Monograph:

McCaffrey, D. and Banks, S.P. (1999) Lagrangian Manifolds, Viscosity Solutions and Maslov Index. Research Report. ACSE Research Report 749 . Department of Automatic Control and Systems Engineering

X

DEPARTMENT OF AUTOMATIC CONTROL AND SYSTEMS ENGINEERING

Research Report No. 749

Title - Lagrangian Manifolds, Viscosity Solutions and Maslov Index

Authors - D. McCaffrey and S.P. Banks

Abstract -

This paper considers a geometrical construction for stationary viscosity solutions based on Lagrangian manifolds. This has been proposed recently in the literature by M.V. Day. The construction involves a key assumption of Lipschitz continuity. We discuss in this paper how this assumption follows from the vanishing of the Maslov index on closed curves on the manifold, at least for low dimensional manifolds. This is naturally satisfied for large portions of stable and unstable manifolds corresponding to hyperbolic equilibrium points. We discuss examples where this situation arises including infinite time optimal control and nonlinear filtering. We also give an introduction to the Maslov index and outline the connections with canonical tunnelling operators and idempotent analysis approaches to deriving stationary solutions to Hamilton-Jacobi-Bellman equations.

Lagrangian Manifolds, Viscosity Solutions and Maslov Index*

D. McCaffrey[†] and S.P. Banks[‡]

3rd July, 1999

Abstract

This paper considers a geometrical construction for stationary viscosity solutions based on Lagrangian manifolds. This has been proposed recently in the literature by M.V. Day. The construction involves a key assumption of Lipschitz continuity. We discuss in this paper how this assumption follows from the vanishing of the Maslov index on closed curves on the manifold, at least for low dimensional manifolds. This is naturally satisfied for large portions of stable and unstable manifolds corresponding to hyperbolic equilibrium points. We discuss examples where this situation arises including infinite time optimal control and nonlinear filtering. We also give an introduction to the Maslov index and outline the connections with canonical tunnelling operators and idempotent analysis approaches to deriving stationary solutions to Hamilton-Jacobi-Bellman equations.

1 Introduction

This paper considers a geometrical approach to constructing stationary viscosity solutions to Cauchy problems involving Hamilton-Jacobi-Bellman (HJB) equations. This approach has recently been put forward by M.V. Day in [12]. The geometry involved is that of the Lagrangian manifolds in phase space on which the characteristic curves of the Cauchy problem lie. A certain regularity assumption, namely Lipschitz continuity, had to be made by Day in order to show that the function he constructs is a viscosity solution. The main point of this paper is to investigate how this assumption follows from the topological properties of the relevant Lagrangian manifold in the stationary case. There are several strands in the literature which have to be introduced in order to place this problem in its proper context. We do this now.

The first of these strands is optimal control and, more generally, Bolza type variational problems. Sufficient conditions for a solution to such a problem usually reduce to solving an HJB partial differential equation, see for instance

*This work was begun under the support of EPSRC grant number GR/J 75241 while the first author was a postdoc at the University of Sheffield.

[†]MDC Technology Ltd, Startforth Road, Riverside Park, Middlesbrough TS2 1PT, UK. email: caffreyd@mdctech.com

[‡]University of Sheffield, Dept. of Automatic Control and Systems Engineering, Mappin Street, Sheffield, S1 3JD, UK. email: s.banks@sheffield.ac.uk

Chapter 4 of [17] or Chapter 1 of [18]. In particular, optimising over a finite time horizon leads to a Cauchy problem where the initial condition (or alternatively the final condition) is determined by the initial cost term in the cost functional. The characteristic curves for the Cauchy problem are the trajectories of the Hamiltonian system corresponding to the HJB equation. In the case of optimal control, this Hamiltonian system is given by some form of Pontryagin's maximum principle. The well known difficulty in formulating such problems is that, even with smooth initial data, the solution to the HJB equation is generally non-smooth. The characteristic curves start to cross at a finite distance from the initial manifold and the solution then becomes multi-valued.

This leads to the second strand, namely viscosity solutions, for which the best introductory reference is probably still [10]. One of the main motivations for their introduction was to provide an acceptable definition of the sense in which the value function for an optimal control problem can be said to solve Bellman's equation - see for instance Theorem 5.1 of [18], Theorem 2.3 of [20] or Theorem 1.10 of [26].

