1. Introduction

Recent activity in nonlinear $H_\infty$ control has sparked renewed interest in the theory of dissipative nonlinear systems. This theory was initiated by Willems (1972) and developed further by Hill & Moylan (1976). Central to this theory are the so-called storage functions. These functions are generalized energy functions defined in terms of dissipation inequalities. The natural infinitesimal versions of such dissipation inequalities (obtained by dynamic programming) are partial differential inequalities (PDIs), special cases of which relate to nonlinear generalizations of the Positive Real Lemma and the Bounded Real Lemma. Such PDIs are important since one hopes to obtain storage functions by solving them. However, the PDIs involve first order partial derivatives of the storage functions, which unfortunately need not exist in general. The purpose of these notes is to explain (i) how functions which need not be differentiable can solve such PDIs in a generalized sense (viscosity sense), and (ii) how to compute numerical approximations to solutions (i.e. storage functions) of such PDIs.

In §2 we consider a simple optimal control problem and present the corresponding PDE (a nonlinear first-order equation) which arises in dynamic programming. In §3 we discuss such nonlinear PDEs, pointing out the difficulty concerning the correct definition of solution and uniqueness. The definition of viscosity solution is given in §4. A simple finite difference numerical approximation scheme is discussed in §5, and an example is given in §6 which illustrates the scheme and a consequence of the lack of smoothness. In §7 the discussion turns to dissipative systems, PDIs, and the viscosity solution interpretation. These results are applied in §8 to the problem of computing the “$H_\infty$ norm” of nonlinear systems. Two examples are given in §§9, 10.

2. Optimal Control

Consider the problem of minimizing the cost functional

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

(2.1)

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

$$ \begin{cases} 
\dot{x}(s) = f(x(s), u(s)), & t \leq s \leq t_1, \\
 x(t) = x. 
\end{cases} 
$$

(2.2)

Here, $u(\cdot)$ is a control defined on $[t, t_1]$ taking values in, say, a compact set $U \subset \mathbb{R}^m$, and $x(\cdot)$ is the state trajectory in $\mathbb{R}^n$. Let’s use dynamic programming to solve this problem.
The optimal cost or value function is defined by

$$V(t, x) = \inf_{u(\cdot)} J(t, x; u)$$

(2.3)

for \((t, x) \in [t_0, t_1] \times \mathbb{R}^n\), and the dynamic programming principle states that for every \(r \in [t, t_1]\),

$$V(t, x) = \inf_{u(\cdot)} \left[ \int_t^r L(x(s), u(s)) \, ds + V(r, x(r)) \right].$$

(2.4)

From this, one can derive formally the equation

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

(2.5)

with terminal data

$$V(t_1, x) = \psi(x) \quad \text{in} \quad \mathbb{R}^n.$$  

(2.6)

Here, the Hamiltonian is given by

$$H(x, p) = \inf_{u \in U} \{ p \cdot f(x, u) + L(x, u) \}. $$

(2.7)

The nonlinear first order PDE (2.5) is the dynamic programming equation (DPE), or Hamilton-Jacobi-Bellman (HJB) equation. The utility of the value function and the DPE is as follows. If there exists a \(C^1([t_0, t_1] \times \mathbb{R}^n)\) solution \(W\) to (2.5), (2.6) (i.e. the DPE has a classical solution), and if \(u^*(\cdot) \in L^\infty\) is such that

$$u^*(s) \in \arg\max_{u \in U} \{ \nabla_x W(s, x^*(s)) \cdot f(x^*(s), u) + L(x^*(s), u) \},$$

then \(u^*(\cdot)\) is optimal, and \(x^*(\cdot)\) is the corresponding optimal state trajectory. In addition, \(W(t, x) = V(t, x)\). This is the content of the well known verification theorem Fleming & Rishel (1975), Fleming & Soner (1993).

In general \(V \notin C^1([t_0, t_1] \times \mathbb{R}^n)\); indeed \(V\) may not be everywhere differentiable, in which case (2.5) does not have a classical solution, and the verification theorem does not hold (at least in this form, see Clarke (1983)). In spite of this, the DPE is still of fundamental importance, and in fact \(V\) turns out to be the unique viscosity solution of (2.5), (2.6).

3. Nonlinear PDE

As we saw in §2, dynamic programming leads to nonlinear PDEs. In this section, we take a quick look at such equations.

The method of characteristics is a classical method for solving equations of the type (2.5), as well as other first-order equations. This method constructs smooth functions which satisfy the PDE locally, in certain regions by integrating along the characteristic curves (obtained by solving Hamilton’s equations). In general, a globally defined smooth classical solution cannot be obtained. Typically, the derivatives may suffer discontinuities across certain lower dimensional sets. This phenomenon is well-known in the calculus of variations, see Fleming & Rishel (1975), Fleming & Soner (1993).

