)$, expectations of functions $f(X)$ (e.g. moments) can be calculated via

$$\mathbf{E}[f(X)] = \int_{\mathbf{R}} f(x)\mathbf{P}_X(dx) = \int_{-\infty}^{\infty} f(x)p_X(x)dx. \tag{99}$$

Note that the cumulative distribution function is given by $F_X(x) = \mathbf{P}_X((-\infty, x])$, and the density $p_X(x) = dF_X(x)/dx$ (when it exists).

Let Y be a random variable, and $\mathcal{Y} = \sigma(Y)$ be the σ-algebra generated by Y; i.e. the smallest σ-algebra with respect to which Y is measurable. In general $\mathcal{Y} \subset \mathcal{F}$. If Z is a random variable that is also $\mathcal{Y}$-measurable, then there is a function g_Z such that $Z(\omega) = g_Z(Y(\omega))$. Thus, Z is a function of the values of Y.

For $0 < p < \infty$, the set $L^p(\Omega, \mathcal{F}, \mathbf{P})$ is the vector space of complex-valued random variables X such that $\mathbf{E}[|X|^p]$ is finite. It is a Banach space with respect to the norm $\| X \|_p = \mathbf{E}[|X|^p]^{1/p}$. The case $p = 2$ is of special interest, since $L^2(\Omega, \mathcal{F}, \mathbf{P})$ is a Hilbert space with inner product

$$\langle X, Y \rangle = \mathbf{E}[X^* Y].$$

For $p = \infty$, the space of essentially bounded random variables $L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ is a Banach space with norm $\| X \|_\infty = \text{ess.sup}_\omega |X(\omega)|$.

Example 3.1 Consider the classical probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where $\Omega = \{1, 2\}$, $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2\}\}$, and $\mathbf{P} = (p_1, p_2)$. A random variable X has the form $X = (x_1, x_2)$, where $x_1, x_2 \in \mathbf{C}$, and in this example, the spaces $L^\infty(\Omega, \mathcal{F}, \mathbf{P})$ (equal to $L^2(\Omega, \mathcal{F}, \mathbf{P})$ as a set) consist of all such random variables. The expected value of X is given by

$$\mathbf{E}[X] = \int_\Omega X(\omega)\mathbf{P}(d\omega) = x_1 p_1 + x_2 p_2. \tag{100}$$

□

3.1.2 Conditional Expectations

Let $\mathcal{Y} \subset \mathcal{F}$ be a sub-σ-algebra. The *conditional expectation* $\mathbf{E}[f(X)|\mathcal{Y}]$ of $f(X)$ given $\mathcal{Y}$ is the unique $\mathcal{Y}$-measurable function such that

$$\mathbf{E}[f(X)I_F] = \mathbf{E}[\mathbf{E}f(X)|\mathcal{Y}]I_F] \text{ for all } F \in \mathcal{Y}. \tag{101}$$

Here, I_F is the *indicator function* of the set F defined by $I_F(\omega) = 1$ if $\omega \in F$, and $I_F(\omega) = 0$ otherwise.

This definition may seem abstract, and so we attempt to clarify it by describing what happens in the language of elementary probability and give examples.

Recall from elementary probability the notion of conditional density $p_{X|Y}(x|y)$ of a random variable X given another random variable Y. It is given by

$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x, y)}{p_Y(y)} \tag{102}$$

where $p_{X,Y}(x,y)$ is the joint density of X and Y and $p_Y(y)$ is the density of Y. Using this conditional density, conditional expectations can be computed by

$$\mathbf{E}[f(X)|Y](y) = \int_{-\infty}^{\infty} f(x)p_{X|Y}(x|y)dx. \tag{103}$$

Here, we have emphasized the fact that the conditional expectation is a function of the values y of the random variable Y. If $\mathcal{Y}$ is the σ-algebra generated by Y, then we write $\mathbf{E}[f(X)|\mathcal{Y}] = \mathbf{E}[f(X)|Y]$. Property (101) can be checked as follows:

$$\begin{aligned}
\mathbf{E}[\mathbf{E}f([X)|\mathcal{Y}]I_F] &= \int_F \mathbf{E}f([X)|Y](y)p_Y(y)dy \\
&= \int_F \int_{-\infty}^{\infty} f(x)p_{X|Y}(x|y)p_Y(y)dxdy \\
&= \int_F \int_{-\infty}^{\infty} f(x)p_{X,Y}(x,y)dxdy \\
&= \int_F f(x)p_X(x)dx \\
&= \mathbf{E}[f(X)I_F].
\end{aligned}$$

Example 3.2 Let X and W be independent random variables with densities $p_X(x)$ and $p_W(w)$, respectively. Let Y be a random variable defined by

$$Y = X + W. \tag{104}$$

The joint density of X and Y is $p_{X,Y}(x,y) = p_X(x)p_W(y-x)$, and this defines a probability measure $\mathbf{P}(dx) = p_{X,Y}(x,y)dx$ on the Borel subsets of $\Omega = \mathbf{R}\times\mathbf{R}$. The conditional density is given explicitly by

$$p_{X|Y}(x|y) = \frac{p_X(x)p_W(y-x)}{\displaystyle\int p_X(x)p_W(y-x)dx},$$

which defines a conditional distribution $\pi(y)(dx) = p_{X|Y}(x|y)dx$.

In the absence of measurement information, expectations of functions $f(X)$ are evaluated by

$$\langle \mathbf{P}, f\rangle = \mathbf{E}[f] = \int f(x)\mathbf{P}(dx),$$

in accordance with (99). Measurement of Y provides information about X, allowing us to revise expectations using a conditional probability measure. Suppose we know that the values of Y occurred in a set F ($F \in \mathcal{Y} = \sigma(Y)$). The revised or conditional probability measure is

$$\pi(F)(dx) = \frac{\displaystyle\int I_F(y)p(x,y)dy}{p(F)}dx, \tag{105}$$

where $p(F) = \text{Prob}(Y \in F)$. Then the conditional average of $f(X)$ given F is

$$\begin{aligned} \langle \pi(F), f \rangle &= \int f(x)\pi(F)(dx) \\ &= \frac{\int\int_F p_{X,Y}(x,y)f(x)dydx}{p(F)} \\ &= \frac{\int_F \langle \pi(y), f \rangle dy}{p(F)} \end{aligned} \tag{106}$$

Note that the last equality uses the invariance property (101). If A is a possible set of values of X (i.e. $X \in \sigma(X)$), with $f = I_A$, then (106) reduces to the familiar elementary formula $\text{Prob}(X \in A | Y \in F) = \text{Prob}(X \in A, Y \in F)/\text{Prob}(Y \in F)$. □

Example 3.3 Consider the (finite) discrete set $\Omega = \{1, 2, 3\}$. Let $\mathcal{F}$ be the standard discrete σ-algebra, namely all subsets $\mathcal{F} = \{\emptyset, \Omega, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}$, and let $\mathbf{P}$ be the uniform distribution, $\mathbf{P} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. Let $X = (2, 4, 9)$, i.e. $X(1) = 2$, $X(2) = 4$, $X(3) = 9$, and let $Y = (1, 1, 4)$. Then $\mathcal{Y} = \sigma(Y) = \{\emptyset, \Omega, \{1, 2\}, \{3\}\}$. Using the property (101) with $F = \{1, 2\}$ we find that

$$\mathbf{E}[X|\mathcal{Y}](\omega) = 3 \quad \text{for } \omega \in \{1, 2\}$$

while $F = \{3\}$ gives

$$\mathbf{E}[X|\mathcal{Y}](\omega) = 9 \quad \text{for } \omega \in \{3\}.$$

The random variable X has been averaged over the atoms $\{1, 2\}$, $\{3\}$ of $\mathcal{Y}$. The random variables Y and $\mathbf{E}[X|\mathcal{Y}]$ are constant on these atoms ($\mathcal{Y}$ measurability). The conditional expectation $\mathbf{E}[X|\mathcal{Y}]$ can be viewed as a function of the values $y \in \{1, 4\}$ of Y. Indeed, let $g(1) = 3$ and $g(4) = 9$. Then

$$\mathbf{E}[X|\mathcal{Y}](\omega) = g(Y(\omega)).$$

We write simply $\mathbf{E}[X|\mathcal{Y}](y)$ for $g(y)$ (a slight abuse of notation). □

3.1.3 Stochastic Processes

Heuristics. A stochastic process is a random function of time. e.g. Noise, Figure 3.

Definition 3.4 *A **stochastic process** $\{X_n\}_{n=0}^{\infty}$ is a sequence of random variables X_n defined on $(\Omega, \mathcal{F}, P)$. For each $\omega \in \Omega$,*

$$X(\omega) = \{X_0(\omega), X_1(\omega), \ldots\}$$

denotes a sample path.

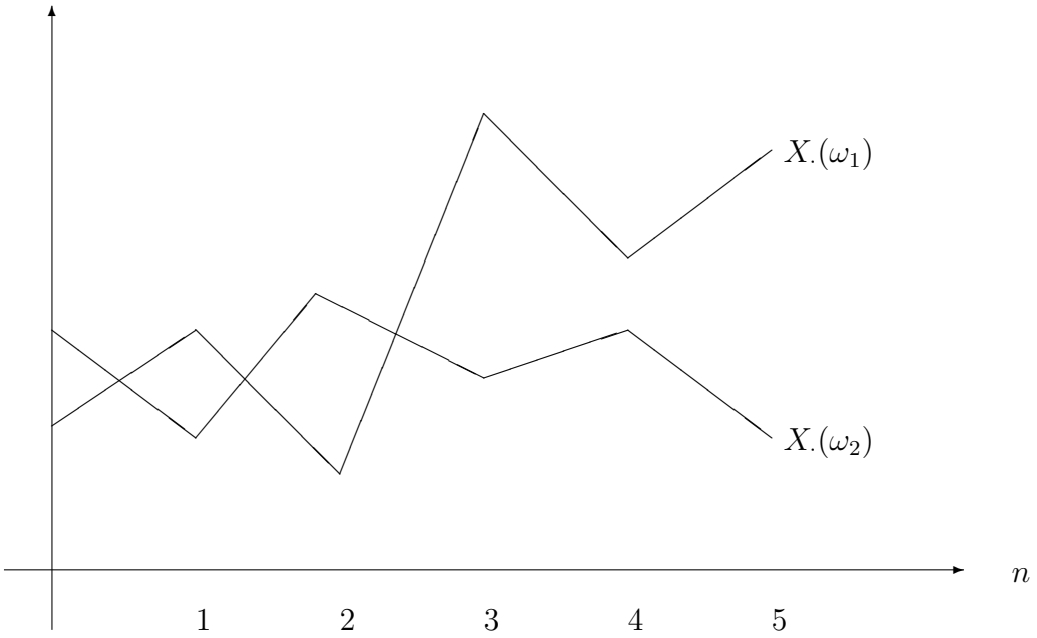

Figure 3: Sample paths of stochastic process X.

If $\{X_n\}_{n=0}^{\infty}$ is real valued, then $\{X_n\}_{n=0}^{\infty}$ can be viewed as a random variable with values in $\mathbf{R}^{\mathbf{N}}$ (path space; here $\mathbf{N}$ is the set of integers).

It is important to quantify the information history or flow corresponding to a stochastic process. This is achieved using **filtration**.

Definition 3.5 *A **filtration** of a measurable space $(\Omega, \mathcal{F})$ is a family $\{\mathcal{F}_n\}_{n=0}^{\infty}$ of sub-σ-algebra $\mathcal{F}_n \subset \mathcal{F}$ such that*

$$\mathcal{F}_n \subset \mathcal{F}_{n+1} \quad , \quad \forall\, n = 0, 1, \ldots$$

Example 3.6 If $\{X_n\}$ is a stochastic process,

$$\mathcal{G}_n \stackrel{\triangle}{=} \sigma(X_0, X_0, \ldots, X_n)$$

defines a filtration $\{\mathcal{G}_n\}$, $\mathcal{G}_n \subset \mathcal{F}$ called the **history** of $\{X_n\}$ or filtration **generated** by $\{X_n\}$. □

Definition 3.7 *A stochastic process $\{X_n\}$ is **adapted** to a filtration $\{\mathcal{F}_n\}$ if X_n is $\mathcal{F}_n$-measurable for all $n = 0, 1, 2, \ldots$.*

Remark 3.8 We can consider $\mathcal{F}_n$ are events that have occurred **up to** time n.

3.1.4 Martingales

Definition 3.9 *A stochastic process $\{X_n\}$ defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}, P)$ is called a supermartingale/martingale/submartingale if*

1. *$\{X_n\}$ is adapted to $\{\mathcal{F}_n\}$*
2. *$E[|X_n|] < +\infty$*
3. *$E[X_{n+1}|\mathcal{F}_n] \leq X_n$ / $E[X_{n+1}|\mathcal{F}_n] = X_n$ / $E[X_{n+1}|\mathcal{F}_n] \geq X_n$, $\forall n$.*

A martingale corresponds intuitively to a fair game of chance.

Note: For $m \geq n$,

$$\begin{aligned} E[X_m|\mathcal{F}_n] &= E[X_{n+1}|\mathcal{F}_n] \\ &= X_n \end{aligned}$$

if $\{X_n\}$ is a martingale. In general, $\forall m \geq n$, $E[X_m|\mathcal{F}_n]$ is the predicted value of X_m given the history $\mathcal{F}_n$.

A supermartingale (submartingale) decreases (increases) in conditional mean.

Example 3.10 Let $\{\mathcal{F}_n\}$ be a filtration on $(\Omega, \mathcal{F})$ and define

$$\mathcal{F}_\infty = \sigma(\mathcal{F}_n, \quad n \in \mathbf{N})$$

Let

$$Y \in L_1(\Omega, \mathcal{F}_\infty, P)$$

then

$$X_n \stackrel{\triangle}{=} E[Y|\mathcal{F}_n]$$

is a martingale. □

Example 3.11 Consider i.i.d (independent identically distributed) r.v's.

$$W_n = \begin{cases} 1, & \text{w.p. } \frac{1}{2}; \\ -1, & \text{w.p. } \frac{1}{2}. \end{cases}$$

and

$$\begin{aligned} X_n &= \sum_{k=0}^{n} W_k \\ &= X_{n-1} + W_n \end{aligned} \quad .$$

X is a martingale w.r.t. $\mathcal{F}_n = \sigma(W_0, W_1, \ldots, W_n)$ since

$$\begin{aligned} E[X_{n+1}|\mathcal{F}_n] &= E[X_n + W_{n+1}|\mathcal{F}_n] \\ &= X_n + E[W_{n+1}|\mathcal{F}_n] && \text{(since } X_n \text{ is } \mathcal{F}_n\text{-msble)} \\ &= X_n + E[W_{n+1}] && \text{(by independence)} \\ &= X_n && \text{(martingale)} \end{aligned}$$

□

3.1.5 Semimartingales

Definition 3.12 *A stochastic process $A = \{A_n\}$ is said to be **predictable** with respect to the filtration $\{\mathcal{F}_n\}$ if A_n is $\mathcal{F}_{n-1}$-measurable.*

Example 3.13 Suppose $\mathcal{F}_n = \sigma(Z_0, Z_1, \ldots, Z_n)$ for some stochastic process $\{Z_n\}$. If A is predictable w.r.t. $\{\mathcal{F}_n\}$, then A_n is a function of $Z_0, Z_1, \ldots, Z_{n-1}$. i.e. the value of A_n is determined by the values of $Z_0, Z_1, \ldots, Z_{n-1}$. □

Definition 3.14 *Let X be a stochastic process adapted to $\{\mathcal{F}_n\}$. If*

$$X = A + M$$

*where A is $\{\mathcal{F}_n\}$-predictable and M is $\{\mathcal{F}_n\}$-martingale, then X is called a **semimartingale**.*

Remark 3.15 Semimartingales are important in engineering and physics :

$$X = \underbrace{\text{“signal” or “trend”}}_{A} + \underbrace{\text{“noise”}}_{M}$$

A is sometimes called the “compensator” of X (relative to $\{\mathcal{F}_n\}$).

Theorem 3.16 *(**Doob Decomposition**). If X is an integrable adapted process, then X is a semimartingale.*

Proof. Define

$$\begin{aligned} A_n &= \sum_{k=1}^{n} E[X_k - X_{k-1}|\mathcal{F}_{k-1}] \\ M_n &= X_n - A_n \end{aligned}$$

By definition, A is $\{\mathcal{F}_n\}$-predictable. Next,

$$\begin{aligned} E[M_n|\mathcal{F}_{n-1}] &= E[X_n|\mathcal{F}_{n-1}] - \sum_{k=1}^{n} E[E[X_k - X_{k-1}|\mathcal{F}_{k-1}]|\mathcal{F}_{n-1}] \\ &= E[X_n|\mathcal{F}_{n-1}] - E[E[X_n - X_{n-1}|\mathcal{F}_{n-1}]|\mathcal{F}_{n-1}] - \sum_{k=1}^{n-1} E[X_k - X_{k-1}|\mathcal{F}_{k-1}] \\ &= X_{n-1} - A_{n-1} \\ &= M_{n-1} \end{aligned}$$

M is a $\{\mathcal{F}_n\}$-martingale. □

Remark 3.17 Write $\Delta X_n = X_n - X_{n-1}$ then

$$\Delta X_n = \underbrace{\Delta A_n}_{\text{prediction}} + \underbrace{\Delta M_n}_{\text{“innovations” or new information}}$$

$$\begin{aligned} \Delta A_n &= E[\Delta X_n | \mathcal{F}_{n-1}] \\ E[\Delta M_n | \mathcal{F}_{n-1}] &= 0 \end{aligned}$$

where ΔM_n is martingale increment. This is called the **innovations representation**.

Example 3.18 Let

$$X_n = \sum_{k=0}^{n} W_k$$

where $\{W_k\} \sim N(0,1)$, i.i.d. By the innovation representation, $\Delta X_n = W_n$, $\Delta A_n = E[\Delta X_n | \mathcal{F}_{n-1}] = 0$ so that $\Delta X_n = \Delta M_n = W_n$ □

3.1.6 Markov Processes

Definition 3.19 *Let X be a stochastic process defined on $(\Omega, \mathcal{F}, P)$ and taking values in a measurable space $(S, \mathcal{S})$ and adapted to $\{\mathcal{F}_n\}$. X is called a **Markov process** if for $m \geq n$*

$$P(X_m \in A | \mathcal{F}_n) = P(X_m \in A | X_n) \quad , \quad \forall A \in \mathcal{S}$$

That is, conditional probabilities of the future behavior of X given the whole past depend only on the current value.

Example 3.20 Let $S = \mathbf{R}$ and $\mathcal{S} = \mathcal{B}(\mathbf{R})$. W is i.i.d. process.

$$X_{n+1} = b(X_n, W_n)$$

where $b : \mathbf{R}^2 \to \mathbf{R}$. X is a Markov process w.r.t. $\mathcal{F}_n = \sigma(W_0, W_1, \ldots, W_{n-1})$.

$$\begin{aligned} E[\phi(X_{n+1}) | \mathcal{F}_n] &= E[\phi(b(X_n, W_n)) | \mathcal{F}_n] \\ &= \int_{-\infty}^{\infty} \phi(b(X_n, w)) P_{W_n}(dw) \\ &= E[\phi(X_{n+1}) | X_n] \end{aligned}$$

Setting $\phi = I_A$ we see that X is Markov. □

Definition 3.21 *The (one step) **transition probability** of a Markov process X are given by the transition kernel*

$$P_{X_{n+1}|X_n}(A|x) = P(X_{n+1} \in A | X_n = x)$$

Example 3.22 In the last previous example,

$$P_{X_{n+1}|X_n}(A|x) = \int_{-\infty}^{\infty} I_A(b(x,w))P_{W_n}(dw)$$

□

Proposition 3.23 ***(Chapman-Kolmogorov)***.

$$\begin{aligned} P_{X_{n+1}}(A) &= \int_S P_{X_{n+1}|X_n}(A|x)P_{X_n}(dx) \\ E[\phi(X_{n+1})] &= \int_S \phi(x')P_{X_{n+1}}(dx') \\ &= \int_S \phi(x') \int_S P_{X_{n+1}|X_n}(dx'|x)P_{X_n}(dx) \end{aligned}$$

PROOF.

$$\begin{aligned} P_{X_{n+1}}(A) &= E[I_A(X_{n+1})] \\ &= E[E[I_A(X_{n+1})|\mathcal{F}_n]] \\ &= E[E[I_A(X_{n+1})|X_n]] \\ &= E[P_{X_{n+1}|X_n}(A|X_n)] \\ &= \int_S P_{X_{n+1}|X_n}(A|x)P_{X_n}(dx) \end{aligned}$$

□

Example 3.24 (**Markov Chain**). Now suppose $S = \{s_1, s_2, \ldots, s_N\}$. Let $P = (P_{ij})$ be an $N \times N$ matrix such that

$$\sum_{j=1}^{N} P_{ij} = 1 \quad , \quad a_{ij} \geq 0$$

Define the transition kernel

$$P_{X_{n+1}|X_n}(s_j|s_i) = P_{ij}$$

and the probabilities

$$p_n(i) = P_{X_n}(s_i) = P(X_n = s_i)$$

Then, the Chapman-Kolmogorov equation reads

$$p_{n+1}(j) = \sum_{i=1}^{N} P_{ij}p_n(i)$$

i.e.

$$\begin{aligned} p_{n+1} &= P^* p_n \\ p_n &= (P^*)^n p_0 \end{aligned}$$

where $p_0(s_i) = P(X_0 = s_i) = \rho(i)$ is the initial distribution. Thus the probability distribution of X_n satisfies the dynamics

$$\begin{cases} p_{n+1} = P^* p_n \\ p_0 = \rho \end{cases}$$

Here we regard p_n as a column vector. If, instead, we regard p_n as a row vector, then

$$p_{n+1} = p_n P$$

□

Remark 3.25 In general, suppose $P_{X_{n+1}|X_n}$ is independent of time n. We write

$$\begin{aligned} p(A|x) &= P_{X_{n+1}|X_n}(A|x) \\ \rho(A) &= P_{X_0}(A) \\ p_n(A) &= P_{X_n}(A) \end{aligned}$$

then

$$\begin{cases} p_{n+1} = P^* p_n \\ p_0 = \rho, \end{cases}$$

where

$$(P^* p)(A) = \int_S p(A|x) p(dx)$$

3.1.7 Observation Processes

Definition 3.26 *Let X be a Markov process with values in S. An **observation process** is an O valued process defined by the **observation probabilities** $P_{Y|X}$*

$$P(Y_n \in B) = \int_S P_{Y|X}(B|x) P_{X_n}(dx) \quad , \quad B \in \mathcal{B}(\mathbf{R})$$

Note that for each $x \in S$, $P_{Y|X}(\cdot|x)$ is a probability measure on $(O, \mathcal{B}(O))$.