As already stated, the aspect of a viscosity solution which concerns us here is its connection with the geometry of the underlying Lagrangian manifold in phase space. This connection has already been explored in several places, so in order to make clear the interesting features of Day's construction, it is worth considering briefly how it differs from these other contributions.

Firstly, Hopf's original formula (Theorem 5a of [28]) for a generalised solution to an HJB equation in the case of a convex Hamiltonian is essentially geometric - the value of the solution at a point x is related to the value of the initial condition at a point y, the relationship being that y is the initial point on the minimising trajectory through x for the variational problem giving rise to the HJB equation. This formula was shown by Evans in Theorem 6.1 of [16] to be a viscosity solution. In this finite time variational setting, Lipschitz continuity of the generalised solution follows immediately from the corresponding Lipschitz continuity of the initial condition. In Examples 4.1.1 and 4.1.2 of [12], Day shows that his construction gives the same function as Hopf's formula. The Lipschitz property he requires in order to deduce that it is a viscosity solution follows from the above variational interpretation. However as the function constructed is already known to be a viscosity solution, this example doesn't expose the real application of his construction, which is to infinite time variational problems.

The other recent exploration of the geometry of viscosity solutions has been in papers such as [8] which have considered how to express classical second order conjugate point type necessary and sufficient conditions in the viscosity setting. These results essentially identify the points at which nearby trajectories start to cross and are no longer locally optimal. Again the analysis studies the evolution of an initial non-smooth manifold along the trajectories of a Hamiltonian system.

The interesting aspect of Day's construction is that it holds in the absence of any variational interpretation. In some sense it is independent of the evolution of initial manifolds along characteristics and it furnishes information on global optimality as opposed to local optimality. He makes this point himself in his paper, but we hope here to give an indication of the sense in which this is true.

This brings us to the third strand of the introduction. The particular case where Day's construction is of interest is in looking at stationary solutions to HJB Cauchy problems. In this case one cannot deduce the Lipschitz continuity of the solution from the evolution of a Lipschitz continuous initial condition, as the initial condition is the solution. No more is known about one than the other. Instead stationary solutions are usually arrived at as some sort of limit of solutions to finite time variational problems and regularity is difficult to prove. The key point of this paper is that the Lipschitz property actually follows from the topology of the relevant Lagrangian manifold in the stationary case.

To introduce this, we briefly review the various approaches to finding stationary solutions. In linear quadratic optimal control, given appropriate stabilizability and detectability conditions, stationary solutions are found as the limit of the value functions for finite time optimal control problems as the final time tends to infinity - see for instance [34]. The corresponding Hamiltonian dynamics in phase space (i.e. the dynamics arising from the maximum principle) are hyperbolic and the limiting value function for the infinite time problem is the generating function for the stable plane at the equilibrium point. This idea is extended to non-linear infinite time optimal control in [4, 27] and to non-linear H_∞ control in [37, 38]. The key point is that the stable manifold theorem implies the existence of a stable Lagrangian manifold whose tangent plane at the origin is the stable plane corresponding to the phase space dynamics of the linearised control problem. In the region around the origin where the stable manifold has a well-defined projection onto state space, the value function for the infinite time problem is smooth and is the generating function for the stable manifold.

The extension to viscosity solutions of infinite time optimal control problems and H_∞ problems is done in, for instance, [18, 26] and [35] respectively. However the approach is not explicitly geometric. The stable manifold approach, at least as far as viscosity solutions are concerned, has until now been stuck at the point at which the projection onto state space becomes ill-defined. The connection is probably in there implicitly in the above cited works in that they all consider an exponentially discounted infinite cost function while the stable manifold theorem gives exponential bounds on the approach to the equilibrium point. So a transformation between the two viewpoints may be possible.