In general \(V \notin C^1([t_0, t_1] \times \mathbb{R}^n)\); indeed \(V\) may not be everywhere differentiable, in which case (2.5) does not have a classical solution, and the verification theorem does not hold (at least in this form, see Clarke (1983)). In spite of this, the DPE is still of fundamental importance, and in fact \(V\) turns out to be the unique viscosity solution of (2.5), (2.6).
Consider the equation
\[
\begin{aligned}
\frac{\partial}{\partial t} V - |\nabla_x V|^2 &= 0 \quad \text{in } (0, 1) \times \mathbb{R}, \\
V(x, 1) &= 0 \quad \text{in } \mathbb{R}.
\end{aligned}
\] (3.1)

The function
\[
V^1(t, x) \equiv 0
\]
is a classical solution, while
\[
V^2(t, x) = \begin{cases}
0 & \text{if } |x| > 1 - t \\
x - 1 + t & \text{if } 0 < x < 1 - t \\
-x - 1 + t & \text{if } -1 + t < x < 0
\end{cases}
\]
is a locally Lipschitz generalized solution.

A satisfactory notion of solution should enjoy a uniqueness property. One pre-viscosity uniqueness result concerns equations with convex Hamiltonian $H$, as in the example. Equation (3.1) has a unique semi-concave generalized solution, namely $V^1$. In fact, this solution is given by Hopf’s formula:
\[
V^1(t, x) = \min_{y \in \mathbb{R}} \left[ (1 - t) \frac{1}{4} \frac{x - y}{1 - t} \right] = 0.
\]
This variational formula is actually a special case of an optimal control or calculus of variations representation. Equation (3.1) can be interpreted as a DPE for an optimal control problem with $f(x, u) = u, L(x, u) = u^2/4, \psi = 0$, with value function $V^1$:
\[
V^1(t, x) = \inf_{u(\cdot)} \left\{ \int_{t}^{1} \frac{1}{4} |u(s)|^2 \, ds \right\} = 0.
\]
Thus variational formulae “pick out the right solution”. This observation is relevant to the theory of viscosity solutions, as value functions turn out to be viscosity solutions. The viscosity solution definition also singles out the “right solution”, and it is very important to note that convexity or semi-concavity assumptions are not required, and hence the concept applies more generally.

4. Viscosity Solutions

The definition of viscosity solution first appeared in Lions (1982), Crandall & Lions (1984), Crandall, Evans & Lions (1984). A function $W \in C([t_0, t_1] \times \mathbb{R}^n)$ is a viscosity subsolution (resp. supersolution) of (2.5) if for all $\phi \in C^\infty((t_0, t_1) \times \mathbb{R}^n)$,
\[
\frac{\partial}{\partial t} \phi(t', x') + H(x', \nabla x \phi(t', x')) \geq 0 \quad \text{(resp. } \leq 0) \quad (4.1)
\]
at every point $(t', x')$ where $W - \phi$ attains a local maximum (resp. minimum). $W$ is a viscosity solution if it is both a subsolution and a supersolution.

There are a number of equivalent definitions, and the one just stated was given in Crandall, Evans & Lions (1984), and offers advantages such as ease of use (despite apparent awkwardness). The functions $\phi \in C^\infty$ play the role of “test functions”, and note that the derivatives of $\phi$ replace those of $W$, which need not exist. Thus differentiation has been transferred to smooth functions, at the expense of a pair of inequalities. (Recall that in the theory of linear PDE, this is achieved using integration by parts and generalized derivatives or distributions.)
What is happening in (4.1) is that the test functions are characterizing the super- and subdifferentials of $W$: $D^+ W$ and $D^- W$. The superdifferential is defined by

$$D^+ W(t, x) = \left\{ (q, p) : \limsup_{(s, y) \to (t, x)} \frac{W(s, y) - W(t, x) - q(s - t) - p(y - x)}{|s - t| + |y - x|} \leq 0 \right\},$$

and analogously for the subdifferential $D^- W$. The super- and subdifferentials can be expressed equivalently in terms of the test functions:

$$D^+ W(t, x) = \left\{ (q, p) : (q, p) = (\frac{\partial}{\partial n}\phi(t, x), \nabla_x \phi(t, x)) \text{ where } (t, x) \text{ is a local maximum for } W - \phi, \text{ for some } \phi \in C^\infty \right\}$$

Inequalities (4.1) mean that

$$q + H(x, p) \geq 0 \text{ (resp. } \leq 0) \text{ for all } (q, p) \in D^+ W(t, x) \text{ (resp. } D^- W(t, x)).$$

If $W$ is differentiable at $(t', x')$, then (2.5) holds at this point, and a classical solution is also a viscosity solution.

Refering to the above example, the function $V^1 \equiv 0$ is a classical and hence a viscosity solution of (3.1). However, $V^2$ is not a viscosity solution because the supersolution property fails. Indeed, consider the point $(t', x') = (\frac{1}{2}, 0)$. Then $D^- V(\frac{1}{2}, 0) = \{(1, p) : p \in [-1, 1]\}$, and with $(q, p) = (1, 0) \in D^- W(\frac{1}{2}, 0)$ we see that

$$1 + 0 \not\geq 0.$$

The subsolution property does hold (in fact, $D^+ W(\frac{1}{2}, 0) = \emptyset$, and there is nothing to check at the point $t', 0$).