Example 3.27 Let $\{V_n\}$ be i.i.d. independent of X. Define

$$Y_n = h(X_n, V_n)$$

then

$$\begin{aligned} P_{Y|X}(B|X_n) &= P(Y_n \in B|X_n) \\ &= P(h(X_n, V_n) \in B|X_n) \\ &= \int I_B(h(X_n, V_n)) P_{V_n}(dv) \end{aligned}$$

□

3.1.8 Linear Representation of a Markov Chain

Suppose we replace S by $S' = \{e_1, e_2, \ldots, e_N\}$ where e_i is the i-th unit vector in $\mathbf{R}^N$. Then $X_n = e_i$ for some $i = 1, 2, \ldots, N$

Lemma 3.28 *Let X_n be a Markov chain with probability matrix P. Then*

$$\Delta M_{n+1} \stackrel{\triangle}{=} X_{n+1} - P^* X_n$$

is a martingale increment.

Corollary 3.29 *X has the semimartingale or innovations representation*

$$X_{n+1} = P^* X_n + \Delta M_{n+1}$$

or

$$\Delta X_{n+1} = (P^* - I) X_n + \Delta M_{n+1}$$

3.2 Controlled State Space Models

We will define controlled stochastic systems in terms of controlled transition probabilities and output probabilities.

Definition 3.30 *A **stochastic system** (S, U, O, P, Q, ρ) is specified by*

- *Borel spaces S, U, O*
  (state, control or input, observation or output)
- *Transition probability P*
  $P(\cdot|x, u)$ is a probability measure on S , $\forall (x, u) \in S \times U$.
- *Output probability Q*
  $Q(\cdot|x)$ is a probability measure on O , $\forall x \in S$
- *Initial distribution ρ of X_0.*

The evolution of the system is described as follows. If $X_n = x \in S$ is the state at time n and if $U_n = u \in U$ is the control input applied at the time n, then the system moves to a new state X_{n+1} according to the probability measure $P(\cdot|x, u)$ and produces an output Y_n according to $Q(\cdot|x)$.

Example 3.31 Nonlinear Systems.
Let $S = \mathbf{R}^n$, $U = \mathbf{R}^m$ and $O = \mathbf{R}^p$. The processes $\{W_n\}$ and $\{V_n\}$ are i.i.d., independent, and taking values in $\mathbf{R}^r$ and $\mathbf{R}^s$ respectively. $X_0 \sim \rho$ is independent of W and V. Define

$$b : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^r \to \mathbf{R}^n$$

and

$$h : \mathbf{R}^n \times \mathbf{R}^s \to \mathbf{R}^p$$

such that

$$\begin{cases} X_{n+1} = b(X_n, U_n, W_n) \\ Y_n = h(X_n, V_n) \end{cases}$$

The transition and output probability measures are

$$\begin{aligned} P(A|x,u) &= \int_{\mathbf{R}^r} I_A(b(x,u,w)) P_{W_n}(dw) \\ &\forall\ A \in \mathcal{S} = \mathcal{B}(\mathbf{R}^n) \quad , \quad (x,w) \in \mathbf{R}^n \times U \end{aligned}$$

$$\begin{aligned} Q(B|x) &= \int_{\mathbf{R}^s} I_B(h(x,v)) P_{V_n}(dv) \\ &\forall\ B \in \mathcal{O} = \mathcal{B}(\mathbf{R}^p) \quad , \quad x \in S \end{aligned}$$

□

Example 3.32 Linear Systems.
This is a special case of nonlinear systems with b, h are linear and W, V and ρ are Gaussian : $W_n \sim N(0, Q)$, $V_n \sim N(0, R)$ and $\rho \sim N(\overline{x}_0, \Sigma_0)$:

$$\begin{aligned} X_{n+1} &= AX_n + BU_n + GW_n \\ Y_n &= CX_n + HV_n \end{aligned}$$

□

Example 3.33 Controlled Markov Chain (Finite State).
Consider $S = \{s_1, s_2, \ldots, s_N\}$, $U = \{u_1, u_2, \ldots, u_M\}$ and $O = \{o_1, o_2, \ldots, o_P\}$. $P(\cdot|x,u)$ is defined by a controlled transition matrix $P(u)$ where

$$\begin{aligned} \sum_{j=1}^{N} P_{ij}(u) &= 1 \quad , \quad P_{ij}(u) \geq 0\ ,\ \forall\ u \in U\ ,\ s_i \in S \\ P(s_j|s_i,u) &= P_{ij}(u) \end{aligned}$$

$Q(\cdot|x)$ is defined by an output probability matrix Q :

$$\begin{aligned} P(o_j|s_i) &= Q_{ij} \\ \sum_{j=1}^{P} Q_{ij} &= 1 \quad , \quad \forall\ s_i \in S \quad , \quad Q_{ij} \geq 0 \end{aligned}$$

□

3.2.1 Feedback Control Laws or Policies

A deterministic sequence $u = \{u_0, u_1, u_2, \ldots\}$ of control values is called an **open loop control**. A **feedback** or **closed loop** control depends on the observations. More precisely, let $g = \{g_0, g_1, g_2, \ldots\}$ be a sequence of functions

$$g_n : O^{n+1} \to U$$

The value u_n of the control at time n is given by

$$u_n = g_n(y_{0n})$$

where $y_{0n} = \{y_0, y_1, \ldots, y_n\}$. Such a sequence g is called a **feedback** or **closed loop control policy** or **law**. Note that an open loop control is a special case of a closed loop policy.

A policy g determines a probability measure P^g on the canonical sample or path space $(\Omega, \mathcal{F})$,

$$\begin{aligned} \Omega &= (S \times U \times Y)^\infty \\ \mathcal{F} &= \mathcal{B}(\Omega) \end{aligned}$$

Note $\omega \in \Omega$ is of the form

$$\omega = \{x_0, u_0, y_0, x_1, u_1, y_1 \ldots\}$$

Let

$$\begin{aligned} \mathcal{F}_n &= \sigma(X_0, Y_0, X_1, Y_1, \ldots, X_n, Y_n) \\ \mathcal{Y}_n &= \sigma(Y_0, Y_1, \ldots, Y_n) \end{aligned}$$

Then $U_n = g_n(Y_{0n})$ is adapted to $\mathcal{Y}_n$ and X_n, Y_n is adapted to $\mathcal{F}_n$. Note that now $\mathcal{Y}_n = \sigma(Y_0, Y_1, \ldots, Y_n, U_0, U_1, \ldots, U_n)$ since g is used.

3.2.2 Partial and Full State Information

In general, the only information available to the controller is contained in the output or observation process Y; X is in general unavailable. This is the case of **partial information**. If $O = S$ and $Q(\cdot|x) = \delta_x$ (Dirac measure), then $Y \equiv X$, and one has **full state information**.

Control problems are much easier to solve when one has full state information. In the general case of partial information, one must do some **filtering** as well.

3.3 Filtering

3.3.1 Introduction

Consider the (non-dynamic) filtering problem

$$\begin{cases} X &= \mu + W \qquad \text{(signal)} \\ Y &= CX + V \qquad \text{(observation)} \end{cases} \tag{107}$$

where $W \sim N(0, Q)$, $V \sim N(0, R)$ are independent. Let's compute $\widehat{X} = E[X|Y]$ using two methods:

1. The "innovations" method. We will see that $\widehat{X}$ has the representation

$$\widehat{X} = \mu + \alpha \widetilde{Y}$$

where $\widetilde{Y} \triangleq Y - C\mu$ is the "innovation".

Note: We use the following notation:
Means :

$$\begin{aligned} \overline{X} &\triangleq E[X] = \mu \\ \overline{Y} &\triangleq E[Y] = C\mu \end{aligned}$$

Covariances :

$$\begin{aligned} \Sigma_X = \quad Cov(X) &= E[(X-\mu)(X-\mu)'] \\ &= Q \\ \Sigma_Y = \quad Cov(Y) &= E[(Y-C\mu)(Y-C\mu)'] \\ &= CQC' + R \\ \Sigma_{XY} = Cov(X,Y) &= E[(X-\mu)(Y-C\mu)'] \\ &= QC' \end{aligned}$$

We claim

$$\widehat{X} = \mu + \Sigma_{XY}\Sigma_Y^{-1}(Y - C\mu).$$

Proof. Write

$$\begin{aligned} \widehat{\nu} &= \mu + \Sigma_{XY}\Sigma_Y^{-1}(Y - C\mu) \\ \widetilde{\nu} &= X - \widehat{\nu} \end{aligned}$$

then

$$\begin{aligned} E[\widetilde{\nu}] &= 0 \\ E[\widetilde{\nu}(Y - \overline{Y})'] &= 0. \end{aligned}$$

Then since these random variables are Gaussian, $\widetilde{\nu}$ and Y are independent (orthogonal). Then,

$$\begin{aligned} \widehat{X} &= E[X - \widehat{\nu} + \widehat{\nu}|Y] \\ &= E[\widetilde{\nu}|Y] + E[\widehat{\nu}|Y] \\ &= 0 + \widehat{\nu} \end{aligned}$$

This gives the "*Kalman Filter*" for the problem :

$$\widehat{X} = \mu + QC'(CQC' + R)^{-1}(Y - C\mu).$$

□

Note that if we write error estimates by $e = X - \widehat{X}$, the conditional error covariance is given by

$$\begin{aligned} \Sigma_e &= \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{XY}' \\ &= Q - QC'(CQC' + R)^{-1}CQ \end{aligned}$$

Remark 3.34 $\widehat{X} - \mu$ is the projection of $\widetilde{X} = X - \mu$ onto the subspace

$$\mathcal{Y} = \{\alpha \widetilde{Y} \ : \ \alpha \in \mathbf{R}\}$$

relative to the inner product $E[\widetilde{X}\widetilde{Y}] = Cov(X, Y)$, see Figure 4.

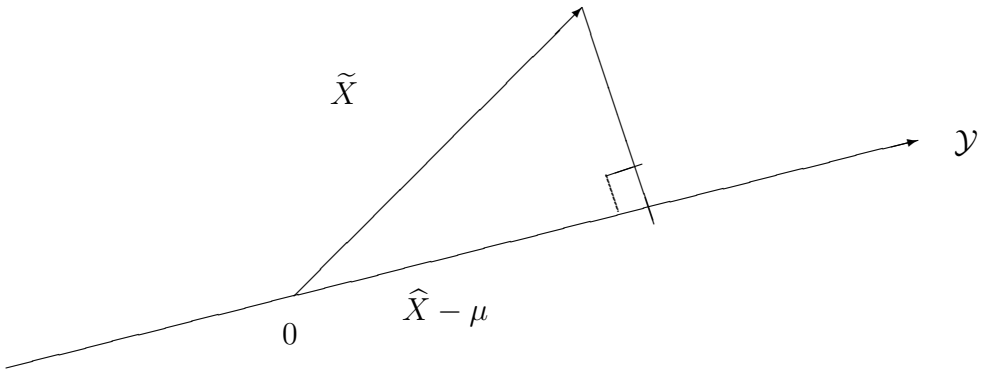

Figure 4: Projection interpretation.

i.e. $\widehat{X}$ must satisfy $\widetilde{X} - (\widehat{X} - \mu) \ \perp \ \widetilde{Y}$
i.e. $(\widetilde{X} - (\widehat{X} - \mu) \ , \ \widetilde{Y}) = 0$. Since $\widehat{X} - \mu = \alpha \widetilde{Y}$, we get $(\widetilde{X}, \widetilde{Y}) = \alpha(\widetilde{Y}, \widetilde{Y})$. Thus,

$$\begin{aligned} \alpha &= (\widetilde{X}, \widetilde{Y})(\widetilde{Y}, \widetilde{Y})^{-1} \\ &= \Sigma_{XY} \Sigma_Y^{-1} \end{aligned}$$

2. Using the "reference probability" approach.
We consider the following processes are under a probability measure P,

$$\begin{cases} X &= \mu + W \\ Y &= CX + V \end{cases}$$

The joint density of X, Y is

$$p_{X,Y}(x, y) = p_X(x) p_{Y|X}(y|x)$$

in which

$$\begin{aligned} p_X(x) &= (2\pi|Q|)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}(x - \mu)' Q^{-1}(x - \mu)) \\ p_{Y|X}(y|x) &= (2\pi|R|)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}(y - Cx)' R^{-1}(y - Cx)) \end{aligned}$$

Using a new "*reference*" probability measure P^+, we have the new processes

$$\begin{cases} X &= \mu + W \\ Y &= V^+ \end{cases}$$

where $V^+ \sim N(0,R)$, $W \sim N(0,Q)$, V^+ and W are independent. The joint density of X, Y is now

$$p^+_{X,Y}(x,y) = p_X(x)p_V(y)$$

where

$$p_V(y) = (2\pi|R|)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}y'R^{-1}y)$$

Write $\dfrac{dP}{dP^+} = \Lambda$, where

$$\Lambda(x,y) = \exp(-\tfrac{1}{2}|R^{-1}Cx|^2 + y'R^{-1}Cx)$$

By Bayes' rule,[1]

$$E[X|Y] = \frac{E^+[X\Lambda|Y]}{E^+[\Lambda|Y]}$$

the conditional density is given by

$$\begin{aligned} p_{X|Y}(x|y) &= \frac{\Lambda(x,y)\, p_X(x)}{\displaystyle\int_{\mathbf{R}^n} \Lambda(x,y)p_X(x)\, dx} \\ &= \frac{q_{X|Y}(x|y)}{\displaystyle\int_{\mathbf{R}^n} q_{X|Y}(x|y)\, dx} \end{aligned}$$

We call $q_{X|Y}$ the "*unnormalised conditional density*"

$$q_{X|Y}(x|y) = \Lambda(x,y)p_X(x)$$

Now we compute,

$$\begin{aligned} \widehat{X} = E[X|Y=y] &= \int_{\mathbf{R}^n} x\, p_{X|Y}(x|y)\, dx \\ &= \frac{\displaystyle\int_{\mathbf{R}^n} x\, q_{X|Y}(x|y)\, dx}{\displaystyle\int_{\mathbf{R}^n} q_{X|Y}(x|y)\, dx} \\ &= \mu + QC'(CQC' + R)^{-1}(Y - C\mu). \end{aligned}$$

Remark : The main techniques for nonlinear filtering are the innovations approach and the reference probability method.

[1]Prove this using the invariance property (101) by showing that $E[X|Y]E^+[\Lambda|Y] = E^+[X\Lambda|Y]$.

3.3.2 The Kalman Filter

Consider the linear gaussian stochastic system :

$$\left\{ \begin{array}{rl} X_{k+1} & = AX_k + GW_k \\ Y_k & = CX_k + HV_k \end{array} \right.$$

where the initial condition $X_0 \sim N(\overline{X}_0, \Sigma_0)$, $W_k \sim N(0, Q)$and $V_k \sim N(0, R)$ are all mutually independent with $Q > 0, R > 0$ and $\Sigma_0 > 0$.

Filtering problem : to compute conditional expectations of the type

$$\widehat{\phi}_k = E[\phi(X_k)|Y_{0,k}]$$

Since all the random variables are Gaussian, the conditional distribution

$$P_{X_k|Y_{0,k}}(A|Y_{0,k}) = P(X_k \in A|Y_{0,k})$$

is Gaussian with density

$$\begin{aligned} p_{X_k|Y_{0:k}}(x|Y_{0:k}) &\stackrel{\triangle}{=} p_{k|k}(x|Y_{0:k}) \\ &= (2\pi|\Sigma_{k|k}|)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}(x - X_{k|k})' \, \Sigma_{k|k}^{-1} \, (x - X_{k|k})) \end{aligned}$$

where

$$\begin{array}{rl} X_{k|k} & = E[X_k|Y_{0,k}] \\ \Sigma_{k|k} & = E[(X - X_{k|k})(X - X_{k|k})'|Y_{0,k}] \end{array}$$

Note that the conditional distribution is completely determined by the *finite dimensional* quantities $X_{k|k} \in \mathbf{R}^n$ and $\Sigma_{k|k} \in \mathbf{R}^{n^2}$.

The filtering problem will be solved if we can find simple recursion for these two quantities. Thus

$$\widehat{\phi}_k = \int_{\mathbf{R}^n} \phi(x) p_{k|k}(x|Y_{0,k}) \, dx$$

Theorem 3.35 *(**Kalman Filter**, [38, chapter 7], [1, chapter3]). The conditional density $p_{k|k} \sim N(X_{k|k}, \Sigma_{k|k})$ is obtained from*

$$\begin{array}{rl} X_{k+1|k+1} & = AX_{k|k} + L_{k+1}(Y_{k+1} - CAX_{k|k}) \\ X_{0|0} & = L_0 \widetilde{Y}_0 + X_0 \\ \Sigma_{k+1|k+1} & = (I - L_{k+1}C)\Sigma_{k+1|k} \\ \Sigma_{k+1|k} & = A\Sigma_{k|k}A' + GQG' \\ \Sigma_{0|0} & = (I - L_0 C)\Sigma_0 \\ L_k & \stackrel{\triangle}{=} \Sigma_{k|k-1}C'(C\Sigma_{k|k-1}C' + HRH')^{-1} \\ L_0 & \stackrel{\triangle}{=} \Sigma_0 C'(C\Sigma_0 C' + HRH')^{-1} \end{array}$$

Notation. $p_{k+1|k}(x|Y_{0,k})$ denotes the conditional density of X_{k+1} given $Y_{0,k}$.

$$\begin{array}{rl} X_{k+1|k} & = E[X_{k+1}|Y_{0,k}] \\ \Sigma_{k+1|k} & = E[(X_{k+1} - X_{k+1|k})(X_{k+1} - X_{k+1|k})'|Y_{0,k}] \\ p_{k+1|k} & \sim N(X_{k+1|k}, \Sigma_{k+1|k}) \end{array}$$

3.3.3 The Kalman Filter for Controlled Linear Systems

Let $g = \{g_0, g_1, \dots, \}$ be a feedback policy, possibly nonlinear. This determines :

$$\begin{aligned} X^g_{k+1} &= AX^g_k + BU^g_k + GW_k \\ Y^g_k &= CX^g_k + HV_k \\ U^g_k &= g_k(Y^g_{0,k}) \end{aligned}$$

The processes X^g, Y^g need not be Gaussian. However, the conditional densities $p_{k+1|k}, p_{k|k}$ are Gaussian.

Theorem 3.36 *The conditional means $X^g_{k|k}$ and $X^g_{k+1|k}$ are given by*

$$\begin{aligned} X^g_{k+1|k+1} &= AX^g_{k|k} + BU^g_k + L_{k+1}(Y^g_{k+1} - C(AX^g_{k+1} + BU^g_{k+1})) \\ X^g_{k+1|k} &= AX^g_{k|k-1} + BU^g_k + AL_k(Y^g_k - CX^g_{k|k-1}) \\ X^g_{0|0} &= \overline{X}_0 + L_0(Y_0 - C\overline{X}_0) \\ X^g_{0|-1} &= \overline{X}_0 \end{aligned}$$

$\Sigma_{k|k}$, $\Sigma_{k+1|k}$ and L_k are given as before.

Proof. Write

$$X^g_k = \overline{X}^g_k + X_k \quad , \quad Y^g_k = \overline{Y}^g_k + Y_k$$

where

$$\begin{aligned} \overline{X}^g_{k+1} &= A\overline{X}^g_k + BU^g_k \quad , \quad \overline{X}^g_0 = \overline{X}_0 \\ \overline{Y}^g_k &= C\overline{X}^g_k \\ X_{k+1} &= AX_k + GW_k \quad , \quad X_0 \sim N(0, \Sigma_0) \\ Y_k &= CX_k + GV_k \end{aligned}$$

Then the conditional density of X^g_k given $(Y^g_{0,k}, U^g_{0,k-1})$ is the conditional density of X_k given $Y_{0,k}$ shifted by $\overline{X}^g_k$. (See [38, page 102-3] for details.) □

3.3.4 The HMM Filter (Markov Chain)

Let X, Y be a HMM defined by a transition matrix P and an output probability matrix Q :

$$\begin{aligned} P(X_{k+1} = s_j | X_k = s_i) &= p_{ij} \\ P(Y_k = o_j | X_k = s_i) &= q_{ij} \end{aligned}$$

This determines a measure P on a measurable space $(\Omega, \mathcal{F})$ (canonical path space, say).

Filtering : We wish to compute the conditional probabilities

$$P(X_k = s_j | Y_{0,k}).$$

If we use the linear representation of Markov chain, we want to compute the (column) vector

$$p_{k|k} \stackrel{\triangle}{=} E[X_k | Y_{0,k}]$$

Recall that in the linear representation, $X_k = P^* X_{k-1} + \Delta M_k$ where ΔM_k is a martingale increment w.r.t. $\mathcal{F}_k = \sigma(X_{0,k}, Y_{0,k})$.

Since Y_k depends on ΔM_k, it is not clear how to proceed. We will use a **reference probability** P^+ under which Y_k is independent of ΔM_k.