However, Day's construction does make explicit the link between the stable manifold and the stationary viscosity solution in the case of infinite time optimal control and L_2 gain problems. (The full H_∞ problem is currently outside the scope of his construction as convexity of the Hamiltonian is required.) The relevant Lagrangian manifold in the stationary case is the stable manifold for the Hamiltonian system. This is simply connected - at least that portion of it which can be connected by Hamiltonian trajectories to a neighbourhood of the equilibrium point is. Closed curves on this manifold therefore have zero Maslov index. The purpose of this paper is to explain these terms, particularly Maslov index, and then show that for low dimensional cases this implies Lipschitz continuity of Day's construction and hence a stationary viscosity solution.

Before outlining the contents of each section, we mention one of the other major approaches to defining unique generalised solutions to HJB equations, namely idempotent analysis. This is based on the observation that HJB equations become linear when considered with respect to the arithmetic operations of (max,+) rather than $(+,\times)$, see [31, 24, 25, 22] for instance. This observation arose out of the study of logarithmic limits of short wave length asymtotic solutions to quantum tunnelling equations. It lead to the construction of quantum tunnelling canonical operators, usually based on heat transforms. We indicate in this paper the connection between Day's construction and a quantum tunnelling canonical operator based on the Laplace transform and also some connections between idempotent solutions and viscosity solutions in the stationary case.

We now outline the contents of the paper. The next section briefly introduces the required background ideas on Lagrangian manifolds, HJB equations and Day's construction. Section 3 then briefly introduces the various definitions of the Maslov index and illustrates the connection with optimal control by addressing some questions raised by van der Schaft in [39]. It also introduces the connection with idempotent solutions. Section 4 presents the main results of the paper - namely that for low dimensional stable Lagrangian manifolds, Day's construction is automatically Lipschitz. Hence, from his results, it is a stationary viscosity solution to the corresponding HJB equation. Lastly Section 5, suggests some applications to infinite time optimal control, L_2 gain problems, small noise asymptotics, nonlinear filtering and idempotent analysis.

2 Lagrangian manifolds and HJB equations

This section reviews Day's proposed construction of viscosity solutions to HJB equations. It starts with a very brief survey of the required elements of symplectic geometry. There are many references in the literature for this material. Day himself gives a nice introduction in Sections 1 and 2 of [12]. Another control theoretic perspective is given in [37, 38]. More detailed expositions from a classical mechanics or geometry perspective can be found in, for instance, [3, 29, 33, 36].

Define phase space to be the real $2n$-dimensional vector space with coordinates (x, p) where $x \in \mathbf{R}^n$ and $p \in \mathbf{R}^n$. (Much of the following applies to general symplectic spaces but our ultimate applications all live in $\mathbf{R}^{2n}$.) On phase space there exists a canonical two-form $\omega = dp \wedge dx$. Let $\phi : M \to \mathbf{R}^{2n}$ be an n-dimensional submanifold of phase space on which the restriction of the canonical two-form vanishes, i.e.

$$\phi^*(dp \wedge dx) \equiv 0.$$

The M is said to be a Lagrangian submanifold of phase space. This means that the one form pdx is locally exact on M.

Let I denote a subset of the set $\{1, \ldots, n\}$ and $\bar{I}$ denote its compliment. Let x^I denote the set of coordinates $\{x^i : i \in I\}$ and $p_{\bar{I}}$ denote the set $\{p_k : k \in \bar{I}\}$. Then it follows from the Lagrangian property that, at any point on M, there exists a collection of indices I such that $(x^I, p_{\bar{I}})$ form a local system of coordinates on M - see Section 2.1 of [33] or Proposition 4.6 of [29]. This means that M can be covered by an atlas of so-called canonical coordinate charts U_I in each of which its immersion into phase space is given by the relations

$$x^{\bar{I}} = x^{\bar{I}}\left(x^I, p_{\bar{I}}\right) \qquad p_I = p_I\left(x^I, p_{\bar{I}}\right). \tag{1}$$

Furthermore, it follows from the Lagrangian property that in each coordinate neighbourhood U_I, there exists a function $S_I(x^I, p_{\bar{I}})$ satisfying the equation

$$dS_I = \phi^*\left(p_I dx^I - x^{\bar{I}} dp_{\bar{I}}\right) \tag{2}$$

- see again Section 2.1 of [33] or Theorem 4.21 of [29]. (For brevity we cease to distinguish from now on between forms defined on M and on all of phase space.)