Existence results are available, using a variety of methods, for example, optimal control and games, vanishing viscosity, and Perron’s method. In particular, under very general conditions, value functions for optimal control or game problems are viscosity solutions. The dynamic programming principle is the key to this, and is discussed at length in the book, in both deterministic and stochastic contexts. To see how this works, suppose that $V - \phi$ attains a maximum (resp. minimum) at $(t', x')$. Set $r = t' + h, h > 0$, and rearrange (2.4) to yield

$$\inf_{W(t')} \left[ \frac{1}{h} (V(t' + h, x(t' + h)) - V(t', x')) + \frac{1}{h} \int_t^{t' + h} L(x(s), u(s)) \, ds \right] = 0.$$

But also

$$(\phi(t' + h, x(t' + h)) - \phi(t', x')) \geq (\text{resp. } \leq) (V(t' + h, x(t' + h)) - V(t', x')).$$

Combining these two displays and sending $h \to 0$ yields (4.1).

Uniqueness theorems are often presented in the form of a comparison theorem, as follows. Let $W$ and $V$ be (bounded) viscosity subsolutions and supersolutions, respectively, of (2.5). Then, assuming some technical conditions are met,

$$\sup_{(t, x) \in [t_0, t_1] \times \mathbb{R}^n} (W(t, x) - V(t, x)) = \sup_{x \in \mathbb{R}^n} (W(t_1, x) - V(t_1, x)) = 0.$$  

In particular, if $W(t_1, t) \leq V(t_1, t)$, then $W \leq V$. Uniqueness is deduced from (4.2). If $W$ and $V$ are both viscosity solutions of (2.5) satisfying (2.6), then (4.2) implies $W \leq V$ and $V \leq W$; hence $W = V$. In our example (equation (3.1)), the function $V^1 \equiv 0$ is the unique viscosity solution of (3.1). In general, the value function $V$ defined by (2.3) is the unique viscosity solution of (2.5).
5. Numerical Solution

Finite difference and finite element schemes are commonly employed to solve both linear and nonlinear PDEs. The method is discussed in a number of references, e.g., Capuzzo Dolcetta & Falcone (1989), Kushner & Dupuis (1993).

We consider a stationary version of (2.5), viz.,

$$\min_u [\nabla_x V \cdot f(x, u) + L(x, u)] = 0 \text{ in } \mathbb{R}^n,$$

(5.1)

and, assuming validity of the verification theorem, the optimal controller \( u^*(x) \) is the control value achieving the minimum in (5.1):

$$u^*(t) = u^*(x^*(t)), \quad t \geq 0.$$

(5.2)

Assume that a solution \( V(x) \) exists.

For \( \delta > 0 \) define the grid \( \delta \mathbb{Z}^n = \{ \delta z : z_i \in \mathbb{Z}, i = 1, \ldots, n \} \). We wish to construct an approximation \( V^\delta(x) \) to \( V(x) \) on this grid. The basic idea is to approximate the derivative \( \nabla_x V(x) \cdot f(x, u) \) by the finite difference

$$\sum_{i=1}^n f^\pm_i(x, u)(V(x \pm \delta e_i) - V(x))/\delta,$$

(5.3)

where \( e_1, \ldots, e_n \) are the standard unit vectors in \( \mathbb{R}^n \), \( f^+_i = \max(f_i, 0) \) and \( f^-_i = -\min(f_i, 0) \). A finite difference replacement of (5.1) is

$$V^\delta(x) = \min_u \left\{ \sum_{z} p^\delta(x, z|u)V^\delta(z) + \Delta t(x)L(x, u) \right\}.$$

(5.4)

Here,

$$p^\delta(x, z|u) = \begin{cases} 1 - \frac{\| f(x, u) \|_1}{\max_u \| f(x, u) \|_1} & \text{if } z = x \\ \frac{f^+_i(x, u)}{\max_u \| f(x, u) \|_1} & \text{if } z = x \pm \delta e_i \\ 0 & \text{otherwise}, \end{cases}$$

(5.5)

\( \| v \|_1 = \sum_{i=1}^n |v_i| \), and \( \Delta t(x) = \delta/\max_u \| f(x, u) \|_1 \). Equation (5.4) is obtained by substituting (5.3) for \( \nabla_x V \cdot f \) in (5.1) and some manipulation. In Kushner & Dupuis (1993), the quantities \( p^\delta(x, z|u) \) are interpreted as transition probabilities for a controlled Markov chain, with (5.4) as the corresponding DPE. The optimal state feedback policy \( u^*_{\text{opt}}(x) \) for this discrete optimal control problem is obtained by finding the control value attaining the minimum in (5.4). Using this interpretation, convergence results can be proven (viscosity solution methods can also be used, see, e.g., Theorem 8.2 below). In our case, it is interesting that a deterministic system is approximated by a (discrete) stochastic system. See Figure 1. One can imagine a particle moving randomly on the grid trying to follow the deterministic motion or flow, on average. When the particle is at a grid point, it jumps to one of its neighbouring grid points according to the transition probability.

In practice, one must use only a finite set of grid points, say a set forming a box \( D^\delta \) centred at the origin. On the boundary \( \partial D^\delta \), the vector field \( f(x, u) \) is modified by projection, so if \( f(x, u) \) points out of the box, then the relevant components are set to zero. This has the effect of constraining the dynamics to a bounded region. From now on we take (5.4) to be defined on \( D^\delta \), with \( p^\delta \) modified accordingly.