Under P, the joint density of $X_{0,k}, Y_{0,k}$ is

$$p_{X_{0,k},Y_{0,k}}(x_{0,k}, y_{0,k}) = p_{X_{0,k}}(x_{0,k})\, p_{Y_{0,k}|X_{0,k}}(y_{0,k}|x_{0,k})$$

where

$$\begin{aligned} p_{X_{0,k}}(x_{0,k}) &= \prod_{l=0}^{k-1} P_{x_l,x_{l+1}} \rho_{x_0} \\ p_{Y_{0,k}|X_{0,k}}(y_{0,k}|x_{0,k}) &= \prod_{l=0}^{k} Q_{x_l,y_l} \end{aligned}$$

ρ_{x_0} is initial distribution of X.

The new measure P^+ corresponds to the joint density

$$p^+_{X_{0,k},Y_{0,k}}(x_{0,k}, y_{0,k}) = p_{X_{0,k}}(x_{0,k})\, p^+_{Y_{0,k}|X_{0,k}}(y_{0,k}|x_{0,k})$$

where

$$p^+_{Y_{0,k}|X_{0,k}}(y_{0,k}|x_{0,k}) = \tfrac{1}{p^{k+1}}$$

where $p = \#O$ is the number of output values. Then,

$$\begin{aligned} \Lambda_k &= \left.\frac{dP}{dP^+}\right|_{\mathcal{F}_k} \\ &= p^{k+1} \prod_{l=0}^{k} Q_{x_l,y_l} \end{aligned}$$

Notation. We write Ψ_k for the diagonal matrix with entries pQ_{i,y_k}, $i = 1, \ldots, n$. And we write

$$\sigma_k = E^+[X_k \Lambda_k | Y_{0,k}]$$

(an n-vector since X_k is an n-vector using the linear representation

Note: Under P^+, X is a Markov chain with transition matrix P and initial distribution ρ, Y is i.i.d. uniform and, X and Y are independent.

Bayes' rule :

$$p_{k|k} = \frac{\sigma_k}{\langle \sigma_k, 1 \rangle}$$

where $\langle \sigma_k, 1 \rangle = \sum_{i=1}^n \sigma_k(i) = E^+[\Lambda_k | Y_{0,k}]$. $\langle ., . \rangle$ = inner product in $\mathbf{R}^n$.

Theorem 3.37 *(HMM Filter)*

$$\begin{cases} p_{k|k} &= \frac{1}{c_k} \Psi_k P^* p_{k-1|k-1} \\ p_{0|0} &= \rho \\ c_k &= \langle \Psi_k P^* p_{k-1|k-1}, 1 \rangle \\ c_0 &= 1 \end{cases}$$

Remark 3.38 In unnormalized form we have

$$\begin{cases} \sigma_k &= \Psi_k P^* \sigma_{k-1} \\ \sigma_0 &= \rho \end{cases}$$

which is a *linear* recursion.

PROOF.

- Step 1
  Recursion for σ_k.

$$\begin{aligned} \sigma_k(j) &= E^+[X_k(j)\Lambda_k | Y_{0,k}] \\ &= E^+[(P^* X_{k-1} + \Delta M_k)_j \, pQ_{j,y_k} \Lambda_{k-1} | Y_{0,k}] \\ &= \Psi_k(j) \textstyle\sum_{i=1}^n p_{ij} \, E^+[X_{k-1}(i) \ \Lambda_{k-1} | Y_{0,k-1}] \\ &\quad \text{(by independence).} \\ &= \Psi_k(j) \textstyle\sum_{i=1}^n p_{ij} \ \sigma_{k-1}(i) \end{aligned}$$

- Step 2
  By induction,

$$\langle \sigma_k, 1 \rangle = \prod_{l=0}^{k} c_l$$

  To see this,

$$\begin{aligned} p_{k|k} &= \frac{\sigma_k}{\langle \sigma_k, 1 \rangle} \\ &= \frac{\Psi_k P^* \sigma_{k-1}}{\langle \sigma_k, 1 \rangle} \\ &= \frac{\Psi_k P^* p_{k-1|k-1} \langle \sigma_{k-1}, 1 \rangle}{\langle \sigma_k, 1 \rangle} \end{aligned}$$

  This gives

$$c_k = \frac{\langle \sigma_k, 1 \rangle}{\langle \sigma_{k-1}, 1 \rangle}$$

  and clearly

$$\begin{aligned} \prod_{l=0}^{k} c_l &= \tfrac{\langle \sigma_k, 1 \rangle}{\langle \sigma_{k-1}, 1 \rangle} \ \tfrac{\langle \sigma_{k-1}, 1 \rangle}{\langle \sigma_{k-2}, 1 \rangle} \cdots \tfrac{\langle \sigma_1, 1 \rangle}{\langle \sigma_0, 1 \rangle} \\ &= \langle \sigma_k, 1 \rangle \end{aligned}$$

Note that the (nonlinear) recursion for $p_{k|k}$ is just formula (6.2) in [38, page 83]. □

3.3.5 Filter for Controlled HMM

Consider a controlled Markov chain with transition matrix $P(u)$ and output probability matrix Q. Let g be a feedback policy

$$U_k = g_k(Y_{0,k})$$

We want to compute the conditional expectation

$$P(X_k = s_j | Y_{0,k}, U_{0,k-1})$$

i.e.

$$p^g_{k|k} = E[X_k | Y_{0,k}, U_{0,k-1}]$$

Theorem 3.39 *([38, page 81]). $p^g_{k|k}$ does not depend on g. It satisfies*

$$\begin{cases} p^g_{k|k} &= \frac{1}{c_k} \Psi_k P^*(u_{k-1})\, p^g_{k-1|k-1} \\ p^g_{0|0} &= \rho \\ c_k &= \langle \Psi_k P^* p_{k-1|k-1}, 1 \rangle \\ c_0 &= 1 \end{cases}$$

as before.

Definition 3.40 *A stochastic process $\xi = \{\xi_k\}$ is called an **information state** if*

- *ξ_k is a function of $Y_{0,k}, U_{0,k-1}$*
- *ξ_{k+1} can be determined from ξ_k, y_{k+1} and u_k.*

Example 3.41 $p_{k|k} = p^g_{k|k}$ is an information state. $\sigma_{k|k} = \sigma p^g_{k|k}$ is also an information state. For linear systems, $X_{k|k}$ is an information state □

3.4 Dynamic Programming - Case I : Complete State Information

The stochastic system is described by (S, U, P, ρ). A control policy $g = \{g_0, g_1, \ldots\}$ determines control U and state X processes :

$$\begin{aligned} &u_k = g_k(X_{0,k}) \\ &P(X_{k+1} \in A | X_{0,k}, U_{0,k}) = P(A | X_k, U_k) \end{aligned}$$

Let

$$G = \{g \ : \ g \text{ is a state feedback policy}\}$$

denote the set of admissible controllers.

Cost function :

$$J(g) = E^g[\sum_{l=0}^{M-1} L(X_l, U_l) + \Phi(X_M)]$$

where

$$\begin{aligned} L &: S \times U \to \mathbf{R} \\ \Phi &: S \to \mathbf{R} \\ L &\geq 0 \quad , \quad \Phi \geq 0 \end{aligned}$$

3.4.1 Optimal Control Problem

Find $g^* \in G$ minimizing J; i.e.

$$J(g^*) \leq J(g) \ , \ \forall g \in G$$

We solve this problem using **dynamic programming**.

To this end, we define the value function

$$\begin{aligned} V_k(x) &= \inf_{g \in G_{k,M-1}} E^g_{x,k}[\sum_{l=k}^{M-1} L(X_l, U_l) + \Phi(X_M)] \\ &\quad \text{for } k = 0, 1, \ldots, M-1 \\ V_M(x) &= \Phi(x) \end{aligned}$$

where $G_{k,l}$ denotes policies $g_{k,l} = \{g_k, g_{k+1}, \ldots, g_l\}$

$V_k(x)$ is the optimal "cost to go", given that we start at x at time k; i.e. $X_k = x$.

Theorem 3.42 *([38, chapter 6]) V satisfies the dynamic programming equation :*

$$\text{(DPE)} \quad \begin{cases} V_k(x) &= \inf_{u \in U} \{L(x,u) + \int_S V_{k+1}(z) P(dz|x,u)\} \\ &\quad \text{for } k = 0, 1, \ldots, M-1 \\ V_M(x) &= \Phi(x) \end{cases}$$

Further, let $\overline{u}^*_k(x)$ denotes a control value which achieves the minimum in (DPE), for $k = 0, 1, \ldots, M-1$. Then the policy g^* defined by

$$u_k = g^*_k(X_{0,k}) = \overline{u}^*_k(X_k)$$

*is **optimal**, and $J(g^*) = \int_S V_0(z)\rho(dz)$.*

Remark 3.43 Policies g for which g_k is only a function of X_k (and not $X_{0,k-1}$) are called **Markov** policies. Then w.r.t. P^g, when g is a Markov policy, X is a Markov process.

So the optimal policy g^* above is Markov and the optimal state process X_k is Markov :

$$P^{g^*}(X_{k+1} \in A | X_{0,k}, U_{0,k}) = P(A|x, g^*_k(x))$$

Remark 3.44 Algorithm for finding the optimal state feedback controller :

1. Set $k = M$, $V_M(x) = \Phi(x)$
2. Set $k = M - 1$. Then solve (DPE) to get $V_{M-1}(x)$ and $\overline{u}^*_{M-1}(x)$
3. Set $k = M - 2$. Then solve (DPE) to get $V_{M-2}(x)$ and $\overline{u}^*_{M-2}(x)$
4. Continue, until $k = 0$
5. Set

$$\begin{aligned} g^*_0(X_0) &= \overline{u}^*_0(X_0) \\ g^*_1(X_{0,1}) &= \overline{u}^*_1(X_1) \\ &\vdots \\ g^*_k(X_{0,k}) &= \overline{u}^*_k(X_k) \\ &\vdots \\ g^*_{M-1}(X_{0,M-1}) &= \overline{u}^*_{M-1}(X_{M-1}) \end{aligned}$$

Proof. Define W by

$$\begin{cases} W_k(x) &= \inf_{u \in U}\{L(x,u) + \int_S W_{k+1}(z)P(dz|x,u)\} \\ & \text{for } k = 0, 1, \ldots, M-1 \\ W_M(x) &= \Phi(x) \end{cases}$$

Our goal is to prove $W_k(x) = V_k(x)$, and g^* optimal. Define

$$\overline{V}_k(x,g) = E^g_{x,k}[\sum_{l=k}^{M-1} L(X_l, U_l) + \Phi(X_M)]$$

Claim $(*)$:

$$\begin{aligned} W_k(x) &= \inf_{g \in G_{k,M-1}} \overline{V}_k(x,g) \stackrel{\triangle}{=} V_k(x) \\ &= \overline{V}_k(x, g^*) \end{aligned}$$

Assume $(*)$ true for $k+1, \ldots, M$ (trivial for $k = M$). Then for any $g \in G_{k,M-1}$

$$\begin{aligned} \overline{V}_k(x,g) &= E^g_{x,k}[L(x,g_k) + \textstyle\sum_{l=k+1}^{M-1} L(X_l, U_l) + \Phi(X_M)] \\ &= E^g_{x,k}[L(x,g_k) + \overline{V}_{k+1}(X_{k+1}, g_{k+1,M-1})] \\ &\geq E^g_{x,k}[L(x,g_k) + W_{k+1}(X_{k+1})] \\ &\geq W_k(x) \end{aligned}$$

by induction, with equality if $g = g^*$

Example 3.45 Controlled Markov chain with transition matrix $P(u)$.
The dynamic programming equation is :

$$\begin{cases} V_k(s_i) & = \min_{u \in U} \{L(s_i, u) + \sum_{j=1}^{n} V_{k+1}(s_j) P_{ij}(u)\} \\ & \text{for } k = 0, 1, \ldots, M-1 \\ V_M(s_i) & = \Phi(s_i) \end{cases}$$

If we regard V_k as a column vector in $\mathbf{R}^n$, etc, this is just

$$\begin{cases} V_k & = \min_{u \in U} \{L^u + P(u) V_{k+1}\} \\ & \text{for } k = 0, 1, \ldots, M-1 \\ V_M & = \Phi \end{cases}$$

This is a **nonlinear backward** recursion. The optimal Markov policy g^* is

$$g_k^*(X_{0,k}) = \overline{u}_k^*(X_k)$$

where $\overline{u}_k^*(s_i)$ achieves min in (DPE) and $\overline{u}_k^*$ can be viewed as a vector. □

Example 3.46 Nonlinear Systems

$$X_{k+1} = b(X_k, U_k, W_k)$$

where $W = \{W_k\}$ i.i.d., independent of $X_0 \sim \rho$. The DPE

$$\begin{cases} V_k(x) & = \inf_{u \in U} \{L(x, u) + \int_{\mathbf{R}^n} V_{k+1}(b(x, u, w)) P_W(dw)\} \\ & \text{for } k = 0, 1, \ldots, M-1 \\ V_M(x) & = \Phi(x) \end{cases}$$

V_k is a function of $x \in \mathbf{R}^n$. $\overline{u}_k^*(x)$ is also a function of $x \in \mathbf{R}^n$. □

Example 3.47 Linear Systems (LQG)

$$X_{k+1} = AX_k + BU_k + GW_k$$

Then the DPE

$$\begin{cases} V_k(x) & = \inf_{u \in \mathbf{R}^m} \{L(x, u) + \int_{\mathbf{R}^n} V_{k+1}(Ax + Bu + Gw)(2\pi |Q|)^{-\frac{1}{2}} \exp(-\frac{1}{2} w' Q^{-1} w)\, dw\} \\ & \text{for } k = 0, 1, \ldots, M-1 \\ V_M(x) & = \Phi(x) \end{cases}$$

Quadratic cost :

$$\begin{aligned} L(x,u) &= \tfrac{1}{2}x'\Gamma x + \tfrac{1}{2}u'\Lambda u \\ \Phi(x) &= \tfrac{1}{2}x'\overline{P}x \end{aligned}$$

where $\Gamma \geq 0$, $\overline{P} \geq 0$, $\Lambda > 0$. To simplify, assume

$$\begin{aligned} G = I \quad &, \quad Q = I \\ X_{k+1} &= AX_k + BU_k + W_k \end{aligned}$$

$\{W_k\}$ i.i.d. $\sim N(0, I)$

Theorem 3.48 *(LQG)*

$$\begin{aligned} V_k(x) &= \tfrac{1}{2}x'P_kx + \alpha_k \\ \overline{u}_k^*(x) &= L_kx \end{aligned}$$

where

$$\text{(R)} \begin{cases} P_k &= \Gamma + A'P_{k+1}A - A'P_{k+1}B[\Lambda + B'P_{k+1}B]^{-1}B'P_{k+1}A \\ & \text{for } k = 0, 1, \ldots, M-1 \\ P_M &= \overline{P} \\ L_k &= -[\Lambda + B'P_{k+1}B]^{-1}B'P_{k+1}A \\ \alpha_k &= \tfrac{1}{2}\sum_{l=k+1}^{M} TrP_l \quad , \quad k = 0, 1, \ldots, M-1 \quad , \quad \alpha_M = 0 \end{cases}$$

Note that :

1. The Riccati equation (R) is a **nonlinear backward** matrix recursion.
2. The optimal Markov policy is

$$u_k = g_k^*(X_{0,k}) = L_kX_k$$

which is **linear**.

Proof. By induction. It it true for $k = M$, (trivial). Assume true for $k+1, \ldots, M$. Plug $V_{k+1}(z) = \frac{1}{2}z'P_{k+1}z + \alpha_{k+1}$ into (DPE) and evaluate the Gaussian integral (do this). Then minimize over u. This gives $\overline{u}_k^*$. Plug back in to evaluate $V_k(x)$. This gives the Riccati equation (do this!). □

With complete state information, the optimal controller is **state feedback**, Figure 5.

3.5 Dynamic Programming - Case II : Partial State Information

Consider general stochastic system

$$(S, U, O, P, Q, \rho)$$

Control policies $g \in G$ are **output feedback**:

$$u_k = g_k(Y_{0,k})$$

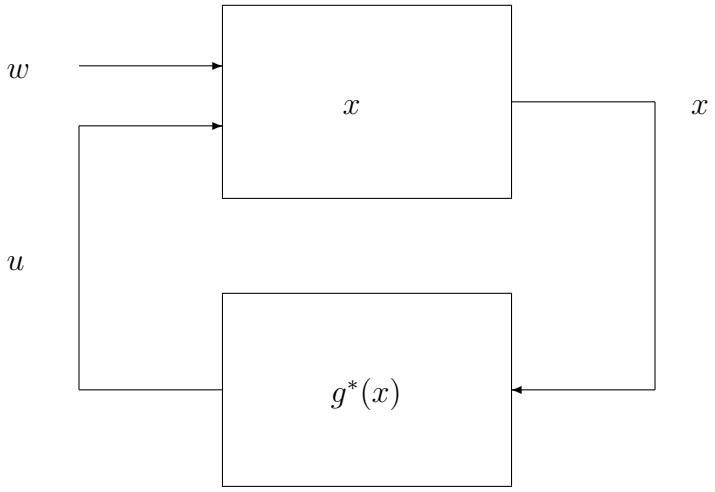

Figure 5: State feedback.

Cost function :

$$J(g) = E^g[\sum_{l=0}^{M-1} L(X_l, U_l) + \Phi(X_M)]$$

as before, except g is a function of Y only.

Partially observed optimal control :
Find an output feedback controller $g^* \in G$ such that

$$J(g^*) \leq J(g) \ , \ \forall g \in G$$

Rather that treat the general case, we first do HMM's, and then linear systems.

3.5.1 Optimal Control of HMM's

Consider controlled HMM's with transition matrix $P(u)$ and output probability matrix Q. To solve this problem, we could use any suitable **information state**. Here we will use

$$\pi_k = p_{k|k}$$

Recall that

$$\text{(IS)} \begin{cases} \pi_k &= T(\pi_{k-1}, U_{k-1}, Y_k) \\ \pi_0 &= \rho \end{cases}$$

where

$$T(\pi, u, y) = \frac{\Psi(y)P^*(u)\pi}{\langle \Psi(y)P^*(u)\pi, 1 \rangle}.$$

We will show that optimal output feedback control of the HMM is equivalent to optimal control of (IS), a problem with full state information (π_k is the new state, which is observed!).

To see how this possible, write

$$\begin{aligned} J(g) &= E^g[\textstyle\sum_{i=0}^{M-1} L(X_i, U_i) + \Phi(X_M)] \\ &= E^g[\textstyle\sum_{i=0}^{M-1} E^g[L(X_i, U_i)|Y_{0,i}] + E^g[\Phi(X_M)|Y_{0,M}]] \\ &= E^g[\textstyle\sum_{i=0}^{M-1} \langle \pi_i, L^{u_i}\rangle + \langle \pi_M, \Phi\rangle] \end{aligned}$$

where

$$\begin{aligned} L^u &= L(., u) \\ \langle \pi, f\rangle &= \textstyle\sum_{i=1}^{n} \pi(i) f(i) \end{aligned}$$

This $J(g)$ has been expressed purely in terms of the information state π_k.

Theorem 3.49 *([38, chapter 6]) Define*

$$\begin{cases} V_k(\pi) &= \inf\limits_{g \in G_{k,M-1}} E^g_{\pi,k}[\sum\limits_{l=k}^{M-1} \langle \pi_l, L^{u_l}\rangle + \langle \pi_M, \Phi\rangle] \\ & \text{for } k = 0, 1, \ldots, M-1 \\ V_M(\pi) &= \langle \pi, \Phi\rangle \end{cases}$$

Then V_k satisfies the dynamic programming equation

$$\text{(DPE)} \begin{cases} V_k(\pi) &= \min\limits_{u \in U}\{\langle \pi, L^u\rangle + \sum\limits_{y \in O} V_{k+1}(T(\pi, u, y)) p(y|\pi, u)\} \\ & \text{for } k = 0, 1, \ldots, M-1 \\ V_M(\pi) &= \langle \pi, \Phi\rangle \end{cases}$$

where

$$p(y|\pi, u) = \sum_{i,j=1}^{n} Q_{j,y}\, P_{ij}(u) \pi(i).$$

Further, if $\overline{u}^*_k(\pi)$ achieves the minimum in (DPE), then the policy

$$u_k = g^*_k(Y_{0,k}) = \overline{u}^*_k(\pi_k)$$

*is **optimal**.*

Proof. See [38, page 85]. □

Note: The optimal controller depends on the observations $Y_{0,k}$ through the information state π_k. Such controllers are called **separated**, i.e. separated into a filter plus a controller, Figure 6.

3.5.2 Optimal Control of Linear Systems (LQG)

Consider a linear system

$$\begin{cases} X_{k+1} &= AX_k + BU_k + GW_k \\ Y_k &= CX_k + HV_k \end{cases}$$

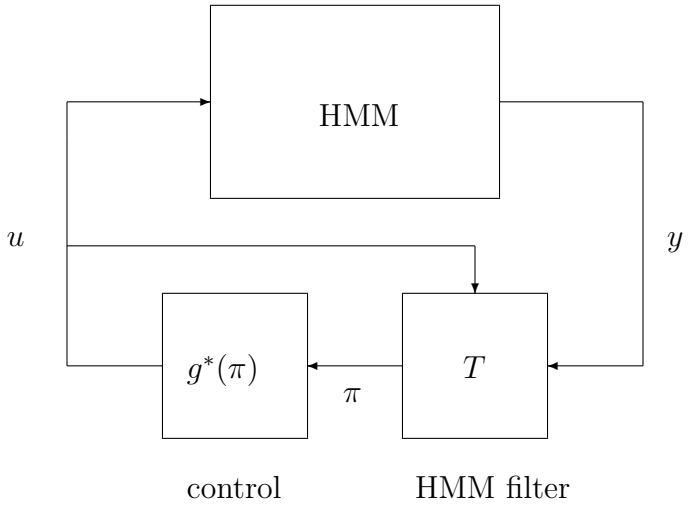

Figure 6: Separation structure for HMM controller.

with quadratic cost

$$\begin{aligned} L(x,u) &= \tfrac{1}{2}x'\Gamma x + \tfrac{1}{2}u'\Lambda u \\ \Phi(x) &= \tfrac{1}{2}x'\overline{P}x \end{aligned}$$

where $\Gamma \geq 0$, $\overline{P} \geq 0$, $\Lambda > 0$. Assume for simplicity, $G = I$, $H = I$, $Q = I$, $R = I$.