The above equation means that M can be locally represented in the form

$$x^{\bar{I}} = -\frac{\partial S_I}{\partial p_{\bar{I}}} \qquad p_I = \frac{\partial S_I}{\partial x^I}. \tag{3}$$

S_I is called a generating function for M .

Now consider a C^2 real valued Hamiltonian function $H(x,p)$ defined on phase space. Associated with H is a Hamiltonian vector field X_H defined by the relation

$$i(X_H)\omega = dH$$

or in coordinate terms

$$\dot{x} = \frac{\partial H}{\partial p} \qquad \dot{p} = -\frac{\partial H}{\partial x}.$$

It follows again from the Lagrangian property that if H is constant on M then the vector field X_H is tangent to M, i.e. M is invariant with respect to the Hamiltonian flow corresponding to H.

This is the principle reason why Lagrangian manifolds are so important. They are formed by collections of Hamiltonian trajectories or characteristic curves for the HJB equation

$$H(x, \partial S/\partial x) = 0 \tag{4}$$

where $\partial S/\partial x$ denotes the vector of partial derivatives of S with respect to x. Suppose $H(x,p) = 0$ on M. Then, on any of the specific coordinate neighbourhoods U_I where $I = \{1, \ldots, n\}$, it follows from the relation (3) above that the local solution to (4) is given by the relevant S_I, up to an additive constant. Note that a HJB equation of the form

$$H(t, x, \partial S/\partial x) = -\partial S/\partial t \tag{5}$$

can be transformed into one of the form (4) by taking t and $-H$ to be canonical coordinates. This is equivalent to considering the Lagrangian manifold traced out in $\mathbf{R}^{2n+2}$ phase space by the evolution under the Hamiltonian flow of the Lagrangian manifold given by the initial condition.

The solution given by the above method of characteristics is only local. Day's construction attempts to extend it beyond the coordinate neighbourhood in which it is defined. He makes the following basic hypotheses in order to obtain his construction.

1. M is a Lagrangian submanifold of $\mathbf{R}^{2n}$.

2. H is a C^2 real valued function on $\mathbf{R}^{2n}$ with $H(x,p) = 0$ for all $(x,p) \in M$.

3. pdx is globally exact on M.

4. M is embedded in $\mathbf{R}^{2n}$.

5. M is locally bounded.

6. M covers an open region Ω of state space $\mathbf{R}^n$ and has no boundary points over Ω.

The technical reasons for making these assumptions can be found in [12]. We only comment on two of them. Firstly, point 5. means that for each $x_0 \in \Omega$ there exists a $\delta > 0$ and $K < \infty$ such that $|p| \leq K$ for all $(x,p) \in M$ with $x \in B_\delta(x_0)$.

Secondly, the fact that M is Lagrangian means that the one-form pdx is closed when restricted to M. It thus defines a cohomomology class $[\phi^*(pdx)] \in H^1(M)$. Condition 3. says that this class is trivial, or equivalently that there exists a function $S(x,p)$ defined globally on M which satisfies $dS = pdx$. For our analysis later on, it is useful to express this in terms of the local generating functions on canonical coordinate neighbourhoods U_I defined above. Let Φ_I denote the restriction of S toU_I. Then the generating function $S_I(x^I, p_{\bar{I}})$ is defined as

$$S_I = \Phi_I - x^{\bar{I}} p_{\bar{I}}.$$

It follows that this function S_I satisfies (2).

Conversely, we can express condition 3. in terms of the local generating functions S_I. Let U_I and U_J be any two canonical coordinate charts with a non-trivial intersection. Let $I_1 = I \cap J$, $I_2 = I \backslash J$, $I_3 = J \backslash I$ and $I_4 = \{1, \ldots, n\} \backslash (I \cup J)$. Then condition 3. is equivalent to the equation

$$S_I - S_J = p_{I_2} x^{I_2} - p_{I_3} x^{I_3} \tag{6}$$

holding in $U_I \cap U_J$. For then if the functions

$$\Phi_I = S_I + x^{\bar{I}} p_{\bar{I}} \tag{7}$$

are defined in each of the respective charts U_I, it follows from (6) that they agree on pairwise intersections. Thus they glue togther to give a smooth function S defined on the whole of M which coincides with Φ_I on each U_I and which satisfies on each chart

$$dS = d\Phi_I = pdx. \tag{8}$$

Equation (6) is the reformulation in Cech cohomology of the requirement that the class $[pdx]$ be trivial. It is known in quantum mechanics as Maslov's first quantization condition. We will return to this point below.