Equation (5.4) is a nonlinear implicit relation for \( V^\delta \), and iterative methods are used
to solve it (approximately). There are two main types of methods, viz., *value space iteration*, which generates a sequence of approximations

\[ V_\delta^k(x) = \min_u \left\{ \sum_z p_\delta(x, z|u)V_{k-1}^\delta(z) + \Delta t(x)L(x, u) \right\} \]

(5.6)

to \( V^\delta \), and *policy space iteration*, which involves a sequence of approximations \( u^*_\delta(x) \) to the optimal feedback policy \( u^*(x) \). There are many variations and combinations of these methods, acceleration procedures, and multigrid algorithms. Note that because of the local structure of equation (5.4), *parallel computation* is natural.

6. Example

We illustrate the finite difference scheme with an example which shows that value functions can fail to be smooth and have discontinuous optimal controls, viz., the classical minimum time problem. The dynamics are \( \dot{x}_1 = x_2, \ \dot{x}_2 = u \), and the controls take values in the compact set \( U = [-1, 1] \). The problem is to find the controller which steers the system from an initial state \( x = (x_1, x_2) \) to the origin in minimum time. The value function is simply the minimum time possible:

\[ T(x) = \inf_{u(.)} \left\{ \int_0^{t_f} 1 \, dt : x(0) = x, \ x(t_f) = 0 \right\}. \]

(6.1)

For a number of reasons, it is convenient to use a transformed minimum time function

\[ S(x) = 1 - e^{-T(x)}, \]

with DPE

\[
\begin{align*}
S(x) &= \min_{u \in [-1,1]} \{ \nabla_x S(x) \cdot f(x,u) + 1 \} \\
S(0) &= 0.
\end{align*}
\]

(6.2)
Note that $0 < S(x) \leq 1$ for $x \neq 0$. The presence of the term $S(x)$ on the LHS of (6.2) changes slightly the numerical scheme, but applying the same principles as above gives

$$
\begin{align*}
S^\delta(x) &= \min_{u \in [-1, 1]} \left\{ \frac{1}{1 + \Delta t(x)} \left( \sum_z p^\delta(x, z; u) S^\delta(z) + \Delta t(x) \right) \right\} \\
S^\delta(0) &= 0.
\end{align*}
$$

(6.3)

The value function $S(x)$ is not differentiable at $x = 0$, rather, it is only Holder continuous there (with exponent $\frac{1}{2}$), see Figure 2. An implication of this lack of smoothness is that the rate of convergence is of order $\delta^{\frac{1}{2}}$, which is rather slow; this has been verified empirically, see Figure 3 (from Cahill et al (1994)), showing the error as a function of $\delta$.

The optimal state feedback $u^*(x)$ equals either $-1$ or $+1$, and has a single discontinuity across the so-called switching curve, see Figure 4.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{A non-smooth value function.}
\end{figure}

7. Dissipative Systems

The control systems $\Sigma$ we consider are described by models of the form

$$
\begin{align*}
\dot{x}(t) &= f(x(t)) + g(x(t))u(t), \quad t > 0, \quad x(0) = x_0, \\
y(t) &= h(x(t)), \quad t \geq 0,
\end{align*}
$$

(7.1)

where $x(t) \in \mathbb{R}^n$, $u(t) \in U$, a closed subset of $\mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$. We assume $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$, $g \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and $h \in C^1(\mathbb{R}^n, \mathbb{R}^p)$, and that the first order derivatives $Df$, $Dg$ and $ Dh $ are bounded. In addition, we suppose $f(0) = 0$ and $h(0) = 0$, so that $x = 0$ is an equilibrium for the uncontrolled system. Given a control $u : [0, \infty) \rightarrow U$, the solution at time $t \geq 0$ with initial condition $x_0$ is denoted $x(t) = \gamma_u(t)x_0$; the corresponding output is $y(t) = h(\gamma_u(t)x_0)$. 

We use the notion of dissipative system as introduced by Willems (1972). A locally Lipschitz continuous function \( w : U \times \mathbb{R}^p \to \mathbb{R} \) satisfying the growth condition \(|w(u, y)| \leq C(1 + |u|^2 + |y|^2)\) is called a supply rate. Let \( U \) denote the set of locally square integrable controls \( u : [0, \infty) \to U \) such that \( \int_0^t |w(r)| \, dr < \infty \) for all \( x_0 \in \mathbb{R}^n, \ t \geq 0, \) where \( w(t) = w(u(t), y(t)) \). A real valued function \( \psi \) is locally bounded if for any compact subset \( K \) of the domain, there exists a constant \( C > 0 \) such that \( \sup_{x \in K} |\psi(x)| \leq C \).

**Definition 7.1.** The system \( \Sigma \) with supply rate \( w \) is said to be dissipative if there exists a locally bounded non-negative function \( V : \mathbb{R}^n \to \mathbb{R} \), called a storage function,
such that

\[
V(x) \geq \sup_{t \geq 0, u \in U} \left\{ V(x(t)) - \int_0^t w(r) \, dr : x(0) = x \right\}.
\]

This relation is known as the dissipation inequality, and expresses a constraint on the amount of (generalised) energy which can be extracted from the system. \( V(x) \) is the amount of energy stored in the system when it is in state \( x \), and is a candidate Lyapunov function. In general, a storage function is not uniquely defined, and in fact there is a continuum of storage functions for dissipative systems: \( V_0 \leq V \leq V_r \), where \( V_0 \) is the available storage, and \( V_r \) is the required supply (finite if \( \Sigma \) is reachable from 0) Willems (1972).