The conditional density

$$\pi_k = p_{k|k} \sim N(X_{k|k}, \Sigma_{k|k})$$

is an information state. Since $\Sigma_{k|k}$ is deterministic, $X_{k|k}$ is itself an information state for the linear system. Thus we expect the optimal policy g^* to have the form

$$u_k = g^*(Y_{0,k}) = \overline{u}_k^*(X_{k|k})$$

for a suitable function $\overline{u}_k^*(x)$. It turns out that $\overline{u}_k^*(x)$ is the complete state information controller derived earlier.

Theorem 3.50 *Let $X_{k|k}$ be the conditional mean as determined by the Kalman filter. Let L_k be the gain sequence determined by the state feedback linear quadratic problem. Then*

$$u_k = g^*(Y_{0,k}) = L_k X_{k|k}$$

*is the **optimal** policy for the partially observed LQG problem.*

Note that this optimal controller is separated.

PROOF. See [38, page 105]. □

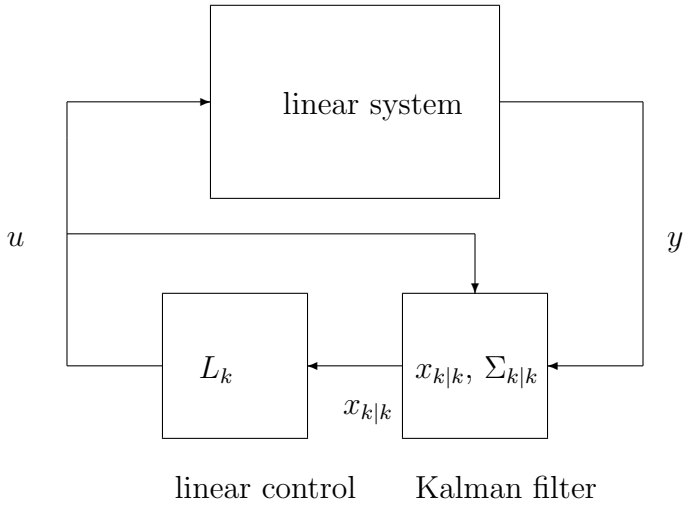

Figure 7: LQG separation structure.

3.6 Two Continuous Time Problems

3.6.1 System and Kalman Filter

Consider the (classical) system model

$$dx_c = (Ax_c + Bu)\,dt + G\,dw \tag{108}$$
$$dy = Cx_c\,dt + dv \tag{109}$$

where

1. $x_c(t)$ is the state of the system (a vector), given by a linear stochastic differential equation driven by Gaussian white noise $\dot{w}(t)$,

2. $\dot{y}(t)$ represents the measured output variables, which are subject to additive Gaussian white noise $\dot{v}(t)$,

3. $u(t)$ is the control variable (a vector), which we take to be a function of the measurement data $y(s)$, $0 \le s \le t$, which we write $u(t) = \mathbf{K}(y(s), 0 \le s \le t)$,

4. A, B, G and C are appropriately sized matrices, and

5. $\dot{w}(t)$ and $\dot{v}(t)$ have zero mean, variance 1, and correlation α.

Throughout we interpret stochastic differential equations in the Ito sense, so that, e.g. $dw(t)dw(t) = dt$, $dv(t)dv(t) = dt$, and $dw(t)dv(t) = \alpha dt$.

The continuous time Kalman filter equations are

$$d\check{x}_c = (A\check{x}_c + Bu)dt + (Y_c C^T + G\alpha^T)(dy - C\check{x}_c dt) \tag{110}$$

and

$$\dot{Y}_c = AY_c + Y_cA^T + GG^T - (Y_cC^T + G\alpha^T)(CY_c + \alpha G^T) \tag{111}$$

with initial conditions $\check{x}_c(0) = x_{c0}$, $Y_c(0) = Y_0$ (the initial Gaussian parameters), see [49, section 4.5], [16, Theorem 4.4.1 (when $\alpha = 0$)].

3.6.2 LQG Control

In continuous time, the classical *linear quadratic Gaussian* (CLQG) problem is to minimize the cost function

$$J(\mathbf{K}) = \mathbf{E}[\int_0^T [\tfrac{1}{2}x_c^T(s)Px_c(s) + \tfrac{1}{2}u^T(s)Qu(s)]\,ds + \tfrac{1}{2}x_c^T(T)Sx_c(T)]. \tag{112}$$

for the linear gaussian model (108), (109).

The optimal control is given in terms of the Kalman filter

$$u^\star(t) = -Q^{-1}B^TX_c(t)\check{x}_c(t), \tag{113}$$

where

$$\dot{X}_c + A^TX_c + X_cA - X_cBQ^{-1}B^TX_c + P = 0, \tag{114}$$

and $X_c(T) = S$, [49, Part I], [16, Theorem 5.3.3 (when $\alpha = 0$)].

3.6.3 LEQG Control

Another important class of stochastic control problems are the so-called *risk-sensitive* problems. The risk-sensitive cost functions attempt to give larger penalties for large deviations from desired values and tend to lead to designs which are more robust than LQG controllers. In this subsection we switch to continuous time.

The classical *linear exponential quadratic Gaussian* (CLEQG) problem (a type of risk-sensitive problem) is specified by modifying the LQG cost using the exponential function as follows:

$$J^\mu(\mathbf{K}) = \mathbf{E}[\exp\mu(\int_0^T [\tfrac{1}{2}x_c^T(s)Px_c(s) + \tfrac{1}{2}u^T(s)Qu(s)]\,ds + \tfrac{1}{2}x_c^T(T)Sx_c(T))]. \tag{115}$$

Here, μ is a parameter indicating the sensitivity to risk. If $\mu > 0$, the problem is *risk-sensitive*, and consequently gives greater weight to the size of the integral compared to the LQG case. If $\mu < 0$ is referred to as *risk-seeking*. The *risk-neutral* case $\mu \to 0$ corresponds to the LQG problem above,

$$\frac{1}{\mu}\log J^\mu(\mathbf{K}) \to J(\mathbf{K}), \tag{116}$$

see [8], [37].

The CLEQG criterion is potentially infinite if $\mu > 0$ is too large. However, when finite the problem is solvable and the optimal solution is

$$u^\star(t) = -Q^{-1}B^T X_c^\mu(t)[I - \mu Y_c^\mu(t) X_c^\mu(t)]^{-1}\check{x}_c^\mu(t), \tag{117}$$

where

$$d\check{x}_c^\mu \;=\; ([A + \mu Y_c^\mu P]\check{x}_c^\mu + Bu)dt + (Y_c^\mu C^T + G\alpha^T)(dy - C\check{x}_c^\mu dt), \tag{118}$$

$$\dot{Y_c^\mu} \;=\; AY_c^\mu + Y_c^\mu A^T + GG^T + \mu Y_c^\mu P Y_c^\mu - (Y_c^\mu C^T + G\alpha^T)(CY_c^\mu + \alpha G^T), \tag{119}$$

with initial conditions $\check{x}_c^\mu(0) = x_{c0}$, $Y_c^\mu(0) = Y_0$, and

$$\dot{X_c^\mu} \;+\; A^T X_c^\mu + X_c^\mu A - X_c^\mu[BQ^{-1}B^T - \mu GG^T]X_c^\mu + P = 0, \tag{120}$$

with terminal condition $X_c^\mu(T) = S$, see [8, Theorem 4.1 (when $\alpha = 0$)], [49, section 8.5], [14, Theorem 3.11].

It is important to note that the CLEQG filter (118), (119) is *not* the LQG Kalman filter, but reduces to it when $\mu = 0$. The CLEQG state $\check{x}_c^\mu(t)$ can still be interpreted as a state estimator for $x_c(t)$, but it is not the optimal minimum mean square error estimator: in general $\check{x}_c^\mu(t) \neq \check{x}_c(t)$. However, $\check{x}_c^\mu(t)$ does serve as an information state for the CLEQG problem, since the cost function (115) can be expressed in terms of it, see [8, Theorem 3.2]:

$$J^\mu(\mathbf{K}) = \mathbf{E}[\int \exp(-\frac{\mu}{2}x^T S x)\pi^\mu(x,t)dx] \tag{121}$$

where $\pi^\mu(x,t)$ is a gaussian density with mean $\check{x}^\mu(t)$ and covariance $Y_c^\mu(t)$. In general, it does not seem possible to represent the CLEQG cost in terms of the Kalman filter state $\check{x}_c(t)$.

4 Robust Control

References for this section include [30], [54], [19], [26], [4], [46], [31].

4.1 Introduction and Background

As mentioned in the Introduction, section 1, robustness concerns the ability of a controller to cope with uncertainty, disturbances, and model errors. Typically, a controller is designed on the basis of an idealized model, which may neglect some features of the real system. The parameters of the model may or may not be well known. Furthermore, the real system may be subject to disturbances, noise and dissipative effects. A robust controller should have good performance under nominal conditions, and adequate performance in other than nominal conditions.

Figure 8 illustrates a common setup for robust control design.

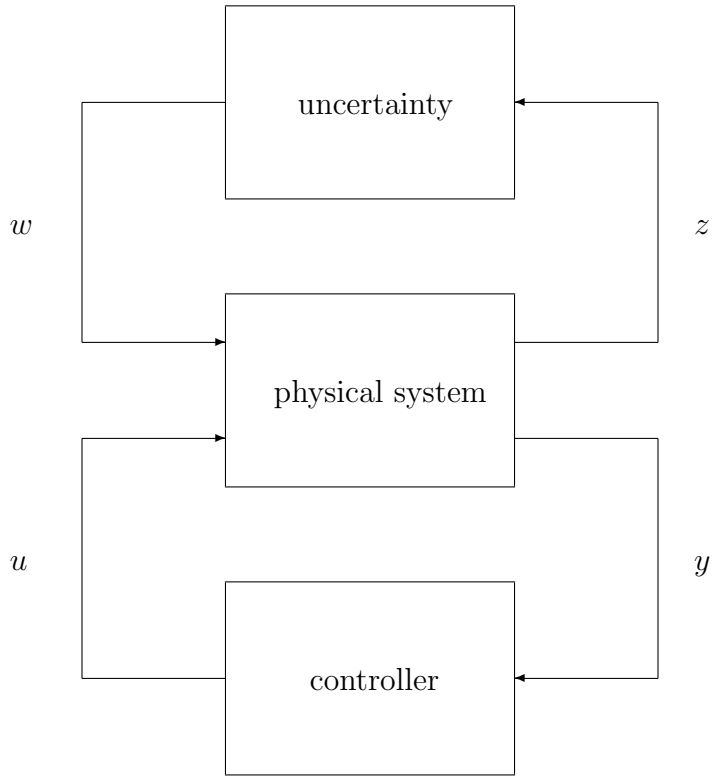

Figure 8: Setup for robust control.

Note that uncertainty has been explicitly included in the setup. The robust control problem is to design a controller achieving the desired performance, explicitly taking into account the presence of uncertainty. Robustness can be guaranteed only for the classes of uncertainty that have been modeled.

Standard approaches to robust control design include the use of H^∞ methods, and integral quadratic constraint (IQC) methods. In H^∞, which originated in the frequency domain [53] and gave rise to the (unwieldy) name, systems are regarded as possibly nonlinear operators acting on signals, and stability and robustness are characterized in terms of the boundedness of the operators. The controller is chosen to make the gain or norm γ of the closed loop physical system-controller combination considered as an operator from disturbance inputs w to disturbance outputs z "small". In view of the *small gain theorem* (see, e.g. [47, Theorem 6.1.1 (1)]), the design is capable of accommodating uncertainties of gain less than $1/\gamma$. Detailed knowledge of the uncertainty is not required, save for the gain limitation and structural connections to the physical system. IQC design is closely related, but here the uncertainty is characterized by an integral quadratic constraint, and the controller is designed subject to this constraint. It is important to know that not all robust control problems are *feasible*, for instance a designer may choose too small a value of γ, and not be able to find a solution.

4.2 The Standard Problem of H_∞ Control

Now we describe in more detail the standard problem of H_∞ control. The problem entails a description of the plant and controller models, and definitions of the control objectives. The standard control problem corresponds to the Figure 9, which we now explain.

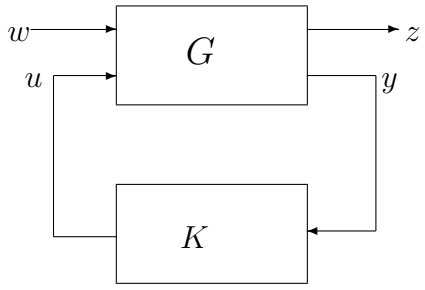

Figure 9: The Closed-Loop System (G, K)

4.2.1 The Plant (Physical System Being Controlled)

Consider plants G with the following structure

$$G: \begin{cases} \dot{x} = A(x) + B_1(x)w + B_2(x)u \\ z = C_1(x) + D_{12}(x)u \\ y = C_2(x) + D_{21}(x)w \end{cases} \tag{122}$$

Here, $x(t) \in \mathbf{R}^n$ denotes the state of the system, and is not in general directly measurable; instead an output $y(t) \in \mathbf{R}^p$ is observed. The additional output quantity $z(t) \in \mathbf{R}^r$ is a

performance measure, depending on the particular problem at hand. The control input is $u(t) \in \mathbf{R}^m$, while $w(t) \in \mathbf{R}^s$ is regarded as an opposing disturbance input. Detailed assumptions concerning the functions appearing in (122) are given in [31]. Here we note that the origin 0 is assumed to be an equilibrium, and $A(0) = 0$, $C_1(0) = 0$, $C_2(0) = 0$.

4.2.2 The Class of Controllers

The plant G was described by an explicit state space model and is assumed given. However, in the spirit of optimal control theory, we do *not* prescribe a state space model for the controller K, since it is an unknown to be determined from the control objectives. Rather, we simply stipulate some basic input-output properties required of any *admissible controller*, namely that the controller must be a causal function of the output

$$K \,:\, y(\cdot) \mapsto u(\cdot)$$

and the resulting closed loop system be well-defined in the sense that trajectories and signals exist and are unique. The controller K will be said to be *null-initialized* if $K(0) = 0$, regardless of whether or not a state space realization of K is given.

4.2.3 Control Objectives

The H_∞ control problem is commonly thought of as having two objectives: find a controller K such that the closed loop system (G, K) is

1. dissipative, and
2. stable.

In §4.3 we define what is meant by these terms in the case of linear systems. We now describe their meanings for nonlinear systems; this gives an extension of H_∞ control to nonlinear systems.[2]

The closed loop system (G, K) is γ-*dissipative* if there exist $\gamma > 0$ and a function $\beta(x_0) \geq 0$ $\beta(0) = 0$, such that

$$\begin{cases} \frac{1}{2}\displaystyle\int_0^T |z(s)|^2\, ds \;\leq\; \gamma^2 \tfrac{1}{2} \int_0^T |w(s)|^2\, ds + \beta(x_0) \\ \\ \text{for all} \;\; w \in L_{2,T} \text{ and all } T \geq 0. \end{cases} \tag{123}$$

This definition is saying that the nonlinear input-output map $(G, K) : w \mapsto z$ defined by the closed loop system has finite L_2 gain with a bias term due to the initial state x_0 of the plant G. This inequality expresses the effect of the uncertainty w on the system performance z.

[2]The term "nonlinear H_∞ control" has no precise mathematical meaning, but it has come into common use in the control engineering community and refers to nonlinear generalizations of H_∞ control (which has precise meaning for linear systems).

While dissipation captures the notion of performance of a control system, another issue with H_∞ control is stability of the system. The closed loop system will be called *weakly internally stable* provided that if G is initialized at any x_0, then if $w(\cdot) \in L_2[0,\infty)$, all signals $u(\cdot), y(\cdot), z(\cdot)$ in the loops as well as $x(\cdot)$ converge to 0 as $t \to \infty$. By *internal stability* we mean that the closed loop is weakly internally stable and in addition if the controller has a state space realization, then the controller state will converge to an equilibrium as $t \to \infty$.

Dissipation and stability are closely related; see, e.g. [50], [32], [33], [46]. Indeed, dissipative systems which enjoy a detectability or observability property also enjoy a stability property. In our context, suppose the system (G, K) is *z-detectable*, that is, $w(\cdot)$ and $z(\cdot) \in L^2[0,\infty)$ imply $x(\cdot) \in L^2[0,\infty)$ and $x(t) \to 0$ as $t \to \infty$. By *z-observable* we mean that if $w(\cdot) = 0$, $(\cdot) = 0$, then $x(\cdot) = 0$. *If (G, K) is γ-dissipative and z-detectable, then (G, K) is weakly internally stable* (see [31, Theorem 2.1.3]).

Solutions to this problem are descrbed in [46], [4], [31].

4.3 The Solution for Linear Systems

We recall here the well-known solution to the H_∞ control problem for linear systems, see [19], [45], [30], etc. The class of linear systems considered are of the form

$$G: \begin{cases} \dot{x} = Ax + B_1 w + B_2 u \\ z = C_1 x + D_{12} u \\ y = C_2 x + D_{21} w. \end{cases} \tag{124}$$

4.3.1 Problem Formulation

The class of admissible controllers K are those with finite dimensional linear state space realizations

$$K: \begin{cases} \dot{\eta} = A_K \eta + B_{1K} y + B_{2K} u \\ u = C_K + D_K y \end{cases}$$

Given $\gamma > 0$, the H_∞ control problem for G is to find, if possible, a controller K such that the resulting closed loop system $(G, K) : w \mapsto z$ satisfies:

(i) *Dissipation:* The required dissipation property is expressed in the frequency domain in terms of the H_∞ norm of the closed loop transfer function $(G, K)(s)$ as follows:

$$\| (G, K) \|_{H_\infty} = \sup_{\omega \in \mathbf{R}} \sigma_{max}[(G, K)(j\omega)] < \gamma.$$

(ii) *Stability*: We require that the closed loop system

$$(G, K) \text{ is } \textit{internally stable}.$$

4.3.2 Background on Riccati Equations

Recall a few facts about Riccati equations. An algebraic Riccati equation

$$\Sigma\, A + A\Sigma + \Sigma\, R\, \Sigma + Q = 0 \tag{125}$$

with real matrix entries A, R, Q and R, Q selfadjoint, meeting suitable positivity and technical conditions (see, e.g., [54, Chapter 13]), has upper and lower solutions Σ_a, Σ_r so that any other self adjoint solution Σ lies between them

$$\Sigma_a \leq \Sigma \leq \Sigma_r.$$

The bottom solution is called the *stabilizing solution* because it has and is characterized by the property

$$A + R\Sigma_a \tag{126}$$

is asymptotically stable. Likewise Σ_r is *antistabilizing* in that

$$-(A + R\Sigma_r) \tag{127}$$

is asymptotically stable.

4.3.3 Standard Assumptions

There are a number of "standard assumptions" that are needed for the necessity and sufficiency theorems about H_∞ control. These can be expressed in various ways and here we follow [45].

The first two conditions concern the rank of the matrices D_{12} and D_{21}:

$$D_{12}'D_{12} \;=\; E_1 > 0, \tag{128}$$
$$D_{21}D_{21}' \;=\; E_2 > 0. \tag{129}$$

Next are two important technical conditions which take the form

$$\text{rank} \begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix} = n + m \quad \text{for all } \omega \geq 0, \tag{130}$$

and

$$\text{rank} \begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix} = n + l \quad \text{for all } \omega \geq 0. \tag{131}$$

These two conditions are commonly used in LQG control and filtering, and concern the controllability and observability of underlying systems.

4.3.4 Problem Solution

The necessary and sufficient conditions for solvability of the H_∞ problem under the standard assumptions are:

Condition 1: *State feedback Control.* There exists $X_e \geq 0$ solving the control-type Riccati equation

$$\begin{aligned}(A - B_2E_1^{-1}D_{12}'C_1)'X_e + X_e(A - B_2E_1^{-1}D_{12}'C_1) \\ +X_e(\tfrac{1}{\gamma^2}B_1B_1' - B_2E_1^{-1}B_2')X_e + C_1'(I - D_{12}E_1^{-1}D_{12}')C_1 = 0\end{aligned} \tag{132}$$

which is stabilizing, i.e.,

$$A - B_2E_1^{-1}D_{12}'C_1 + (\gamma^{-2}B_1B_1' - B_2E_1^{-1}B_2')X_e \quad \text{is asymptotically stable.} \tag{133}$$

Condition 2: *State Estimation.* There exists $Y_e \geq 0$ solving the filter-type Riccati equation

$$\begin{aligned}(A - B_1D_{21}'E_2^{-1}C_2)Y_e + Y_e(A - B_1D_{21}'E_2^{-1}C_2)' \\ +Y_e(\gamma^{-2}C_1'C_1 - C_2'E_2^{-1}C_2)Y_e + B_1[I - D_{21}'E_2^{-1}D_{21}]B_1' = 0\end{aligned} \tag{134}$$

which is stabilizing, i.e.,

$$A - B_1D_{21}'E_2^{-1}C_2 + Y_e(\gamma^{-2}C_1'C_1 - C_2'E_2^{-1}C_2) \quad \text{is asymptotically stable.} \tag{135}$$

Condition 3: *Coupling.* The matrix X_eY_e has spectral radius strictly less than γ.