Given the existence of a smooth function $S(x,p)$ defined globally on M and satisfying $dS = pdx$, Day then proposes the following function

$$W(x) = \inf\{S(x,p) : p \text{ such that } (x,p) \in M\} \tag{9}$$

as a viscosity solution. His actual result (Theorem 3 of [12]) says that if $H(x,p)$ is convex in p for each x and if $W(x)$ is (locally) Lipschitz continuous in Ω, then W is a viscosity solution of $H(x, \partial W(x)/\partial x) = 0$ in Ω. He also shows in Theorem 1 that without the Lipschitz condition, W is a lower semi-continuous viscosity supersolution. It is worth pointing out that condition 5. of his hypotheses ensure that the infimum in (9) is achieved for every $x \in \Omega$ - see [12], Section 2.1.2.

3 Maslov index and idempotent solutions

The point of this paper is to show that the required Lipschitz continuity can follow if all closed curves on the Lagrangian manifold M have Maslov index zero. This section outlines the various definitions of this index. These are illustrated along the way with some suggested applications to H_∞ control and with an analysis of the original application of the Maslov index - namely to the construction of global asymptotic solutions to linear PDEs in mathematical physics. This motivates a brief introduction to the idempotent analytical approach to solving HJB equations. In particular we show that Day's construction has already been used to obtain idempotent solutions to certain forms of stationary HJB equations. This holds out the possibility of showing a connection between viscosity solutions and idempotent solutions in the stationary case, a point returned to in the last section of the paper.

The simplest definition involves the points at which the projection of the Lagrangian manifold M onto state space $\mathbf{R}^n$ is singular. Note for later reference that the images of such singular points in state space are called caustic points. Let Σ denote the set of singular points on M. It in shown in [1] (see also, for instance, Appendix 11 of [3] or Theorem 7.6 of [29]) that, in the generic case, which can always be achieved by a small deformation, Σ is an $(n-1)$-dimensional two-sided cycle in M. This means that a positive and a negative side of Σ can be consistently defined. The defintion of the orientation goes as follows. Recall from above the notion (1) of canonical coodinate charts on M. In the neighbourhood of a simple singular point on Σ (i.e. one at which the rank of the projection onto state space drops by 1), a canonical system U_I can be chosen, where $I = \{1, \ldots, \hat{\imath}, \ldots, n\}$ for some $i \in \{1, \ldots n\}$. This means that M is represented in a neighbourhood of the singular point in the form

$$x^i = x^i(x^I, p_i) \qquad p_I = p_I(x^I, p_i).$$

Singular points near the given one are then defined by the condition $\partial x^i / \partial p_i = 0$. For M in general position (i.e. up to a small deformation), this derivative changes sign on passing from one side of Σ to the other in the neighbourhood of the given simple singular point. The positive side of Σ is then taken to be the side where this derivative is positive. It is shown in [1] that this definition is independent of the chosen coordinate system.

Given this orientation of Σ, the Maslov index of a curve γ on M is then defined to be

$$\operatorname{ind}(\gamma) = \nu_+ - \nu_- \tag{10}$$

where ν_+ is the number of points where γ crosses from the negative to the positive side of Σ and ν_- is the number of points where γ crosses from the positive to the negative side of Σ - see for instance Appendix 11 of [3] or Definition 7.7 of [29]. This definition assumes that the endpoints of γ are non-singular and that γ only intersects Σ transversely in simple singular points. It is then extended to any curve on M by approximating such a curve with one of the form γ - it can be shown that the definition is independent of the approximating curve. As an example, the Maslov index of the circle $x^2 + p^2$ traversed anti-clockwise in 2-dimensional phase space is $+2$.