Willems’ definition Willems (1972) is cast in a very general setting, and regularity requirements are not part of his definition. However, any dissipative system possesses a lower semicontinuous (l.s.c.) storage function:

**Proposition 7.2.** If a locally bounded function \( V \) satisfies the dissipation inequality (7.2), then so does its lower semicontinuous envelope \( V_* \), defined by

\[
V_*(x) = \lim_{z \to x} V(z),
\]

and hence \( V_* \) is a lower semicontinuous storage function.

**Proof.** Fix \( t \geq 0, u \in U \) and \( x_0 \in \mathbb{R}^n \). Select a sequence \( \{x_i\}_{i=1}^{\infty} \) such that \( x_0 = \lim_{i \to \infty} x_i \) and \( V_*(x_0) = \lim_{i \to \infty} V(x_i) \). Since \( V \geq V_* \), (7.2) implies

\[
V(x_i) \geq V_*(\gamma_u(t)x_i) - \int_0^t w(u(r), h(\gamma_u(r)x_i)) \, dr.
\]

Send \( i \to \infty \) to obtain

\[
V_*(x_0) \geq \lim_{i \to \infty} V_*(\gamma_u(t)x_i) - \int_0^t w(u(r), h(\gamma_u(r)x_0)) \, dr.
\]

This inequality holds for all \( u \in U, t \geq 0 \), and implies that \( V_* \) satisfies (7.2).

One can test for dissipativeness by attempting to compute the available storage, defined by

\[
V_a(x) = \sup_{t \geq 0, u \in U} \left\{ -\int_0^t w(r) \, dr : x(0) = x \right\}.
\]

If \( V_a \) is locally bounded, then \( V_a \) is a storage function and \( \Sigma \) is dissipative. This is a variational test. Also, \( V_{a*} \) is a l.s.c. storage function.

Applying dynamic programming methods to the definition of storage function leads to the PDI

\[
H(x, \nabla_x V) = \sup_{u \in U} (\nabla_x V \cdot (f + gu) - w(u, h)) \leq 0 \text{ in } \mathbb{R}^n,
\]

where the Hamiltonian \( H(x, p) = \sup_{u \in U} (p \cdot (f(x) + g(x)u) - w(u, h(x))) \) is an extended–real valued l.s.c. function, locally bounded below: \( H(x, p) \geq C(1 + |p| + |p||x| + |x|^2) \). \( H \) is continuous if \( U \) is compact.

Since value functions, and hence storage functions, are in general not smooth, we use the viscosity solution framework to define solutions to the PDI (7.4). The following theorem characterises dissipativeness in terms of this PDI. Since this PDI is interpreted in the viscosity sense, we do not require \( V \) to be \( C^1 \), as was the case for a classical version of this result due to Hill & Moylan (1976). See James (1993), Ball & Helton (1994).
Since $\Sigma$ is dissipative, from Proposition 7.2 and (7.2) we have

$$H(x_0, \nabla_x \phi(x_0)) = \sup_{u \in U} (\nabla_x \phi(x_0) \cdot (f(x_0) + g(x_0)u) - w(u, h(x_0))) \leq 0. \quad (7.5)$$

Select a constant control $u(t) = u \in U$ for all $t \geq 0$, and let $x(t)$ denote the corresponding trajectory with $x(0) = x_0$. For $t \geq 0$ sufficiently small, we have

$$V_s(x_0) - V_s(x(t)) \leq \phi(x_0) - \phi(x(t)).$$

Since $\Sigma$ is dissipative, from Proposition 7.2 and (7.2) we have

$$V_s(x_0) - V_s(x(t)) \geq -\int_0^t w(u, h(x(r))) \, dr.$$

Combining these two inequalities, we get

$$\frac{\phi(x(t)) - \phi(x_0)}{t} - \frac{1}{t} \int_0^t w(u, h(x(r))) \, dr \leq 0.$$

Send $t \downarrow 0$ to obtain

$$\nabla_x \phi(x_0) \cdot (f(x_0) + g(x_0)u) - w(u, h(x_0)) \leq 0.$$

This inequality holds for all $u \in U$, hence (7.5).

2. To prove the converse, we adapt a technique used in Lions & Souganidis (1985). For $R > 0$, define $U_R = \{u \in U : |u| \leq R\}$ and let $U_R$ denote the set of controls with values in $U_R$. Since $V_s$ is l.s.c., there exist $\{\psi_i\}_{i=1}^\infty \subset C(\mathbb{R}^n)$ such that $\psi_i \leq V_s$ and $\psi_i \uparrow V_s$ as $i \to \infty$.