Theorem 4.1 *([19], [45], [30]) The H_∞ control problem for G, meeting certain technical conditions is solvable if and only if the above three conditions are satisfied. If these conditions are met, one controller, called the central controller, is given by*

$$\begin{aligned}\dot{\bar{x}} &= (A - B_2E_1^{-1}D_{21}'C_1 + (\gamma^{-2}B_1B_1' - B_2E_1^{-1}B_1')X_e)\bar{x} \\ &+(I - \gamma^{-2}Y_eX_e)^{-1}(B_1D_{21}' + Y_eC_2')E_2^{-1}(y - [C_2 + D_{21}B_1'X_e]\bar{x}) \\ &+(I - \gamma^{-2}Y_eX_e)^{-1}(B_2 + \gamma^{-2}Y_eC_1'D_{12})(u + E_1^{-1}[D_{12}'C_1 + B_2'X_e]\bar{x}), \\ u &= -E_1^{-1}[D_{12}'C_1 + B_2'X_e]\bar{x}.\end{aligned} \tag{136}$$

It is important to notice that the filter for the H^∞ controller in (136) is *not* the Kalman filter. This solution was first given in [29], [19].

4.4 Risk-Sensitive Stochastic Control and Robustness

An important connection between the *deterministic* H^∞ problem and the *stochastic* LEQG problem was discovered by Glover and Doyle, [29]. This can be seen by comparing the solutions to each problem: (136) and (117) with the identifications $\gamma = 1/\sqrt{\mu}$,

$$B_1 = \begin{pmatrix} G & 0 \end{pmatrix}, \quad B_2 = B, \tag{137}$$

$$C_1 = \begin{pmatrix} P^{1/2} \\ 0 \end{pmatrix}, \quad D_{12} = \begin{pmatrix} 0 \\ Q^{1/2} \end{pmatrix}, \tag{138}$$

$$C_2 = C, \quad D_{21} = \begin{pmatrix} 0 & I \end{pmatrix}. \tag{139}$$

Both problems are equivalent to dynamic games, [30], [4]. Connections with algebraic Riccati equations and disturbance attenuation were discovered in [44].

Mathematically, what underlies this connection is the interpretation of the risk-sensitive problem as an equivalent stochastic dynamic game. In general terms, this corresponds to a general convex duality formula (e.g., see [20, Chapter 1.4]):

$$\log \mathbf{E}_\mathbf{P}[e^f] = \sup_\mathbf{Q}\{\mathbf{E}_\mathbf{Q}[f] - H(\mathbf{Q} \parallel \mathbf{P})\} \tag{140}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are probability distributions, and where the relative entropy is defined by (e.g., see [42, Chapter 11])

$$H(\mathbf{Q} \parallel \mathbf{P}) = \mathbf{E}_\mathbf{Q}[\log \frac{d\mathbf{Q}}{d\mathbf{P}}].$$

The duality formula (140) implies

$$\log \mathbf{E}_\mathbf{P}[e^f] \geq \mathbf{E}_\mathbf{Q}[f] - H(\mathbf{Q} \parallel \mathbf{P}) \tag{141}$$

for any probability distribution $\mathbf{Q}$ which is absolutely continuous with respect to $\mathbf{P}$.

We let f be an integral of quadratic function

$$f = \mu \int_0^T [x'(t)Px(t) + u'(t)Qu(t)]dt,$$

and let $\mathbf{P}$ be a probability distribution corresponding to a nominal model used for controller design. We also set $\mathbf{Q}$ to be probability distribution corresponding to the true physical system, with deterministic disturbance w shifting the mean of the noise. Then

$$H(\mathbf{Q} \parallel \mathbf{P}) = \int_0^T |w(t)|^2 dt, \tag{142}$$

and (141) implies the gain-type inequality [21], [11]:

$$\mathbf{E}_{true}[\int_0^T [x'(t)Px(t) + u'(t)Qu(t)]dt] \leq \frac{1}{\mu}\{\log J^\mu + \int_0^T |w(t)|^2 dt\}; \tag{143}$$

cf. (123).

This shows that the performance (evaluated as a LQG mean with respect to the the true distribution $\mathbf{Q}$) of the LEQG controller (designed assuming the distribution $\mathbf{P}$) when applied to a real system is bounded from above by the sum of two terms. The first term is a constant determined by the optimal LEQG cost, while the second is a measure of the size of the "uncertainty". This second term is zero under nominal conditions ($w = 0$). See [11], [21] for further details.

5 Optimal Feedback Control of Quantum Systems

References for this section of the notes include [5], [35], [36], [42].

5.1 Preliminaries

An operator T on a Hilbert space $\mathbf{H}$ is *positive* if $\langle \psi, T\psi \rangle \geq 0$ for all $\psi \in \mathbf{H}$. A *generalized measurement* is defined as follows [42, section 2.2.3]. Let $\{M_n\}$ be a collection of operators such that

$$\sum_n M_n^\dagger M_n = I. \tag{144}$$

The probability of outcome n when the system is in state ρ is

$$\text{Prob}(n) = \langle \rho, M_n^\dagger M_n \rangle, \tag{145}$$

and the state of the system after the measurement outcome n is obtained is

$$\rho'(n) = \frac{M_n \rho M_n^\dagger}{\text{Prob}(n)}. \tag{146}$$

A *positive operator valued measure* (POVM) [42, section 2.2.6] is a collection $\{E_m\}$ of positive operators E_n such that

$$\sum_n E_n = I. \tag{147}$$

A generalized measurement $\{M_n\}$ defines a POVM $\{E_n\}$ via $E_n = M_n^\dagger M_n$. A POVM may define many generalized measurements, one of which is obtained by taking square roots.

It is straightforward to see that an orthogonal projective measurement $\{P_n\}$ (a PVM) defines a POVM and a generalized measurement.

Let $\mathfrak{M}_n$ denote the set of $n \times n$ complex matrices.

A linear map Γ mapping operators on $\mathbf{H}$ to operators is called *positive map* if $\Gamma(A)$ is positive for all positive operators A on $\mathbf{H}$. Γ is *completely positive* (c.p.), i.e. $(\Gamma \otimes I_n)(A)$ is positive for any positive operator A on $\mathbf{H} \otimes \mathfrak{M}_n$, for any n.

A quantum *operation* [34, section 3.1.2], [42, section 8.2.4] is a c.p. map Γ such that $0 \leq \text{tr}[\Gamma(A)] \leq \text{tr}[A]$ for all positive operators A on $\mathbf{H}$. A *dynamic map* or *trace-preserving operation* is a quantum operation that satisfies $\text{tr}[\Gamma(A)] = \text{tr}[A]$, [34, section 3.1.2].

Every c.p. map Γ can be expressed in operator-sum form

$$\Gamma(\rho) = \sum_n M_n \rho M_n^\dagger \tag{148}$$

for suitable operators $\{M_n\}$, [34, Corollary 3.1.1]. If Γ is an operation, then $\sum_n M_n^\dagger M_n \leq I$, [42, Theorem 8.1], [34, page 74] and if it is a trace-preserving operation then $\sum_n M_n^\dagger M_n =$

I. Note that if $\{M_n\}$ is a generalized measurement, then the state reduction can be expressed as

$$\rho'(n) = \frac{\Gamma(n)\rho}{\langle \Gamma(n)\rho, I\rangle}, \tag{149}$$

where $\Gamma(n) = M_n \rho M_n^\dagger$ is a quantum operation.

Please note that use of the above terminology varies somewhat in the literature, with some differences in the definitions. For instance, some authors may defined all operations to be normalized $\mathrm{tr}[\Gamma(A)] = \mathrm{tr}[A]$.

Quantum operations and c.p. maps Γ are associated with adjoints $\Gamma^\dagger$ via

$$\langle \Gamma\rho, X\rangle = \langle \rho, \Gamma^\dagger X\rangle \tag{150}$$

for states ρ and operators X. This is the basic Schrodinger-Heisenburg duality.

5.2 The Feedback Control Problem

We consider the optimal control of a quantum system using standard digital or analog electronics, Figure 10.

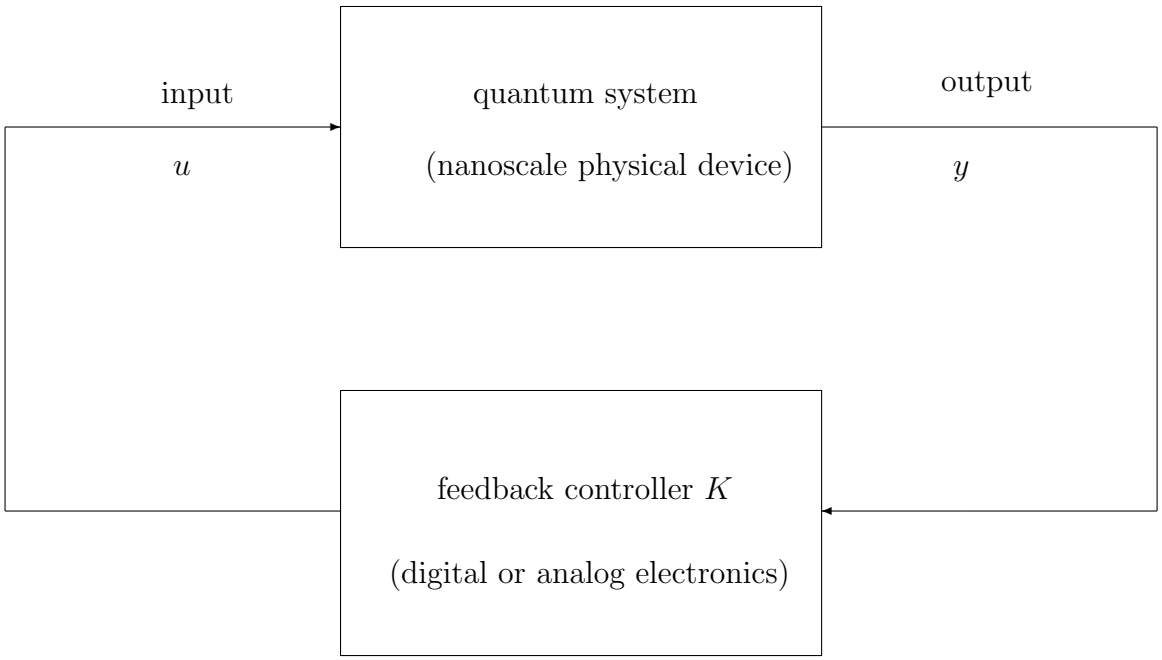

Figure 10: Feedback control of a quantum system

.

The problem is formulated as follows:

1. We use a simple discrete time model, and we assume all variables are discrete-valued.

2. Evolution of the quantum system is interleaved with measurements.

3. The controller is a causal function of the measurement data, so that

$$u_k = K_k(y_1, \ldots, y_k)$$

4. The quantum system dynamics are modelled by a *controlled quantum operation* $\Gamma(u, y)$. This can include the effects of measurement and an external environment. For fixed u, y it is a completely positive map (preserves positivity) and satisfies the normalization condition

$$\sum_{y \in \mathbf{Y}} \langle \Gamma(u, y)\omega, I \rangle = \langle \omega, I \rangle = 1 \tag{151}$$

The quantity

$$p(y|u, \omega) = \langle \Gamma(u, y)\omega, I \rangle. \tag{152}$$

is a conditional probability distribution on the measurement values (depends on current input value and state).

5. The performance objective $J(K)$ can be expressed in terms of the conditional states. The objective is to be minimized to obtain the desired optimal controller K^*.

5.3 Conditional Dynamics

5.3.1 Controlled State Transfer

The idea is that if the quantum system is in state ω_k at time k, and at this time the control value u_k is applied, a measurement outcome y_{k+1} will be recorded, and the system will transfer to a new state ω_{k+1}. The probability of y_{k+1} is $p(y_{k+1}|u_k, \omega_k)$.

Selective or conditional evolution means that the new state ω_{k+1} depends on the value of the measurement y_{k+1}, and we write this dependance as follows:

$$\omega_{k+1} = \Lambda_\Gamma(u_k, y_{k+1})\omega_k, \tag{153}$$

where

$$\Lambda_\Gamma(u, y)\omega = \frac{\Gamma(u, y)\omega}{p(y|u, \omega)}. \tag{154}$$

Equation (153) is a discrete time *stochastic master equation (SME)*, or *quantum filtering equation*.

Example 5.1 Consider the operator $\Gamma(u, y)$ is given by

$$\Gamma(u, y)\omega = \sum_{a,b} q(y|a) P_a E_b^u \omega E_b^{u\dagger} P_a \tag{155}$$

where ω is a state, and the adjoint is given by

$$\Gamma^\dagger(u,y)B = \sum_{a,b} q(y|a) E_b^{u\dagger} P_a B P_a E_b^u \tag{156}$$

where B is an observable. Here:

1. E_b^u are controlled operators that satisfy $\sum_b E_b^{u\dagger} E_b^u = I$. They can be used to model the influence of an environment on the system via

$$\mathcal{E}^u \omega = \sum_b E_b^u \omega E_b^{u\dagger} \tag{157}$$

   The simplest case is $E_b^u = T^u$ for all b, which corresponds simply to closed unitary evolution with no environmental influence:

$$\mathcal{E}^u \omega = T^{u\dagger} \omega T^u.$$

2. An observable A is measured. It is assumed to have a discrete non-degenerate spectrum $\{a\}$. The normalized eigenvectors are $|a\rangle$, with projections $P_a = |a\rangle\langle a|$ ($P_a|\psi\rangle = \langle a|\psi\rangle|a\rangle$).

3. The measurement of A is imperfect; instead values y are measured with probabilities $q(y|a)$. The kernels satisfy $\sum_y q(y|a) = 1$ for all a.

□

Example 5.2 (Two-state system.) We now describe a specific instance of Example 5.1, viz. a two-state system and measurement device, where it is desired to use feedback control to put the system into a given state.

A particle beam is passed through a Stern-Gerlach device, which results in one beam of particles in the up state, and one beam in the down state. It is desired to put all particles into the up state. In the absence of measurement noise, the following simple feedback scheme [52] achieves this objective: the beam of particles in the up state is subsequently left alone, while the beam in the down state is subject to a further device which will result in a change of spin direction from down to up. The final outcome of this feedback arrangement is that all particles are in the up state.

We extend this example by accommodating repeated noisy measurements. Physically, the noisy measurements might arise from imperfectly separated beams, where a proportion of each beam contaminates the other, and/or from interference or noise affecting sensors.

The pure states of the system are of the form

$$|\psi\rangle = c_{-1}|-1\rangle + c_1|1\rangle \;\equiv\; \begin{pmatrix} c_{-1} \\ c_1 \end{pmatrix}.$$

The states $|-1\rangle$ and $|1\rangle$ are eigenstates of the observable

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{158}$$

corresponding to ideal measurement values $a = -1$ and $a = 1$. It is desired to put the system into the state

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ or } |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

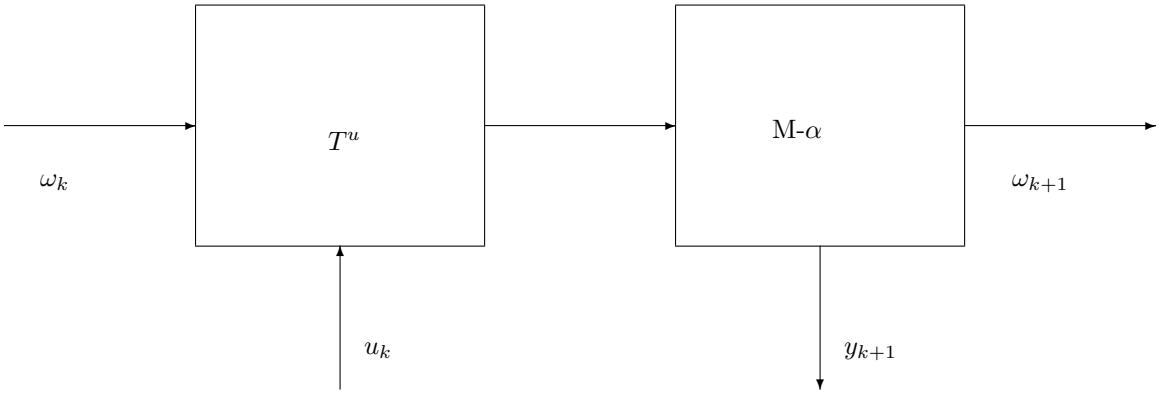

Figure 11: Two-state system example showing the controlled unitary operator T^u and the noisy measurement device M-α with error probability α.

In this example we do not consider the effects on an external environment. We define a controlled transfer operator $\Gamma(u, y)$ as the following physical process, Figure 11. First apply a unitary transformation T^u, where the control value $u = 0$ means do nothing, while $u = 1$ means to flip the states (quantum not gate), i.e.

$$T^u = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } u = 0 \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & \text{if } u = 1. \end{cases}$$

We then make an imperfect measurement corresponding to the observable A. We model this by an ideal device (e.g. Stern-Gerlach) with projection operators

$$P_{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; P_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

followed by a memoryless channel with error probability kernels

$$\begin{aligned} q(-1|-1) &= 1-\alpha \\ q(-1|1) &= \alpha \\ q(1|-1) &= \alpha \\ q(1|1) &= 1-\alpha \end{aligned}$$

where $0 \leq \alpha \leq 1$ is the probability of a measurement error (cf. [42, Figure 8.1]).

The controlled transfer operator is therefore (from (155))

$$\Gamma(u,y)\omega = q(y|-1)P_{-1}T^u\omega T^{u\,\dagger}P_{-1} + q(y|1)P_1T^u\omega T^{u\,\dagger}P_1.$$

In this example, the control u can take the values 0 or 1, and output y has values -1 or 1 ($\mathbf{U} = \{0,1\}$), $\mathbf{Y} = \{-1,1\}$).

If we write a general density matrix as

$$\omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12}^* & \omega_{22} \end{pmatrix}, \tag{159}$$

then the controlled operators $\Gamma(u,y)$ are given explicitly by

$$\begin{aligned}
\Gamma(0,-1)\omega &= \begin{pmatrix} (1-\alpha)\omega_{11} & 0 \\ 0 & \alpha\omega_{22} \end{pmatrix} \\
\Gamma(0,1)\omega &= \begin{pmatrix} \alpha\omega_{11} & 0 \\ 0 & (1-\alpha)\omega_{22} \end{pmatrix} \\
\Gamma(1,-1)\omega &= \begin{pmatrix} (1-\alpha)\omega_{22} & 0 \\ 0 & \alpha\omega_{11} \end{pmatrix} \\
\Gamma(1,1)\omega &= \begin{pmatrix} \alpha\omega_{22} & 0 \\ 0 & (1-\alpha)\omega_{11} \end{pmatrix}
\end{aligned}$$

□

5.3.2 Feedback Control

In the above description of the quantum system (153), we have not described how the controls u_k are determined by the measurements y_k via a feedback controller K. We now do this.

Feedback controllers should be *causal*, i.e., the current control value u_k cannot depend on future values of the measurements $y_{k+1}, y_{k+2}, \ldots$. On a time interval $0 \leq k \leq M-1$ this is expressed as follows:

$$K = \{K_0, K_1, \ldots, K_{M-1}\}$$

where

$$\begin{aligned}
u_0 &= K_0 \\
u_1 &= K_1(y_1) \\
u_2 &= K_2(y_1, y_2) \\
&\text{etc.}
\end{aligned}$$

To simplify notation, we often write sequences $u_{k_1}, u_{k_1+1}, \ldots, u_{k_2}$ as u_{k_1,k_2}. Then we can write $u_k = K_k(y_{1,k})$. A controller K can be restricted to subintervals $k \leq j \leq M$ by fixing (or omitting) the first arguments in the obvious way. We denote by $\mathcal{K}$ the class of all such feedback controllers.

ω_0	p_1	ω_1	p_2	ω_2
ω_0	$\frac{1}{2}$	$\omega_1^{(0,-1)}$	$2\alpha(1-\alpha)$	$\omega_2^{(0,-1),(1,-1)}$
			$\alpha^2+(1-\alpha)^2$	$\omega_2^{(0,-1),(1,1)}$
	$\frac{1}{2}$	$\omega_1^{(0,1)}$	$2\alpha(1-\alpha)$	$\omega_2^{(0,1),(0,-1)}$
			$\alpha^2+(1-\alpha)^2$	$\omega_2^{(0,1),(0,1)}$
$\rho_0=\omega_0$		ρ_1		ρ_2

Table 1: State evolution under the controller $\bar{K}$.

A feedback controller K in closed loop with the quantum system, Figure 10, operates as follows. The given initial state ω_0 and controller K are sufficient to define random sequences of states $\omega_{0,M}$, inputs $u_{0,M-1}$ and outputs $y_{1,M}$ over a given time interval $0 \leq k \leq M$ iteratively as follows. The control value u_0 is determined by K_0 (no observations are involved yet), and it is applied to the quantum system, which responds by selecting y_1 at random according to the distribution $p(y_1|u_0,\omega_0)$. This then determines the next state ω_1 via (153). Next u_1 is given by $K_1(y_1)$, and applied to the system. This process is repeated until the final time.

The controller K therefore determines controlled stochastic processes ω_k, u_k and y_k on the interval $0 \leq k \leq M$. Expectation with respect to the associated probability distribution is denoted $\mathbf{E}^K_{\omega_0,0}$. The state sequence ω_k is a *controlled Markov process*.

One way a controller K can be constructed is using a function

$$u_k = \mathbf{u}(\omega_k, k)$$

where ω_k is given by (153) with initial state ω_0. This controller is denoted $K^{\mathbf{u}}_{\omega_0}$. The SME equation (153) forms part of this controller, viz. its dynamics, and must be implemented with suitable technology (e.g. digital computer). Controllers of this type are said to have a *separation structure*, where the controller can be decomposed into an estimation part (i.e. filtering via (153)) and a control part (i.e. the function $\mathbf{u}$). The separation structure arises naturally from the dynamic programming techniques, as we shall see.

Example 5.3 (Two-state system with feedback, Example 5.2 continued.) We consider a particular feedback controller $\bar{K}$ for a time horizon $M = 2$ defined by

$$u_0 = \bar{K}_0 = 0, \quad u_1 = \bar{K}_1(y_1) = \begin{cases} 0 & \text{if } y_1 = 1 \\ 1 & \text{if } y_1 = -1. \end{cases} \tag{160}$$

We apply $\bar{K}$ to the system with initial pure state

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}|-1\rangle + \frac{1}{\sqrt{2}}|1\rangle, \text{ or } \omega_0 = \tfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{161}$$

The result is shown in Table 1, which displays the resulting conditional states

$$\begin{aligned} \omega_1^{(u_0,y_1)} &= \Lambda_\Gamma(u_0,y_1)\omega_0, \\ \omega_2^{(u_0,y_1),(u_1,y_2)} &= \Lambda_\Gamma(u_1,y_2)\omega_1^{(u_0,y_1)} \end{aligned}$$

and the associated probabilities. Explicitly, the terms shown in Table 1 are:

$$p_1 = p(y_1|u_0, \omega_0), \quad p_2 = p(y_2|u_1, \omega_1)$$

$$\omega_1^{(0,-1)} = \begin{pmatrix} (1-\alpha) & 0 \\ 0 & \alpha \end{pmatrix}, \quad \omega_1^{(0,1)} = \begin{pmatrix} \alpha & 0 \\ 0 & (1-\alpha) \end{pmatrix}$$

$$\omega_2^{(0,-1),(1,-1)} = \omega_2^{(0,1),(0,-1)} = \tfrac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{aligned} \omega_2^{(0,-1),(1,1)} &= \omega_2^{(0,1),(0,1)} \\ &= \frac{1}{\alpha^2 + (1-\alpha)^2} \begin{pmatrix} \alpha^2 & 0 \\ 0 & (1-\alpha)^2 \end{pmatrix}. \end{aligned}$$

Also shown are the non-selective states:

$$\rho_0 = \omega_0$$

$$\begin{aligned} \rho_1 &= p(-1|u_0, \omega_0)\omega_1^{(0,-1)} + p(1|u_0, \omega_0)\omega_1^{(0,1)} \\ &= \tfrac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \rho_2 = &\, p(-1|1, \omega_1^{(0,-1)})\omega_2^{(0,-1),(1,-1)} \\ &+ p(1|1, \omega_1^{(0,-1)})\omega_2^{(0,-1),(1,1)} \\ &+ p(-1|0, \omega_1^{(0,1)})\omega_2^{(0,1),(0,-1)} \\ &+ p(1|0, \omega_1^{(0,1)})\omega_2^{(0,1),(0,1)} \\ = &\, \tfrac{1}{2} \begin{pmatrix} \alpha^2 + \alpha(1-\alpha) & 0 \\ 0 & \alpha(1-\alpha) + (1-\alpha)^2 \end{pmatrix}. \end{aligned} \tag{162}$$

At time $k = 0$ the control $u = 0$ is applied. If $y_1 = -1$ is observed, as a result of the imperfect measurement, the system moves to the state $\omega_1^{(0,-1)}$. Since $y_1 = -1$, the controller $\bar{K}$ (160) gives $u_1 = 1$. This results in the states $\omega_2^{(0,-1),(1,-1)}$ or $\omega_2^{(0,-1),(1,1)}$, depending on the outcome of the second measurement y_2. If, on the other hand, $y_1 = 1$ is observed, the system moves to the state $\omega_1^{(0,1)}$. Since $y_1 = 1$, the controller $\bar{K}$ (160) gives $u_1 = 0$, and hence $\omega_2^{(0,1),(0,-1)}$ or $\omega_2^{(0,1),(0,1)}$, again depending on the outcome of the second measurement y_2. This is illustrated in Figure 12.

Note that when $\alpha = 0$ (perfect measurements), the feedback system terminates in the desired pure state $\rho_2 = |1\rangle\langle 1|$, as discussed above. The role of feedback control is clearly demonstrated here. With imperfect measurements, $0 < \alpha < 1$, the system terminates in the mixed state ρ_2 given by (162), with the degree of mixing (indicating the expected degradation in performance) depending on the measurement error probability parameter α:

$$\begin{aligned} \mathrm{tr}\rho_2^2 &= (\alpha^2 + \alpha(1-\alpha))^2 + (\alpha(1-\alpha) + (1-\alpha)^2)^2 \\ &< 1 \ \text{ if } 0 < \alpha < 1 \\ &= 1 \text{ if } \alpha = 0. \end{aligned}$$

□

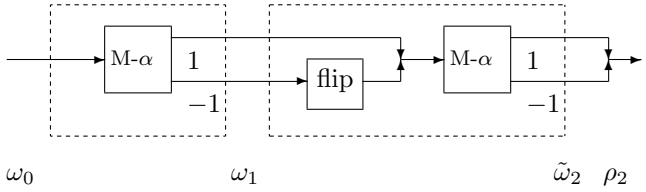

Figure 12: Physical realization of the two stages of the two-state system with feedback using controller $\bar{K}$. Due to the merging of the beams in the second stage, we have the intermediate state $\tilde{\omega}_2 = \frac{1}{2}\omega_2^{(0,-1),(1,-1)} + \frac{1}{2}\omega_2^{(0,1),(0,-1)}$ if $y_2 = -1$ (with probability $2\alpha(1-\alpha)$), or $\tilde{\omega}_2 = \frac{1}{2}\omega_2^{(0,-1),(1,1)} + \frac{1}{2}\omega_2^{(0,1),(0,1)}$ if $y_2 = 1$ (with probability $\alpha^2 + (1-\alpha)^2$).

5.4 Optimal Control

In this section we summarize dynamic programming results for a well-known type of finite time horizon optimal control problem, [5, 38]. The optimal control problem discussed here can be considered to be a prototype problem illustrating measurement feedback in the quantum context. The dynamic programming methods used in this paper for solving the optimal control problems are standard.

We define a *cost function* to be a non-negative observable $L(u)$ that can depend on the control u. The cost function encodes the designer's control objective. We also use a non-negative observable N to define a cost for the final state.

Example 5.4 (Two-state system with feedback, Example 5.3 continued.) To set up the cost function $L(u)$ to reflect our objective of regulating the system to the desired pure state $|1\rangle$, we define

$$X = \frac{1}{2}(A - 1.I) = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

where A is the observable corresponding to the projective measurement (158). We note that the expected value of X^2 is

$$\begin{aligned} \langle 1|X^2|1\rangle &= \mathrm{tr}[X^2|1\rangle\langle 1|] &&= 0 \\ \langle -1|X^2|-1\rangle &= \mathrm{tr}[X^2|-1\rangle\langle -1|] &&= 1 \end{aligned}$$

which gives zero cost to the desired state, and nonzero cost to the undesired state. We shall also introduce a cost of control action, as follows:

$$c(u) = \begin{cases} 0 & \text{if } u = 0 \\ p & \text{if } u = 1 \end{cases}$$

where $p > 0$. This gives zero cost for doing nothing, and a nonzero cost for the flip operation. Thus we define the cost function to be

$$L(u) = X^2 + c(u)I \tag{163}$$

and the cost for the final state is defined to be

$$N = X^2.$$

This modifies our earlier objective of putting the system into the desired state by including a penalty for control action. □

Let $M > 0$ be a positive integer indicating a finite time interval $k = 0, \ldots, M$. Given a sequence of control values $u_{0,M-1} = u_0, \ldots, u_{M-1}$ and measurements $y_{1,M} = y_1, \ldots, y_M$, define the *risk-neutral cost functional*

$$J_{\omega,0}(K) = \mathbf{E}^K_{\omega,0}[\sum_{i=0}^{M-1} \langle \omega_i, L(u_i)\rangle + \langle \omega_M, N\rangle], \tag{164}$$

where ω_i, $i = 0, \ldots, M$ is the solution of the system dynamics (153) with initial state $\omega_0 = \omega$ under the action of a controller K. This is an appropriate quantum generalization of the classical LQG type cost. The objective is to minimize this functional over all measurement feedback controllers $K \in \mathcal{K}$.

Following [5] it is convenient to rewrite the cost functional (164). For each k, given a sequence of control values $u_{k,M-1} = u_k, \ldots, u_{M-1}$ and measurements $y_{k+1,M} = y_{k+1}, \ldots, y_M$, define a random sequence of observables Q_k by the recursion ([5, equation (3.1)])

$$\begin{aligned} Q_k &= \Gamma^\dagger(u_k, y_{k+1})Q_{k+1} + L(u_k), \quad 0 \le k \le M-1 \\ Q_M &= N \end{aligned} \tag{165}$$

When useful, we write

$$Q_k = Q_k(u_{k,M-1}, y_{k+1,M})$$

to indicate dependence on the input and outputs. Q_k may be called a *cost observable*. The cost functional (164) is given by

$$J_{\omega,0}(K) = \sum_{y_{1,M} \in \mathbf{Y}^M} \langle \omega, Q_0(K(y_{1,M})_{0,M-1}, y_{1,M})\rangle \tag{166}$$

Here and elsewhere we use abbreviations of the form

$$K(y_{1,M})_{0,M-1} = (K_0, K_1(y_1), \ldots, K_{M-1}(y_{1,M-1}))$$

Remark 5.5 The cost observable Q_k given by (165) and the expression in (166) is analogous to the familiar Heisenberg picture used in quantum physics. It is very natural from the point of view of dynamic programming, and indeed (164) and (166) are related by iterating (165). Here is the first step:

$$\begin{aligned} \langle \omega_0, Q_0\rangle &= \langle \omega_0, \Gamma^\dagger(u_0, y_1)Q_1 + L(u_0)\rangle \\ &= \langle \omega_0, L(u_0)\rangle + \langle \Gamma(u_0, y_1)\omega_0, Q_1\rangle \\ &= \langle \omega_0, L(u_0)\rangle + \langle \omega_1, Q_1\rangle p(y_1|u_0, \omega_0) \end{aligned}$$

where $\omega_1 = \Lambda_\Gamma(u_0, y_1)\omega_0$ and $p(y_1|u_0, \omega_0)$ is given by (152). □

The key idea of dynamic programming is to look at the current state at a current time $0 \leq k \leq M-1$ and to optimize the remaining cost from the current time to the final time. This leads to an iterative solution. Accordingly, we define, for each $0 \leq k \leq M$, the cost to go incurred by a controller K (restricted to $k \leq l \leq M-1$) to be

$$J_{\omega,k}(K) = \sum_{y_{k+1,M} \in \mathbf{Y}^{M-k}} \langle \omega, Q_k(K(y_{k+1,M})_{k,M-1}, y_{k+1,M}) \rangle \tag{167}$$

The dynamic programming equation associated with this risk-neutral problem is

$$\begin{aligned} V(\omega,k) &= \inf_{u \in \mathbf{U}} \{ \langle \omega, L(u) \rangle + \sum_{y \in \mathbf{Y}} V(\Lambda_\Gamma(u,y)\omega, k+1) p(y|u,\omega) \}, \\ V(\omega,M) &= \langle \omega, N \rangle \end{aligned} \tag{168}$$

where $0 \leq k \leq M-1$. This is the fundamental equation from which optimality or otherwise of a controller can be determined.

Let V be the solution to the dynamic programming equation (168). Then for any controller $K \in \mathcal{K}$ we have

$$V(\omega,k) \leq J_{\omega,k}(K). \tag{169}$$

If we assume in addition that a minimizer

$$\mathbf{u}^*(\omega,k) \in \operatorname*{argmin}_{u \in \mathbf{U}} \{ \langle \omega, L(u) \rangle + \sum_{y \in \mathbf{Y}} V(\Lambda_\Gamma(u,y)\omega, k+1) p(y|u,\omega)) \} \tag{170}$$

exists[3] for all ω, $0 \leq k \leq M-1$, then the separation structure controller $K^{\mathbf{u}^*}_{\omega_0}$ defined by (170) is optimal, i.e.

$$J_{\omega_0,0}(K^{\mathbf{u}^*}_{\omega_0}) = V(\omega_0,0) \leq J_{\omega_0,0}(K) \tag{171}$$

for all $K \in \mathcal{K}$.

Example 5.6 (Two-state system with feedback, Example 5.4 continued.) We solve the dynamic programming equation (168) and determine the optimal feedback controls as follows. For $k = M = 2$ we have

$$V(\omega,2) = \langle \omega, X^2 \rangle = \omega_{11}$$

and hence for $k = 1$

$$V(\omega,1) = \omega_{11} + \min[V_0(\omega,1), V_1(\omega,1)]$$

where where $V_0(\omega,1), V_1(\omega,1)$ are given in Appendix 5.5. Hence we obtain

$$\mathbf{u}^*(\omega,1) = \begin{cases} 0 & \text{if } V_0(\omega,1) \leq V_1(\omega,1) \\ 1 & \text{if } V_0(\omega,1) > V_1(\omega,1). \end{cases}$$

At time $k = 0$ we have

$$V(\omega,0) = \omega_{11} + \min[V_0(\omega,0), V_1(\omega,0)]$$

[3]The notation $\operatorname{argmin}_{u \in \mathbf{U}} f(u)$ means the subset of values from $\mathbf{U}$ minimizing f.

where $V_0(\omega,0), V_1(\omega,0)$ are given in Appendix 5.5, which gives

$$\mathbf{u}^*(\omega,0) = \begin{cases} 0 & \text{if } V_0(\omega,0) \leq V_1(\omega,0) \\ 1 & \text{if } V_0(\omega,0) > V_1(\omega,0). \end{cases}$$

The optimal risk-neutral feedback controller is given by

$$u_0 = K^{\mathbf{u}^*}_{\omega_0,0} = \mathbf{u}^*(\omega_0,0), \quad u_1 = K^{\mathbf{u}^*}_{\omega_0,1}(y_1) = \mathbf{u}^*(\omega_1,1)$$

where $\omega_1 = \Lambda_\Gamma(u_0,y_1)\omega_0$. *Note that the control u_1 depends on y_1 through the conditional state ω_1 (separation structure).* A physical implementation of the quantum system with optimal risk-neutral feedback is shown in Figure 13.

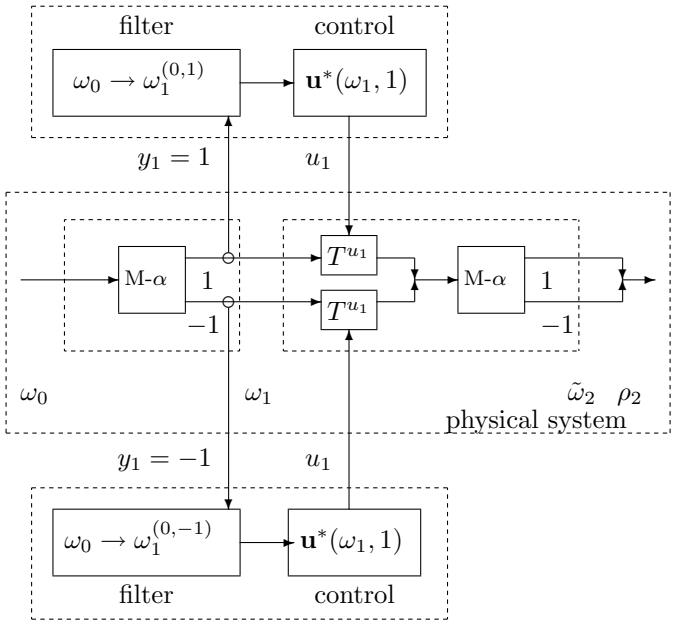

Figure 13: Physical realization of the two stages of the two-state system with feedback using the optimal risk-neutral controller $K^{\mathbf{u}^*}_{\omega_0}$ (with ω_0 given by (161), we have $u_0 = \mathbf{u}^*(\omega_0,0) = 0$, $u_1 = \mathbf{u}^*(\omega_1,1)$).

Let's consider the special case $\alpha = 0$ and $p = 0$, with initial state (161). We then find that $V_0(\omega_0,0) = V_1(\omega_0,0) = 0.5$, and hence we take $\mathbf{u}^*(\omega_0,0) = 0$; i.e. $u_0 = 0$.

Next, if $y_1 = -1$ is observed, we have $\omega_1 = |-1\rangle\langle -1|$, $V_0(\omega_1,1) = 1$ and $V_1(\omega_1,1) = 0$. Hence we take $\mathbf{u}^*(\omega_1,1) = 1$, i.e. $u_1 = 1$. However, if $y_1 = 1$ is observed, we have $\omega_1 = |1\rangle\langle 1|$, $V_0(\omega_1,1) = 0$ and $V_1(\omega_1,1) = 1$; and hence we take $\mathbf{u}^*(\omega_1,1) = 0$, i.e. $u_1 = 0$. In either case we achieve the desired state $\rho_2 = \omega_2 = |1\rangle\langle 1|$.

This action is the same as that seen before for the controller $\bar{K}$. The same controller is obtained for $0 < \alpha < 0.5$ and $p = 0$, but ω_2 will be a mixed state. If $p \neq 0$ the optimal controller $K^{\mathbf{u}^*}_{\omega_0}$ will result in control actions that in general differ from those of $\bar{K}$. $\square$

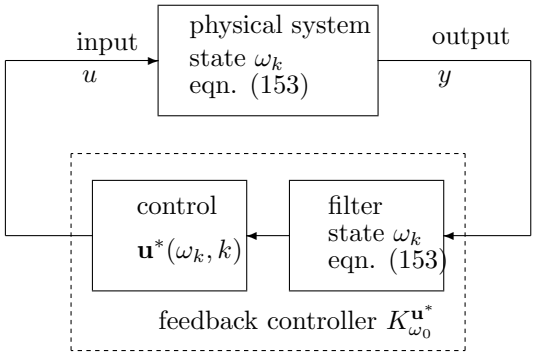

Figure 14: Optimal risk-neutral controller $K^{\mathbf{u}^*}_{\omega_0}$ showing separation structure and states of the physical system ω_k and filter ω_k.

5.5 Appendix: Formulas for the Two-State System with Feedback Example

The following quantities were used in the solution of the risk-neutral problem, Example 5.6:

$$V_0(\omega, 1) = \omega_{11}, \quad V_1(\omega, 1) = \omega_{22} + p$$

$$\begin{aligned} V_0(\omega, 0) = \omega_{11} &+ \min[\alpha\omega_{11}, p + \omega_{22} - \alpha\omega_{22}] \\ &+ \min[\omega_{11} - \alpha\omega_{11}, p + \alpha\omega_{22}] \end{aligned}$$

$$\begin{aligned} V_1(\omega, 0) = p + \alpha\omega_{11} + \omega_{22} - \alpha\omega_{22} \\ + \min[\alpha\omega_{11}, p + \omega_{22} - \alpha\omega_{22}] \\ + \min[p + \alpha\omega_{11}, \omega_{22} - \alpha\omega_{22}] \end{aligned}$$

6 Optimal Risk-Sensitive Feedback Control of Quantum Systems

References for this section include [36], [7], [12], [41], [43], [34], [13].

6.1 System Model

We consider the problem of controlling an open quantum system S (e.g. an atom) that interacts with an electromagnetic field B. We suppose that the evolution can be influenced by control variables u that enter the system Hamiltonian $H(u)$. The field B is continuously monitored, providing weak measurements of the system S, the results y of which are available to the controller $\mathbf{K}$, which processes this information to produce the control actions u. The control in general is allowed to be a causal function of the measurement trajectory. The problem is to find a controller $\mathbf{K}$ so that it minimizes a risk-sensitive cost function $J^\mu(\mathbf{K})$, which will be defined below.

In order to describe the quantum model for this problem, we need to consider how system variables evolve with time. System variables (e.g. position, momentum) are operators X defined on the Hilbert space underlying the system S. The evolution is determined by interaction with the electromagnetic field, the influence of which is modelled by quantum white noise. The quantum equations analogous to classical models (e.g. (108), (109)) are

$$\begin{aligned} dX(t) \;=\; & (-X(t)K(t) - K^\dagger(t)X(t) + M^\dagger(t)X(t)M(t))dt \\ & +[X(t), M(t)]dB^\dagger(t) - [X(t), M^\dagger(t)]dB(t) \end{aligned}$$

and

$$dY(t) \;=\; (M(t) + M^\dagger(t))dt + dQ(t). \tag{172}$$

We now explain these equations which use the framework of quantum stochastic differential equations, [28], [43].

Let $u = u(\cdot)$ be a control signal (a function of time t with values $u(t) \in \mathbf{U}$). The quantum model we use considers the system plus field as a total closed system with unitary operator $U(t) = U^u(t)$ (interaction picture), which solves the quantum stochastic differential equation (QSDE) [28, eq. (11.2.7)], [43, sec. 26],

$$dU(t) = \{-K(u(t))dt + MdB^\dagger(t) - M^\dagger dB(t)\}U(t) \tag{173}$$

with initial condition $U(0) = I$, where

$$K(u) = \frac{i}{\hbar}H(u) + \frac{1}{2}M^\dagger M.$$

Here, M is a system operator which together with the field operator $b(t) = \dot{B}(t)$, model the interaction of the system with the channel. We denote adjoints with the symbol †. Note that equation (173) is written in Ito form (see, e.g. [27, Chapter 4]), as will all

stochastic differential equations in these notes. With vacuum initialization of the field channel, the non-zero Ito product is, [28, eq. (11.2.6)], $dB(t)dB^\dagger(t) = dt$. Then system operators X evolve according to

$$X(t) = j_t(u, X) = U^\dagger(t)XU(t),$$

and satisfy the quantum Langevin equation (QLE) (172), where $M(t) = j_t(u, M)$, and $K(t) = j_t(u, K(u(t)))$.