Let $T > 0$ and define

$$Z^i_R(x, s) = \sup_{u \in U_R} \left\{ \psi_i(x(T)) - \int_s^T w(r) \, dr : x(s) = x \right\}.$$

Then $Z^i_R$ is continuous and is the unique solution of

$$\begin{cases}
\frac{\partial Z^i_R}{\partial t} + \sup_{u \in U_R} (\nabla_x Z^i_R \cdot (f + gu) - w(u, h)) = 0 & \text{in } \mathbb{R}^n \times (0, T) \\
Z^i_R(x, T) = \psi_i(x) & \text{in } \mathbb{R}^n.
\end{cases}$$

Now $V_s$ is a supersolution of this PDE, and so the comparison theorem implies, for all $i = 1, 2, \ldots$

$$V_s(x) \geq Z^i_R(x, s) \text{ for all } (x, s) \in \mathbb{R}^n \times [0, T].$$

Setting $s = 0$ yields

$$V_s(x) \geq \sup_{u \in U_R} \left\{ \psi_i(x(T)) - \int_0^T w(r) \, dr : x(0) = x \right\}.$$

Let $i \to \infty$ to obtain

$$V_s(x) \geq \sup_{u \in U_R} \left\{ V_s(x(T)) - \int_0^T w(r) \, dr : x(0) = x \right\}.$$
Sending $R \to \infty$ gives

$$V_s(x) \geq \sup_{u \in U} \left\{ V_s(x(T)) - \int_0^T w(r) \, dr : x(0) = x \right\}.$$  

Since this inequality holds for all $T \geq 0$, we obtain (7.2). Therefore, $\Sigma$ is dissipative and $V_s$ is a storage function.

The next theorem says that the property of dissipativeness is stable under certain perturbations in the data defining the dynamical system.

**Theorem 7.4.** Let $\Sigma^\varepsilon = (f^\varepsilon, g^\varepsilon, h^\varepsilon)$, $\varepsilon > 0$, denote a family of systems of the form (7.1) ($n$, $m$ and $p$ are fixed). Let $f^\varepsilon \to f$, $g^\varepsilon \to g$, and $h^\varepsilon \to h$ locally uniformly as $\varepsilon \to 0$, for some limit system $\Sigma = (f, g, h)$. Assume that:

(i) Each system $\Sigma^\varepsilon$ is dissipative with respect to supply rates $w^\varepsilon$, and the corresponding l.s.c. storage functions satisfy

$$\sup_{\varepsilon > 0} \| V^\varepsilon \|_{L^\infty(R^n)} < \infty$$

(ii) The corresponding Hamiltonians satisfy

$$\liminf_{\varepsilon \to 0, z \to x, q \to p} H^\varepsilon(z, q) \geq H(x, p)$$

Then the limit system $\Sigma$ is dissipative, and $V(x) = \liminf_{\varepsilon \to 0, z \to x} V^\varepsilon(z)$ is a storage function for $\Sigma$.

**Proof.** Note first that $V$ is well defined and l.s.c.

Let $\phi \in C^1(R^n)$ and assume that, without loss of generality, $V - \phi$ attains a strict local minimum at $x_0 \in R^n$. There is a subsequence $\varepsilon_i \to 0$ as $i \to \infty$ such that

$$V^\varepsilon_i(x^\varepsilon_i) \to V(x_0), \quad x^\varepsilon_i \to x_0$$

as $i \to \infty$, and $V^\varepsilon_i - \phi$ attains a local minimum at $x^\varepsilon_i$. Barles & Perthame (1987). Since each $\Sigma^\varepsilon_i$ is dissipative, by Theorem 7.3 we have

$$H^\varepsilon_i(x^\varepsilon_i, \nabla_x \phi(x^\varepsilon_i)) \leq 0, \quad i = 1, 2, \ldots$$

Send $i \to \infty$ to obtain

$$H(x_0, \nabla_x \phi(x_0)) \leq 0.$$  

By Theorem 7.3, $\Sigma$ is dissipative and $V$ is a storage function.

We noted above the fundamental lack of uniqueness of storage functions for dissipative systems. Consequently, we cannot expect uniqueness of solutions of the PDI (7.4). The available storage $V_a$ (defined by (7.3)) is the minimum solution of the PDI (7.4): $V_a \leq V$ for all locally bounded non–negative l.s.c. solutions $V$ of (7.4). One can show that $V_a$ solves (7.4) with equality, i.e., $V_a$ is a solution of the PDE

$$H(x, \nabla_x V) = 0 \text{ in } R^n. \quad (7.6)$$

This means that $V_a$ is a supersolution of (7.6) (i.e. solution of (7.4), by Theorem 7.3), and $V_a$ is a subsolution of (7.6). A proof of this assertion involves a representation of $V^*_a$ in terms of a relaxed version of the optimisation problem.
8. Computation of the $H_{\infty}$ Norm or $L_2$ Gain

We turn now to the problem of computing the so-called $H_{\infty}$ norm or $L_2$ gain of the nonlinear system $\Sigma$; see James & Yuliar (1994). In the model (7.1), the input $u$ is a disturbance whose effect on a given output quantity $y$ is to be measured, for example using the "$H_{\infty}$ norm" of the input-output map $\Sigma$ relating $u$ and $y$, given $x(0) = 0$. $\Sigma$ is a map defined by

$$(\Sigma u)(t) = h(\gamma u(t)0) \quad \text{for all } t \geq 0,$$

and if $\Sigma$ maps $L_2([0, \infty), U)$ into $L_2([0, \infty), \mathbb{R}^p)$, the $H_{\infty}$ norm of $\Sigma$ could perhaps be defined by analogy with the corresponding linear systems definition in the time domain:

$$\| \Sigma \|_{\infty} = \sup_{u \in L_2([0, \infty), U), u \neq 0} \frac{\| \Sigma u \|_2}{\| u \|_2},$$  \hspace{1cm} (8.1)

where $\| \cdot \|_2$ is the usual $L_2$ norm.