Let v denote the vacuum state for the field, and let π_0 be an arbitrary initial system state. The initial state of the total system is $\rho_0 = \pi_0 \otimes vv^\dagger$. We write, for operators A and B,

$$\langle A, B\rangle = \mathrm{tr}[A^\dagger B].$$

We regard the field operator $B(t)$ as an input field [28, Section 11.3.2], with corresponding output field $A(t) = j_t(u, B(t))$. The self adjoint real quadratures of the input field is defined by $Q(t) = B(t) + B^\dagger(t)$. For the output field real quadrature we write $Y(t) = j_t(u, Q(t))$, also self adjoint. This process satisfies the QSDE (172). The input quadrature is self commutative $[Q(t), Q(s)] = 0$ for all s, t, and corresponds to quantum Brownian motion when the field is initialized in the vacuum state. Indeed, if Ω_T denotes the set of all Wiener paths, the probability of a subset $F \subset \Omega_T$ of paths is $\mathbf{P}^0(F) = \langle vv^\dagger, P_T^Q(F)\rangle$, where $P_T^Q(F)$ is the projection operator associated with F and $Q(s)$, $0 \le s \le T$. The probability distribution $\mathbf{P}^0$ is the Wiener distribution, under which the increments $q(t) - q(s)$, $0 \le s \le t \le T$, are independent, Gaussian, with zero mean and covariance $(t - s)$. The output fields $Y(t)$ are also self commutative, and define a probability measure $\mathbf{P}$ by $\mathbf{P}(F) = \langle \rho_0, P_T^Y(F)\rangle$, where $P_T^Y(F)$ is the corresponding projection operator.

The input and output fields satisfy the non-demolition conditions $[Q(t), X] = 0$, for all $t \ge 0$, and $[Y(s), X(t)] = 0$ for all $s \le t$. This means that we can continuously monitor the output field $Y(t)$ without demolishing the system, say by homodyne detection. The results of the measurement is a real record $y(\cdot) \in \Omega_T$, which is used by a (classical) controller $\mathbf{K}$ to produce the input control signal $u(t)$ by $u(t) = \mathbf{K}(t, y_{[0,t]})$. The notation used here is meant to indicate the causal dependence of the control on the measurements; $y_{[0,t]}$ indicates the segment of the measurement signal on the time interval $[0, t]$, so in effect the controller $\mathbf{K} = \{\mathbf{K}(t, \cdot)\}$ is a family of functions. The measurement record is given by the SDE

$$dy(t) = \mathrm{tr}[(M + M^\dagger)\pi_t]dt + dw(t), \tag{174}$$

where $w(t)$ is a Wiener process under $\mathbf{P}$ and

$$d\pi_t = -\frac{i}{\hbar}[H(u(t)), \pi_t]dt + \mathcal{D}[M]\pi_t dt \quad + \mathcal{H}[M]\pi_t dw(t), \tag{175}$$

where $\mathcal{D}[c]\rho = c\rho c^\dagger - \frac{1}{2}(c^\dagger c\rho + \rho c^\dagger c)$, and $\mathcal{H}[c]\rho = c\rho + \rho c^\dagger - \rho\mathrm{tr}(c\rho + \rho c^\dagger)$. Equation (175) stochastic master equation or *Belavkin quantum filtering equation* (e.g. [34, Chapter 5.2.5]).

In unnormalised form, the Belavkin filter is

$$d\sigma_t = (-K(u(t))\sigma_t - \sigma_t K^\dagger(u(t)) + M\sigma_t M^\dagger)dt + (M\sigma_t + \sigma_t M^\dagger)dy(t). \tag{176}$$

Here, y is a standard Wiener process under the reference distribution $\mathbf{P}^0$. The normalized conditional state an be recovered by

$$\pi_t = \frac{\sigma_t}{\langle \sigma_t, 1\rangle}. \tag{177}$$

6.2 Risk-Neutral Optimal Control

In this section we briefly discuss a risk-neutral problem of the type that has been studied by [5], [6], [18], [13].

Let $C_1(u)$ be a non-negative self-adjoint system operator depending on the control value u, and let C_2 be a non-negative self-adjoint system operator. These so-called *cost operators* are chosen to reflect the performance objectives, and explicitly include the control so that a balance between performance and control cost can be achieved. The quantity

$$\int_0^T C_1(t)dt + C_2(T), \tag{178}$$

where $C_1(t) = j_t(u, C_1(u(t)))$, $C_2(t) = j_t(u, C_2)$, accumulates cost over the given time interval and provides a penalty for the final time. The *risk-neutral* problem defined by the quantum expectation

$$J(\mathbf{K}) = \langle \rho_0, \int_0^T C_1(t)dt + C_2(T)\rangle, \tag{179}$$

where $\rho_0 = \pi_0 \otimes v \otimes v^\dagger$.

The key step in solving the optimal control problem specified by (179) is a stochastic representation followed by classical conditional expectation, which results in (recall section 3.5.1)

$$\begin{aligned} J(\mathbf{K}) &= \mathbf{E}[\int_0^T \langle \pi_t, C_1(u(t))\rangle dt + \langle \pi_T, C_2\rangle] \\ &= \mathbf{E}^0[\int_0^T \langle \sigma_t, C_1(u(t))\rangle dt + \langle \sigma_T, C_2\rangle] \end{aligned} \tag{180}$$

where π_t and σ_t are the conditional states, assuming interchanges of expectations and integrals are justified.

Dynamic programming. The risk-neutral value function is defined by

$$W(\sigma, t) = \inf_{\mathbf{K}} \mathbf{E}^0_{\sigma,t}[\int_t^T \langle \sigma_s, C_1\rangle ds + \langle \sigma_T, C_2\rangle] \tag{181}$$

and the corresponding dynamic programming equation reads

$$\begin{aligned} \tfrac{\partial}{\partial t}W(\sigma,t) + \inf_{u\in\mathbf{U}}\{\mathcal{L}^u W(\sigma,t) + C_1(u)\} &= 0, \quad 0 \le t < T, \\ W(\sigma,T) &= \langle \sigma, C_2\rangle. \end{aligned} \tag{182}$$

We explain the meaning of the operator $\mathcal{L}^u$ following classical stochastic control methods [23]. For a fixed constant control value u (i.e. $u(t) = u \in \mathbf{U}$ for all t), σ_t is a Markov process with generator $\mathcal{L}^u$, which is defined, when it exists, by

$$\mathcal{L}^u f(\sigma) = \lim_{t\downarrow 0} \frac{\mathbf{E}^0_{\sigma,0}[f(\sigma_t)] - f(\sigma)}{t} \tag{183}$$

for suitably smooth functions $f(\cdot)$. In fact, $\mathcal{L}^u f(\sigma)$ can be calculated explicitly for f of the form

$$f(\sigma) = g(\langle \sigma, X_1\rangle, \ldots, \langle \sigma, X_J\rangle), \tag{184}$$

where g is a smooth bounded function of vectors of length J, and $X_1, \ldots, X_J$ are system operators. It is given by

$$\begin{aligned} &\mathcal{L}^u f(\sigma) = \\ &\frac{1}{2}\sum_{j,k=1}^{J} g_{jk}(\langle \sigma, X_1\rangle, \ldots, \langle \sigma, X_J\rangle)\langle \sigma, M^\dagger X_j + X_j M\rangle\langle \sigma, M^\dagger X_k + X_k M\rangle \\ &+ \sum_{j=1}^{J} g_j(\langle \sigma, X_1\rangle, \ldots, \langle \sigma, X_J\rangle)\langle \sigma, -K(u)X_j - X_j K(u) + M^\dagger X_j M\rangle \end{aligned} \tag{185}$$

for functions f of the form (184). Here g_j and g_{jk} denote first and second order partial derivatives of g.

If the dynamic programming equation (182) has a sufficiently smooth solution $W(\sigma,t)$, then the optimal controller $\mathbf{K}^\star$ is given by

$$\mathbf{K}^\star \;:\; \begin{aligned} d\sigma_t &= (-K(u(t))\sigma_t - \sigma_t K^\dagger(u(t)) + M\sigma_t M^\dagger)dt + (M\sigma_t + \sigma_t M^\dagger)dy(t) \\ u(t) &= \mathbf{u}^\star(\sigma^\mu_t, t). \end{aligned} \tag{186}$$

where $\mathbf{u}^\star(\sigma,t)$ attains the minimum in (182). The dynamical part of this controller is the Belavkin quantum filter, (176).

6.3 Risk-Sensitive Optimal Control

Instead of using the expected value of the quantity (178) as a cost function as in (179), we consider the average of the exponential of (178) in the following way, since we wish to generalize the LEQG cost (115). Define $R(t)$ to be the solution of the operator differential equation

$$\frac{dR(t)}{dt} = \frac{\mu}{2}C_1(t)R(t) \tag{187}$$

with initial condition $R(0) = I$. Here, $\mu > 0$ is a positive (risk) parameter. The solution of (187) can be expressed as the time-ordered exponential

$$R(t) = \overleftarrow{\exp}\left(\frac{\mu}{2}\int_0^t C_1(s)ds\right).$$

We then define the *risk-sensitive* cost function to be the quantum expectation

$$J^\mu(\mathbf{K}) = \langle \rho_0, R^\dagger(T)e^{\mu C_2(T)}R(T)\rangle. \tag{188}$$

Here, $\rho_0 = \pi_0 \otimes vv^\dagger$, as above.

It is shown in [36] that the quantum expectation can be replaced by an equivalent classical expectation, viz.

$$J^\mu(\mathbf{K}) = \mathbf{E}^0[\langle \sigma_T^\mu, e^{\mu C_2}\rangle], \tag{189}$$

where the unnormalized state σ_t^μ (a density operator, or information state) is the solution of the SDE

$$d\sigma_t^\mu \;=\; (-K^\mu(u(t))\sigma_t^\mu - \sigma_t^\mu K^{\mu\dagger}(u(t)) + M\sigma_t^\mu M^\dagger)dt + (M\sigma_t^\mu + \sigma_t^\mu M^\dagger)dy_2(t), \tag{190}$$

or

$$d\sigma_t^\mu \;=\; -\frac{i}{\hbar}[H(u(t)), \sigma_t^\mu]dt + \mathcal{D}[M]\sigma_t^\mu dt + \frac{\mu}{2}\tilde{\mathcal{H}}[C_1(u(t))]\sigma_t^\mu dt + \tilde{\mathcal{H}}[M]\sigma_t^\mu dy_2(t), \tag{191}$$

where $\tilde{\mathcal{H}}[c]\rho = c\rho + \rho c^\dagger$, and $K^\mu(u) = K(u) - \mu\frac{1}{2}C_1(u)$. The initial condition is $\sigma^\mu(0) = \pi_0$. The expression (189) is similar to the classical forms [8, eq. (3.4)], [37, eq. (2.10)]. Another useful representation is given in terms of the following normalized risk-sensitive state

$$\pi_t^\mu = \frac{\sigma_t^\mu}{\langle \sigma_t^\mu, 1\rangle},$$

namely

$$J^\mu(\mathbf{K}) = \mathbf{E}^\mu[\exp(\mu\int_0^T \mathrm{tr}(C_1(u(t))\pi_t^\mu)dt)\langle \pi_T^\mu, e^{\mu C_2}\rangle] \tag{192}$$

where $\mathbf{E}^\mu$ denotes expectation with respect to the probability distribution $\mathbf{P}^\mu$ defined by $d\mathbf{P}^\mu = \Lambda_T^\mu d\mathbf{P}^0$, where

$$\Lambda_T^\mu \;=\; \exp(-\tfrac{1}{2}\int_0^T |\mathrm{tr}[(M + M^\dagger)\pi_t^\mu]|^2 dt + \int_0^T \mathrm{tr}[(M + M^\dagger)\pi_t^\mu]dy(t)).$$

The SDE satisfied by π_t^μ is

$$d\pi_t^\mu \;=\; -\frac{i}{\hbar}[H(u(t)), \pi_t^\mu]dt + \mathcal{D}[M]\pi_t^\mu dt + \frac{\mu}{2}\mathcal{H}[C_1(u(t))]\pi_t^\mu dt + \mathcal{H}[M]\pi_t^\mu dw^\mu(t), \tag{193}$$

where $w^\mu(t)$ is a standard Wiener process with respect to $\mathbf{P}^\mu$ defined by

$$dy(t) = \mathrm{tr}[(M + M^\dagger)\pi_t^\mu]dt + dw^\mu(t). \tag{194}$$

Dynamic programming. Define the risk-sensitive value function $S^\mu(\sigma,t)$ for an arbitrary initial unnormalized state σ and initial time $0 \le t \le T$ by

$$S^\mu(\sigma,t) = \inf_{\mathbf{K}} \mathbf{E}^0_{\sigma,t}[\langle \sigma^\mu_T, e^{\mu C_2}\rangle], \tag{195}$$

where σ^μ_T denotes the solution of (190) at time T with initial condition $\sigma^\mu_t = \sigma$ (we have made explicit the dependence on the initial state and time in the expectation notation). Note that the cost (188) is given by

$$J^\mu(\mathbf{K}) = \mathbf{E}^0_{\pi_0,0}[\langle \sigma^\mu_T, e^{\mu C_2}\rangle]$$

so that the optimal controller $\mathbf{K}^{\mu,\star}$ is determined by

$$J^\mu(\mathbf{K}^{\mu,\star}) = S^\mu(\pi_0, 0).$$

The method of dynamic programming in this context relates the value function at time t and at a later time $t \le s \le T$ along optimal trajectories via the relation

$$S^\mu(\sigma,t) = \inf_{\mathbf{K}} \mathbf{E}^0_{\sigma,t}[S^\mu(\sigma^\mu_s, s)]. \tag{196}$$

This is the principle of optimality, [24, Chapter VI], [38, Chapter 6]. Note that by definition $S^\mu(\sigma,T) = \langle \sigma, e^{\mu C_2}\rangle$. The dynamic programming equation is

$$\begin{aligned} \tfrac{\partial}{\partial t} S^\mu(\sigma,t) + \inf_{u\in\mathbf{U}} \mathcal{L}^{\mu;u} S^\mu(\sigma,t) &= 0, \quad 0 \le t < T, \\ S^\mu(\sigma,T) &= \langle \sigma, e^{\mu C_2}\rangle. \end{aligned} \tag{197}$$

Note that in the dynamic programming PDE (197), the minimization is over the control values u, whereas in the definitions of the cost (188) and value function (195) the minimizations are over the controllers $\mathbf{K}$.

For fixed $u \in \mathbf{U}$ and for f of the form (184), we have

$$\begin{aligned} \mathcal{L}^{\mu;u} f(\sigma) = \frac{1}{2} \sum_{j,k=1}^J & g_{jk}(\langle\sigma,X_1\rangle,\ldots,\langle\sigma,X_J\rangle)\langle\sigma, M^\dagger X_j + X_j M\rangle\langle\sigma, M^\dagger X_k + X_k M\rangle \\ & + \sum_{j=1}^J g_j(\langle\sigma,X_1\rangle,\ldots,\langle\sigma,X_J\rangle)\langle\sigma, -K^\mu(u)X_j - X_j K^\mu(u) + M^\dagger X_j M\rangle \end{aligned} \tag{198}$$

If the dynamic programming PDE has a sufficiently smooth solution $S^\mu(\sigma,t)$, then the optimal controller $\mathbf{K}^{\mu,\star}$ can be obtained as follows. Let $\mathbf{u}^{\mu,\star}(\sigma,t)$ denote the control value that attains the minimum in (197) for each σ, t. The optimal controller is obtained by combining this function with the dynamics (190):

$$\mathbf{K}^{\mu,\star} \;:\; \begin{aligned} d\sigma^\mu_t &= (-K^\mu(u(t))\sigma^\mu_t - \sigma^\mu_t K^{\mu\dagger}(u(t)) + M\sigma^\mu_t M^\dagger)dt + (M\sigma^\mu_t + \sigma^\mu_t M^\dagger)dy(t) \\ u(t) &= \mathbf{u}^{\mu,*}(\sigma^\mu_t, t). \end{aligned} \tag{199}$$

6.4 Control of a Two Level Atom

In this section we consider the application of the risk-sensitive control problem to the example studied in [13], namely the feedback control of a two-level atom using a laser.

6.4.1 Setup

The amplitude and phase of the input laser can be adjusted, so via the interaction with the laser the atom can be controlled. The real quadrature of a second field channel is continuously monitored, say by homodyne detection, providing an indirect measurement of the atom. The control input is complex, $u = u_r + iu_i = |u|e^{i\arg_u} \in \mathbf{C}$ (the control field channel becomes a coherent state corresponding to u). The measurement signal $y_2(t)$ is real. It is desired to regulate the system in the σ_z up state $|\uparrow\rangle = (1,0)^T$ (the down state is $|\downarrow\rangle = (0,1)^T$, and $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are the Pauli matrices).

To model this example we use an additional unmeasured channel which interacts with the atom via an operator L. The corresponding risk-sensitive filter is [36]

$$\begin{aligned} d\sigma_t^\mu \;=\; & (-K^\mu(u(t))\sigma_t^\mu - \sigma_t^\mu K^{\mu\dagger}(u(t)) + L\sigma_t^\mu L^\dagger + M\sigma_t^\mu M^\dagger)dt \\ & +(M\sigma_t^\mu + \sigma_t^\mu M^\dagger)dy_2(t), \end{aligned} \tag{200}$$

where $K^\mu = \frac{i}{\hbar}H(u) + \frac{1}{2}M^\dagger M + \frac{1}{2}L^\dagger L - \frac{\mu}{2}C_1(u)$.

In terms of the notation used in these notes, we have

$$\begin{aligned} L \;=\; & \kappa_f\sigma_-, \quad M = \kappa_s\sigma_-, \quad H(u) = i(u^\star L - uL^\dagger), \\ & \kappa_f^2 + \kappa_s^2 = 1, \quad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ \mathbf{U} \;=\; & \mathbf{C}, \quad C_1(u) = a\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \tfrac{1}{2}b|u|^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ & C_2 = c\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad a \geq 0, b \geq 0, c \geq 0. \end{aligned}$$

Here, κ_f^2 and κ_s^2 are the decay rates into the control and measurement channels. The parameters a, b and c are weights for the components of the cost. Note that $\langle\downarrow|C_1(u)|\downarrow\rangle > 0$ and $\langle\downarrow|C_2|\downarrow\rangle > 0$ (if $a > 0$ and $c > 0$), while $\langle\uparrow|C_1(0)|\uparrow\rangle = 0$ and $\langle\uparrow|C_2|\uparrow\rangle = 0$, reflecting the control objective.

6.4.2 Information State

We use the framework described in previous sections to solve the optimal risk-sensitive control problem. Since the second (u-dependent) part of $C_1(u)$ is proportional to the identity, that part commutes with all operators and it is convenient to factor its contribution

to the risk-sensitive state by writing

$$\begin{aligned}\sigma_t^\mu &= \tfrac{1}{2}\left(n(t)I + x(t)\sigma_x + y(t)\sigma_y + z(t)\sigma_z\right)\exp\left(\tfrac{1}{2}\mu b\int_0^t |u(s)|^2 ds\right)\\ &= \tfrac{1}{2}\begin{pmatrix} n(t)+z(t) & x(t)-iy(t)\\ x(t)+iy(t) & n(t)-z(t)\end{pmatrix}\exp\left(\tfrac{1}{2}\mu b\int_0^t |u(s)|^2 ds\right).\end{aligned}$$

Then substitution into the SDE (190) shows that the coefficients satisfy the SDEs

$$\begin{aligned} dn(t) &= \tfrac{1}{2}\mu a(n(t)-z(t))dt + \kappa_s x(t)dy_2(t)\\ dx(t) &= -\tfrac{1}{2}(1-\mu a)x(t)dt + 2\kappa_f u_r(t)z(t)dt\\ &\quad +\kappa_s(n(t)+z(t))dy_2(t)\\ dy(t) &= -\tfrac{1}{2}(1-\mu a)y(t)dt - 2\kappa_f u_i(t)z(t)dt\\ dz(t) &= -(1-\tfrac{1}{2}\mu a)z(t)dt - (1+\tfrac{1}{2}\mu a)n(t)dt\\ &\quad -2\kappa_f(u_r(t)x(t)-u_i(t)y(t))dt - \kappa_s x(t)dy_2(t).\end{aligned} \tag{201}$$

The representation (189) reads

$$J^\mu(\mathbf{K}) = \mathbf{E}^0[\exp\left(\tfrac{1}{2}\mu\int_0^T b|u(s)|^2 ds\right)\tfrac{1}{2}(n(T)-z(T))e^{\mu c}]. \tag{202}$$

6.4.3 Dynamic Programming

We consider the value function (195) as a function of the coefficients, i.e. $S^\mu(n,x,y,z,t)$. In terms of these parameters, the dynamic programming equation is

$$\begin{aligned}&\tfrac{\partial}{\partial t}S^\mu(n,x,y,z,t) + \inf_{u\in\mathbf{C}}\{\mathcal{L}^{\mu;u}S^\mu(n,x,y,z,t) + \tfrac{1}{2}\mu b|u|^2 S^\mu(n,x,y,z,t)\} = 0, \quad 0\le t<T,\\ &S^\mu(n,x,y,z,T) = \tfrac{1}{2}(n-z)e^{\mu c},\end{aligned} \tag{203}$$

where the operator $\mathcal{L}^{\mu;u}$ is given, for sufficiently smooth functions $f(n,z,y,z)$, by

$$\begin{aligned}\mathcal{L}^{\mu;u}f &= \tfrac{1}{2}\kappa_s^2 x^2 f_{nn} + \tfrac{1}{2}\kappa_s^2(n+z)^2 f_{xx} + \tfrac{1}{2}\kappa_s^2 x^2 f_{zz}\\ &\quad +\kappa_s^2 x(n+z)f_{nx} - \kappa_s^2 x^2 f_{nz} - \kappa_s^2(n+z)xf_{xz}\\ &\quad +f_n(\tfrac{1}{2}\mu a(n-z)) + f_x(-\tfrac{1}{2}(1-\mu a)x + 2\kappa_f u_r z)\\ &\quad +f_y(-\tfrac{1}{2}(1-\mu a)y - 2\kappa_f u_i z)\\ &\quad +f_z(-(1-\tfrac{1}{2}\mu a)z - (1+\tfrac{1}{2}\mu a)n\\ &\quad -2\kappa_f(u_r x - u_i y).\end{aligned}$$

Here, the subscripts f_{nx}, etc, refer to partial derivatives, and the arguments n,x,y,z have been omitted.