By definition, let us say that $\| \Sigma \|_{\infty} \leq \gamma$ if and only if for $x(0) = 0$ we have

$$\int_0^T \gamma^2 |u(t)|^2 - |y(t)|^2 \, dt \geq 0 \quad \text{for all } T \geq 0 \text{ and all } u \in L_2([0, T], U).$$  \hspace{1cm} (8.2)

for all $T \geq 0$ and all $u \in L_2([0, T], U)$. Then $\| \Sigma \|_{\infty}$ is the smallest number $\gamma$ for which $\| \Sigma \|_{\infty} \leq \gamma$. So in computing the $H_{\infty}$ norm, the main problem is to determine whether or not $\| \Sigma \|_{\infty} \leq \gamma$. This can be done using the Bounded Real Lemma.

Clearly, inequality (8.2) holds if the system (7.1) is dissipative with respect to the supply rate $w(u, y) = \gamma^2 |u|^2 - |y|^2$, i.e. if there exists a non-negative storage function $S$ such that $S(0) = 0$ and

$$S(x) \geq \sup_{T \geq 0, u \in L_2([0, T], U)} \left\{ S(x(T)) - \int_0^T \gamma^2 |u(r)|^2 - |y(r)|^2 \, dr : x(0) = x \right\}. \hspace{1cm} (8.3)$$

Conversely, if the system (7.1) is reachable from 0 and (8.2) holds, then (7.1) is dissipative. Thus determining whether or not $\| \Sigma \|_{\infty} \leq \gamma$ is equivalent to the solvability of (8.3). This in turn is equivalent to the solvability of the PDI

$$\sup_{u \in U} \{ \nabla_x S \cdot f(x, u) - (\gamma^2 |u|^2 - |h(x)|^2) \} \leq 0 \text{ in } \mathbb{R}^n, \hspace{1cm} (8.4)$$

thanks to Theorem 7.3.

This leads to the following theorem, which is essentially a version of the Bounded Real Lemma.

**Theorem 8.1.** Assume that the system (7.1) is reachable from 0. Then $\| \Sigma \|_{\infty} \leq \gamma$ if and only if there exists a non-negative storage function $S$ satisfying $S(0) = 0$ and the PDI (8.4).

The PDI (8.4) is a nonlinear analogue of a Riccati-type matrix inequality (see equation (7.3.10), Anderson & Vongpanitlerd (1973)). One difficulty in using Theorem 8.1 is the lack of uniqueness of storage functions, or equivalently, solutions of the PDI (8.4). This issue is well known for linear systems, where the matrix inequality has infinitely many solutions, and is discussed at length in Anderson & Vongpanitlerd (1973) together with techniques for finding solutions.

We now use the finite difference scheme discussed earlier to compute an approximation to a storage function. A finite difference analogue of the PDI (8.4) is the discrete
otherwise, choose \( \delta > 0 \) sufficiently small and iterating (8.6) forward in time until

\[
\sup_{\delta \in \delta \in \delta^0} \sup_{x \in x^0, x^0 \leq \{ \| \Sigma \|_\infty \leq \gamma \}} \| S^\delta(x) \| < \infty,
\]

for each \( R > 0 \) (where \( B(0, R) = \{ x : |x| \leq R \} \), then \( \| \Sigma \|_\infty \leq \gamma \).

**Proof.** We follow the general convergence technique described by Barles & Souganidis (1991). Define

\[
S(x) = \liminf_{\delta \to 0, x^0 \to x, x^0 \in (\mathbb{R}^n)^0} S^\delta(x).
\]

Then \( S \) is non-negative, l.s.c., and \( S(0) = 0 \). We now show that \( S \) satisfies the PDI (8.4).

Let \( \phi \in C^1(\mathbb{R}^n) \) and assume, without loss of generality, that \( S - \phi \) attains a strict local minimum at \( x_0 \). There is a subsequence \( x^\delta \), which we again index by \( \delta \), such that

\[
S^\delta(x^\delta) \to S(x_0), \quad x^\delta \to x_0
\]
as \( \delta \to 0 \), and \( S^\delta - \phi \) has a local minimum at \( x^\delta \in (\mathbb{R}^n)^\delta \). Then (8.5) and

\[
S^\delta(x) - S^\delta(x^\delta) \geq \phi(x) - \phi(x^\delta),
\]

for \( \delta \) small imply

\[
\sup_{u \in U} \left\{ \frac{1}{\delta} \sum_{i=1}^n \phi(x^\delta + \delta e_i) - \phi(x^\delta) \right\} f_i^\delta(x^\delta, u) - (\gamma^2 |u|^2 - |h(x^\delta)|^2) \leq 0.
\]

Send \( \delta \downarrow 0 \) to obtain

\[
\sup_{u \in U} \left\{ \nabla \phi \cdot f(x_0, u) - (\gamma^2 |u|^2 - |h(x_0)|^2) \right\} \leq 0.
\]

This proves that \( S \) satisfies (7.4). Hence by Theorem 8.1, we conclude that \( \| \Sigma \|_\infty \leq \gamma \).

In order to use Theorem 8.2, an effective numerical procedure is needed for solving the discrete inequality (8.5). As one might expect, there may be many solutions of (8.5).

Consider, for example, a value space iteration defined by

\[
\begin{align*}
S_k^0(x) &= \sup_{u \in U} \left\{ \sum_z p^k(x, z; u) S_{k-1}^0(z) - \Delta t(x) \left( \gamma^2 |u|^2 - |h(x)|^2 \right) \right\} \\
S_0^0(x) &= 0
\end{align*}
\]

for \( x \in \delta \mathbb{Z}^n, k = 0, 1, \ldots \). Then an approximate solution to the discrete inequality (8.5) can be obtained by fixing \( \delta > 0 \) sufficiently small and iterating (8.6) forward in time until a stationary solution is obtained.

The numerical scheme is summarised as follows.

**Step 1.** Select the discretization size \( \delta > 0 \).

**Step 2.** Set \( p = 0 \) and choose \( \gamma_0 > 0 \).

**Step 3.** Set \( \gamma = \gamma_p \) and iterate (8.6) forward in time.

**Step 4.** If a stationary solution is obtained, then \( \| \Sigma \|_\infty \leq \gamma \), and choose \( \gamma_{p+1} < \gamma \). Otherwise, choose \( \gamma_{p+1} > \gamma \).
Step 5. Repeat Steps 2—4, until a desired accuracy is achieved.

Step 6. If necessary, adjust the discretization size $\delta > 0$ and repeat.

To implement this scheme, the grid $\delta \mathbb{Z}^n$ must be truncated to a finite grid, say $D^\delta = D \cap \delta \mathbb{Z}^n$, where $D$ is a bounded domain in $\mathbb{R}^n$. Appropriate boundary conditions must be imposed, such as Neumann type: $\partial S/\partial \nu = 0$ on $\partial D$.

Other schemes can be obtained using different finite difference or finite element discretizations, or implicit versions, and also policy space iterations and acceleration methods could be used.

9. Example

Consider a simple one-dimensional example for which the $H_\infty$ norm can be calculated explicitly. In this example, $f(x) + g(x)u = -0.5x + u$, $h(x) = x$, and $U = \mathbb{R}$ define a linear state space system corresponding to the transfer function $\Sigma(s) = 1/(s + 0.5)$.

For $\gamma > 0$, solutions of the PDI (8.4), if they exist, are of the form

$$S(x) = Px^2,$$

where $P$ satisfies the Riccati inequality

$$-P + P^2/\gamma^2 + 1 \leq 0.$$  \hfill (9.1)

If $\gamma > 2$, all solutions of (9.1) are real and positive. If $\gamma < 2$, no real solutions exist. Therefore $\|\Sigma\|_\infty = 2$.

Figure 5 illustrates the use of the numerical scheme. In the simulations,

- $D = (-1, 1)$ and the condition $\partial S/\partial \nu = 0$ is imposed on the boundary $\partial D = \{-1, 1\}$.
- $\delta = 0.05$, and $D^\delta$ consists of 41 equally spaced points in the interval $[-1, 1]$.
- $U$ is approximated by 201 equally spaced points in the interval $[-50, 50]$.

If $\gamma > 2$, the finite difference algorithm converges and gives a solution of the discrete inequality (8.5), and so we conclude from Theorem 8.2 that $\|\Sigma\|_\infty \leq \gamma$. On the other hand, if $\gamma < 2$ the algorithm diverges. Thus one concludes $\|\Sigma\|_\infty = 2$.

Figure 5. The finite difference algorithm converges for $\gamma > 2$ and diverges for $\gamma < 2$. 
10. Example

The numerical scheme was applied to the two-dimensional nonlinear system with
\[ f_1 + g_1 u = x_2, \quad f_2 + g_2 u = -5.59x_1 - 7.59x_2 - \sqrt{x_1^2 + x_2^2} \sin x_1 + u, \quad h = x_1. \]
The algorithm diverges for \( \gamma \leq 0.17 \) and converges for \( \gamma \geq 0.23 \), and so \( \| \Sigma \|_{\infty} \in (0.17, 0.23) \).

11. Conclusion

Nonlinear PDEs and PDIs arise naturally in optimal control, dissipative systems, and nonlinear \( H_\infty \) theory. Because of difficulties relating to lack of smoothness, solutions are defined in the viscosity sense. Using the viscosity solution definition, one can give precise meaning to such equations and inequalities. In addition, this framework enables one to prove convergence of numerical approximation schemes. We remark, however, that using such schemes is limited in practice to problems of low state dimension, because of the infamous “curse of dimensionality”.

REFERENCES