To construct the optimal risk-sensitive controller $\mathbf{K}^{\mu,\star}$, we suppose that (203) has a smooth solution, which we write as

$$S^\mu(n,x,y,z,t) = n\exp\left(\frac{\mu}{n}W^\mu(n,x,y,z,t)\right). \tag{204}$$

The minimum over u in (203) can be explicitly evaluated by setting the derivatives of the expression in the parentheses (it is concave) with respect to u_r and u_i to zero. The result is

$$\begin{aligned}\mathbf{u}_r^{\mu,\star}(n,x,y,z,t) &= \frac{2\kappa_f}{bn}(xW_z^\mu(n,x,y,z,t) - zW_x^\mu(n,x,y,z,t)) \\ \mathbf{u}_i^{\mu,\star}(n,x,y,z,t) &= \frac{2\kappa_f}{bn}(zW_y^\mu(n,x,y,z,t) - yW_z^\mu(n,x,y,z,t)).\end{aligned} \tag{205}$$

The optimal risk-sensitive controller is then

$$\mathbf{K}^{\mu,\star} \;:\; u(t) = \mathbf{u}_r^{\mu,\star}(n(t),x(t),y(t),z(t),t) + i\mathbf{u}_i^{\mu,\star}(n(t),x(t),y(t),z(t),t), \tag{206}$$

where $n(t)$, $x(t)$, $y(t)$ and $z(t)$ are given by (201).

Note that the dynamic programming equation (203) (which is a partial differential equation of parabolic type) is solved backwards in time, using the terminal condition specified: $S^\mu(n,x,y,z,T) = \frac{1}{2}(n-z)e^{\mu c}$. The infimum in (203) can be removed by substituting in the optimal control values given by the explicit formulas (205), if desired. However, the form (203) is better suited to numerical computation, since the optimal control structure is preserved, [40]. Note that in this example, the risk-sensitive filter (190) is replaced by the finite-dimensional SDE (201); this fact is important for practical computational reasons.

6.4.4 Risk-Neutral Control

For comparison, consider the risk-neutral cost for this problem. Write

$$\sigma_t = \tfrac{1}{2}\left(n(t)I + x(t)\sigma_x + y(t)\sigma_y + z(t)\sigma_z\right). \tag{207}$$

Then from the SDE (176), we find that

$$\begin{aligned}dn(t) &= \kappa_s x(t)dy_2(t) \\ dx(t) &= -\tfrac{1}{2}x(t)dt + 2\kappa_f u_r(t)z(t)dt + \kappa_s(n(t)+z(t))dy_2(t) \\ dy(t) &= -\tfrac{1}{2}y(t)dt - 2\kappa_f u_i(t)z(t)dt \\ dz(t) &= -z(t)dt - n(t)dt - 2\kappa_f(u_r(t)x(t) - u_i(t)y(t))dt - \kappa_s x(t)dy_2(t).\end{aligned} \tag{208}$$

The risk-neutral representation (180) becomes

$$\begin{aligned}J(\mathbf{K}) = \mathbf{E}^0[\tfrac{1}{2}\int_0^T &(a(n(t)-z(t)+b|u(t)|^2)dt \\ &+\tfrac{1}{2}(n(T)-z(T))c],\end{aligned} \tag{209}$$

and the dynamic programming equation is

$$\begin{aligned}&\tfrac{\partial}{\partial t}W(n,x,y,z,t) + \inf_{u\in\mathbf{C}}\{\mathcal{L}^u W(n,x,y,z,t)\tfrac{1}{2}(a(n-z)+b|u|^2)\} = 0, \quad 0 \le t < T, \\ &W(n,x,y,z,T) = \tfrac{1}{2}(n-z)e^c,\end{aligned} \tag{210}$$

where

$$\begin{aligned}\mathcal{L}^u f \;=\; & \tfrac{1}{2}\kappa_s^2 x^2 f_{nn} + \tfrac{1}{2}\kappa_s^2 (n+z)^2 f_{xx} + \tfrac{1}{2}\kappa_s^2 x^2 f_{zz} \\ & + \kappa_s^2 x(n+z) f_{nx} - \kappa_s^2 x^2 f_{nz} - \kappa_s^2 (n+z) x f_{xz} \\ & + f_x(-\tfrac{1}{2}x + 2\kappa_f u_r z) \\ & + f_y(-\tfrac{1}{2}y - 2\kappa_f u_i z) \\ & + f_z(-z - n - 2\kappa_f(u_r x - u_i y)\end{aligned}$$

Evaluating the minimum in (210) gives

$$\begin{aligned}\mathbf{u}_r^\star(n,x,y,z,t) \;&=\; \frac{2\kappa_f}{bn}(xW_z(n,x,y,z,t) - zW_x(n,x,y,z,t)) \\ \mathbf{u}_i^\star(n,x,y,z,t) \;&=\; \frac{2\kappa_f}{bn}(zW_y(n,x,y,z,t) - yW_z(n,x,y,z,t)),\end{aligned} \tag{211}$$

cf. [13, eq. (15)]. The optimal risk-neutral controller is

$$\mathbf{K}^\star \;:\; u(t) \;=\; \mathbf{u}_r^\star(n(t),x(t),y(t),z(t),t) + i\mathbf{u}_i^\star(n(t),x(t),y(t),z(t),t), \tag{212}$$

where $n(t)$, $x(t)$, $y(t)$ and $z(t)$ are given by (208). Note that normalization of (208) results in [13, eq. (7)].

Note that the expressions for the both the risk-sensitive and risk-neutral controllers are similar, and involve a similar level of complexity for implementation. When $a = 0$, the risk-sensitive SDEs (201) reduces to the risk-neutral or standard SDEs (208), though the controllers will be different in general.

6.5 Control of a Trapped Atom

6.5.1 Setup

We now apply risk-sensitive optimal control to the problem of cooling and confinement of a quantum mechanical oscillator undergoing position measurement considered in the context of the linear quadratic Gaussian optimal control problem in [18]. As discussed there, this model corresponds in an appropriate limit to an atom trapped in an optical cavity, where the atom imparts a phase shift on the light that depends on its position and may be detected outside the cavity, see Figure 15.

In suitable coordinates, it is desired to keep the position q near zero (confinement), and the momentum p also near zero (cooling). A homodyne detection scheme on the light reflected from the cavity (assumed perfect) provides information about the position q. The Hamiltonian and measurement operators are

$$H = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 q^2 + b_1 u_1 p - b_2 u_2 q, \quad M = \sqrt{2k}q.$$

Here, q and p are considered as operators. The mass of the atom is m, and ω is the angular frequency of the harmonic motion in of the atom in its confining potential. The parameter

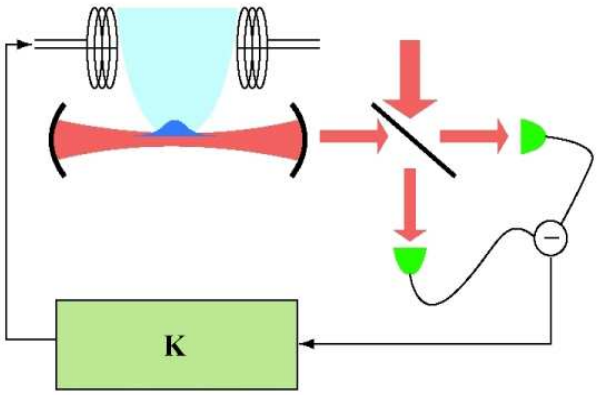

Figure 15: Atom in cavity feedback control.

k describes the sensitivity of the resulting measurement of position, larger values of k corresponding to better measurement sensitivity. This parameter can in principle be tuned by varying the power of the laser driving the cavity. The control input is $u = (u_1, u_2)^T$, with real coefficients b_1, b_2. The cost observables C_1 and C_2 are given by $C_1(u) = \frac{1}{2}(q,p)P(q,p)^T + \frac{1}{2}u^T Q u$, $C_2 = 0$, where P and Q are positive definite symmetric matrices.

6.5.2 Information State

It is shown in [18] that the conditional state π_t is Gaussian if π_0 is Gaussian. It also turns out that the risk-sensitive states σ_t^μ and π_t^μ are Gaussian, as can be shown by lengthy calculations. Indeed, the mean and covariance parameters of π_t^μ are given by

$$q^\mu(t) = \mathrm{tr}[q\pi_t^\mu],\ p^\mu(t) = \mathrm{tr}[p\pi_t^\mu]$$

and

$$Y_q^\mu = \mathrm{tr}[q^2\pi_t^\mu] - (\mathrm{tr}[q\pi_t^\mu])^2,\ Y_p^\mu = \mathrm{tr}[p^2\pi_t^\mu] - (\mathrm{tr}[p\pi_t^\mu])^2,$$
$$Y_{qp}^\mu = \mathrm{tr}[(qp+pq)\pi_t^\mu] - \mathrm{tr}[q\pi_t^\mu]\mathrm{tr}[p\pi_t^\mu].$$

It is important to realize that these are *not* the conditional means and covariances for position and momentum, which correspond to the conditional state π_t and are given by [18, eqs. (55), (48)-(50)].

The differential equations satisfied by these parameters are (when $P_{12} = 0$)

$$\begin{aligned} dq^\mu &= (p^\mu/m + \mu[P_{11}Y_q^\mu q^\mu + P_{22}Y_{qp}^\mu p^\mu] + b_1 u)dt + 2Y_q^\mu dw^\mu \\ dp^\mu &= (-m\omega^2 q^\mu + b_2 u_2)dt + \mu[P_{11}Y_{qp}^\mu q^\mu + P_{22}Y_p^\mu p^\mu] + 2Y_{qp}^\mu dw^\mu \end{aligned} \tag{213}$$

and

$$\begin{aligned} \dot{Y}_q^\mu &= (2/m)Y_{qp}^\mu - 8k(Y_q^\mu)^2 + \mu[P_{11}(Y_q^\mu)^2 + P_{22}(Y_{qp}^\mu)^2 - (\mu\hbar^2/4)P_{22}] \\ \dot{Y}_p^\mu &= -2m\omega^2 Y_{qp}^\mu - 8k(Y_{qp}^\mu)^2 + 2k\hbar^2 + \mu[P_{22}(Y_p^\mu)^2 + P_{11}(Y_{qp}^\mu)^2] - (\mu\hbar^2/4)P_{11} \\ \dot{Y}_{qp}^\mu &= Y_p^\mu/m - m\omega^2 Y_q^\mu - 8kY_{qp}^\mu Y_q^\mu + \mu[Y_{qp}^\mu(P_{11}Y_q^\mu + P_{22}Y_p^\mu)]. \end{aligned}$$

The initial conditions are $q^\mu(0) = q_0$, $p^\mu(0) = p_0$, $Y_q^\mu(0) = Y_{q0}$, $Y_p^\mu(0) = Y_{p0}$, and $Y_{qp}^\mu(0) = Y_{qp0}$, where q_0, p_0, Y_{q0}, Y_{p0} and Y_{qp0} are the Gaussian parameters for π_0.

To facilitate more compact and general notation, we write

$$x^\mu = \begin{pmatrix} q^\mu \\ p^\mu \end{pmatrix}, \; Y^\mu = \begin{pmatrix} Y_q^\mu & Y_{qp}^\mu \\ Y_{qp}^\mu & Y_p^\mu \end{pmatrix}.$$

Then the above equations take the form

$$dx^\mu = ([A + \mu Y^\mu P]x^\mu + Bu)dt + Y^\mu H^T(dy - Hx^\mu dt) \tag{214}$$

where $dw^\mu = dy - Hx^\mu dt$ and

$$\dot{Y^\mu} = AY^\mu + Y^\mu A^T - Y^\mu[H^T H - \mu P]Y^\mu + GG^T - (\mu\hbar^2/4)\Sigma P \Sigma^T \tag{215}$$

where

$$A = \begin{pmatrix} 0 & 1/m \\ -m\omega^2 & 0 \end{pmatrix}, \; B = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix},$$
$$G = \begin{pmatrix} 0 \\ \hbar \end{pmatrix}, \; H = 2\sqrt{2k}\begin{pmatrix} 1 & 0 \end{pmatrix}, \; \Sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

6.5.3 Optimal LEQG Control

We consider now how the optimal control is determined. In terms of the Gaussian parameters, the cost representation (192) takes the form (omitting some immaterial terms)

$$J^\mu(\mathbf{K}) = \mathbf{E}^\mu[\exp(\frac{\mu}{2}\int_0^T ({x^\mu}^T(t)Px^\mu(t) + u^T(t)Qu(t))dt)] \tag{216}$$

Consequently, the problem becomes a standard state feedback LEQG problem for the system (214), (215). The solution is [8], [49], [51]:

$$u^*(t) = -Q^{-1}B^T X^\mu(t)[I - \mu Y^\mu(t)X^\mu(t)]^{-1}x^\mu(t), \tag{217}$$

where

$$\dot{X^\mu} + A^T X^\mu + X^\mu A - X^\mu[BQ^{-1}B^T - \mu(GG^T - (\mu\hbar^2/4))\Sigma P \Sigma^T]X^\mu + P = 0. \tag{218}$$

It is important to note that this solution (217) differs from the classical LEQG solution (117) via the $\mu\hbar^2$ terms.

6.5.4 Robustness

Robustness of the QLEQG controller has been investigated in [51]. It can be seen from the simulation results in Figure 16 that the QLEQG controller provides a more graceful decline in performance (increase in average cost) than the LQG control controller as the nominal model/phyiscal model discrepancy, β, increases. This is typical of the robust properties of a general LEQG controller, recall section 4.4.

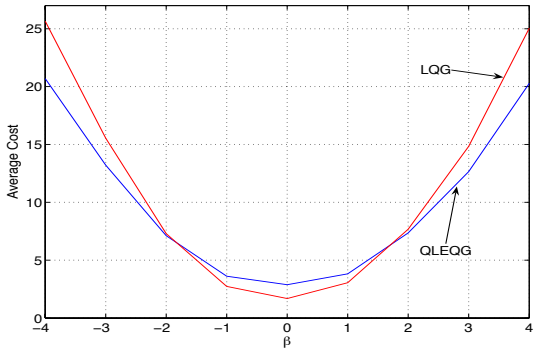

Figure 16: LQG vs. QLEQG (sample graph)

References

[1] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979.

[2] K.J. Astrom. *Introduction to Stochastic Control Theory*. Academic Press, New York, 1970.

[3] M. Bardi and I. Capuzzo-Dolcetta. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*. Birkhauser, Boston, 1997.

[4] T. Basar and P. Bernhard. *H^∞-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Birkhauser, Boston, second edition, 1995.

[5] V.P. Belavkin. On the theory of controlling observable quantum systems. *Automation and Remote Control*, 44(2):178–188, 1983.

[6] V.P. Belavkin. Nondemolition measurements, nonlinear filtering, and dynamic programming of quantum stochastic processes. In A. Blaquiere, editor, *Modelling and Control of Systems in Engineering, Quantum Mechanics, Economics, and Biosciences*, pages 245–265, New York, 1988. Springer Verlag.

[7] V.P. Belavkin. Quantum continual measurements and a posteriori collapse on CCR. *Commun. Math. Phys.*, 146:611–635, 1992.

[8] A. Bensoussan and J.H. van Schuppen. Optimal control of partially observable stochastic systems with an exponential-of-integral performance index. *SIAM Journal on Control and Optimization*, 23:599–613, 1985.

[9] D.P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, Boston, 1995.

[10] P. Billingsley. *Probability and Measure*. Wiley, New York, 1979.

[11] R.K. Boel, M.R. James, and I.R. Petersen. Robustness and risk sensitive filtering. *IEEE Transactions on Automatic Control*, 47(3):451–461, 2002.

[12] L. Bouten. *Filtering and Control in Quantum Optics*. PhD thesis, Catholic University of Nijmegen, 2004.

[13] L. Bouten, S. Edwards, and V.P. Belavkin. Bellman equations for optimal feedback control of qubit states. *J. Phys. B: At. Mol. Opt. Phys.*, 38:151–160, 2005.

[14] C.D. Charalambous. Partially observable nonlinear risk-sensitive control problems: Dynamic programming and verification theorems. *IEEE Trans. Automatic Control*, 42(8):1130–1138, 1997.

[15] M.G. Crandall, L.C. Evans, and P.L. Lions. Some properties of viscosity solutions of hamilton-jacobi equations. *Trans. Amer. Math. Soc.*, 282:487–502, 1984.

[16] M. H. A. Davis. Pathwise solutions and multiplicative functionals in nonlinear filtering. *in Proceeding of the 18th IEEE Conference on Decision and Control*, pages 176–180, 1979.

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

[18] A.C. Doherty and K. Jacobs. Feedback-control of quantum systems using continuous state-estimation. *Phys. Rev. A*, 60:2700, 1999. quant-ph/9812004.

[19] J.C. Doyle, K. Glover, P.P. Khargonekar, and B. Francis. State-space solutions to the standard H_2 and H_∞ control problems. *IEEE Transactions on Automatic Control*, 34(8):831–847, 1989.

[20] P. Dupuis and R.S. Ellis. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York, 1997.

[21] P. Dupuis, M.R. James, and I.R. Petersen. Robust properties of risk-sensitive control. *Math. Control, Systems and Signals*, 13:318–332, 2000.

[22] R.J. Elliott. *Stochastic Calculus and Applications*. Springer Verlag, New York, 1982.

[23] W.H. Fleming. Nonlinear semigroup for controlled partially observed diffusions. *SIAM J. Control*, 20(2):286–301, 1982.

[24] W.H. Fleming and R.W. Rishel. *Deterministic and Stochastic Optimal Control*. Springer Verlag, New York, 1975.

[25] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer Verlag, New York, 1993.

[26] B. A. Francis, J. W. Helton, and G. Zames. H^∞-optimal feedback controller for linear multivariable systems. *IEEE Trans. Auto Comnt.*, AC-29:888–900, 1984.

[27] C.W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. Springer, Berlin, third edition, 2004.

[28] C.W. Gardiner and P. Zoller. *Quantum Noise*. Springer, Berlin, 2000.

[29] K. Glover and J.C. Doyle. State-space formulae for all stabilizing controllers that satisfy an H_∞ norm bound and relations to risk-sensitivity. *Systems and Control Letters*, pages 167–172, 1988.

[30] M. Green and D.J.N. Limebeer. *Linear Robust Cotrol*. Prentice-Hall, Englewood Cliffs, NJ, 1995.

[31] J.W. Helton and M.R. James. *Extending H^∞ Control to Nonlinear Systems: Control of Nonlinear Systems to Achieve Performance Objectives*, volume 1 of *Advances in Design and Control*. SIAM, Philadelphia, 1999.

[32] D. Hill and P. Moylan. The stability of nonlinear dissipative systems. *IEEE Transactions on Automatic Control*, AC-21:708–7811, 1976.

[33] D. Hill and P. Moylan. Stability results for nonlinear feedback systems. *Automatica*, 13:377–382, 1977.

[34] A.S. Holevo. *Statistical Structure of Quantum Theory*. Springer, Berlin, 2001.

[35] M.R. James. Risk-sensitive optimal control of quantum systems. *Phys. Rev. A*, 69:032108, 2004.

[36] M.R. James. A quantum Langevin formulation of risk-sensitive optimal control. *J. Optics B: Semiclassical and Quantum, special issue on quantum control*, to appear 2005.

[37] M.R. James, J.S. Baras, and R.J. Elliott. Risk-sensitive control and dynamic games for partially observed discrete-time nonlinear systems. *IEEE Transactions on Automatic Control*, 39:780–792, 1994.

[38] P.R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Prentice-Hall, Englewood Cliffs, NJ, 1986.

[39] H.J. Kushner. *Introduction to Stochastic Control Theory*. Academic Press, New York, 1972.

[40] H.J. Kushner and P.G. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, New York, 1992.

[41] M. Guta L .Bouten and H. Maassen. Stochastic Schrodinger equations. *J. Phys. A: Math. Gen.*, (37):3189–3209, 2004.

[42] M.A. Nielsen and I.L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[43] K.R. Parthasarathy. *An Introduction to Quantum Stochastic Calculus*. Birkhauser, Berlin, 1992.

[44] I.R. Petersen. Disturbance attenuation and H_∞ optmization: A design method based on the algebraic riccati equation. *IEEE Trans. Aut. Control*, 32(5):427–429, 1987.

[45] I.R. Petersen, B.D.O. Anderson, and E.A. Jonckheere. A first principles solution to the non-singular H^∞ control problem. *International Journal of Robust and Nonlinear Control*, 1(3):171–185, 1991.

[46] A.J. van der Schaft. *L_2-Gain and Passivity Techniques in Nonlinear Control*. Springer Verlag, New York, 1996.

[47] M. Vidyasagar. *Nonlinear Systems Analysis*. Prentice-Hall, Englewood Cliffs, second edition, 1993.

[48] P. Whittle. Risk-sensitive linear/ quadratic/ Gaussian control. *Advances in Applied Probability*, 13:764–777, 1981.

[49] P. Whittle. *Risk-Sensitive Optimal Control*. Wiley, New York, 1990.

[50] J.C. Willems. Dissipative dynamical systems - Part I: General theory. *Archive of Rational Mechanics and Analysis*, 45:321–351, 1972.

[51] S.D. Wilson, C. D'Helon, A.C. Doherty, and M.R. James. Linear exponential quadratic gaussian optimal control for quantum systems. 2005.

[52] H. Wiseman. *Quantum Trajectories and Feedback*. PhD thesis, University of Queensland, Brisbane, Australia, 1994.

[53] G. Zames. Optimal sensitivity and feedback: weighted seminorms, approximate inverses, and plant invariant schemes. In *Proc. Allerton Conf.*, 1979.

[54] K. Zhou, J. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, NJ, 1996.