The Maslov index is related to the Morse index from calculus of variations. Essentially, an extremal for a variational problem with Hamiltonian H corresponds to a phase curve in $\mathbf{R}^{2n}$ phase space. As described above for equation (5) this can be lifted to a phase curve lying on an $(n+1)$-dimensional Lagrangian manifold in $\mathbf{R}^{2n+2}$ phase space by considering its evolution in time under the

Hamiltonian flow. The Morse index of the extremal is equal to the Maslov index of the corresponding phase curve on the $(n+1)$-dimensional Lagrangian manifold - see [1], Theorem 5.2 or Appendix 11 of [3].

The main point of Arnold's paper [1] was to show that the Maslov index can in fact be defined as a topological invariant of M. Let $\mathcal{L}(\mathbf{R}^{2n})$ denote the Grassmanian manifold of Lagrangian subspaces of $\mathbf{R}^{2n}$. This is isomorphic to $U(n)/O(n)$ and its cohomology group $H^1(U(n)/O(n), \mathbf{R}) = \mathbf{R}$ has generator

$$g = (\det{}^2)^* \left[\frac{dz}{2\pi\sqrt{-1}z} \right]$$

where $\det^2 : U(n)/O(n) \rightarrow S^1$ is the square of the determinant of unitary matrices and $dz/(2\pi\sqrt{-1}z)$, $z \in S^1 = \{z \in \mathbf{C} : |z| = 1\}$ is the usual basis of 1-forms on S^1. There is a mapping

$$\psi : M \rightarrow \mathcal{L}(\mathbf{R}^{2n}) \tag{11}$$

defined by $\psi(m) = T_m(M)$ for each $m \in M$, i.e. each point m maps to the tangent space to M at m. The pullback of g under this mapping defines a characteristic class on M

$$g^* = \psi^*(g) \in H^1(M, \mathbf{R}).$$

For a closed curve γ on M, the evaluation of g^* on γ is equal to $\mathrm{ind}(\gamma)$ defined above - Theorem 1.5 of [1]. Again, a summary can be found in Appendix 11 of [3].

A different approach to defining the Maslov index as a characteristic class can be found in [36]. Arnold's definition above can be re-expressed in terms of a symplectic vector bundle on M and two Lagrangian sub-bundles on M. We will outline this approach for the case considered here where M is a Lagrangian submanifold of 2n-dimensional phase space, but it applies also to general symplectic manifolds. The required symplectic vector bundle on M is just $E = T\mathbf{R}^{2n}|_M$, i.e. the tangent bundle on 2n-dimensional phase space restricted to M. The first Lagrangian sub-bundle of E is defined as $L_0 = \{\text{planes } p_i = \text{ const. along } M\}$. The second Lagrangian sub-bundle is $L_1 = TM$, i.e. the tangent bundle on M.

The symplectic form ω on E is the restriction of the canonical two form $dp \wedge dx$ to E. Given this it is possible to choose a compatible positive complex structure J on E. This means that on each fibre E_m, $(m \in M)$ there is a linear mapping $J_m : E_m \rightarrow E_m$ such that $J_m^2 = -Id$ and $\omega_m(J_m v, J_m w) = \omega_m(v, w)$ for all $v, w \in E_m$. Defining $g_m(v, w) = \omega_m(v, J_m w)$ gives rise to a Hermitian metric on each fibre as follows

$$h_m(v, w) = g_m(v, w) - \sqrt{-1}\omega_m(v, w).$$

Over a trivialising neighbourhood $U \subset M$ for the bundle E, it is then possible to choose a complex orthonormal frame field for E, i.e. a n-dimensional field of vectors (e_i), $(i = 1, \ldots, n)$ such that $h(e_i, e_j) = \delta_{ij}$. Then the $2n$-dimensional field (e_i, Je_i) gives both a real g-orthonormal and a symplectic frame for E over U. The corresponding complex form

$$\epsilon_i = \frac{1}{\sqrt{2}} \left(e_i - \sqrt{-1} J e_i \right)$$

gives a complex unitary frame field for E over U. This is equivalent to a reduction of the structure group of E from the symplectic group $Sp(n, \mathbf{R})$ to the unitary group $U(n)$. Details of all this can be found in Chapters 2 and 3 of [36].

For each of the Lagrangian sub-bundles L_α, $(\alpha = 0, 1)$ it is in fact possible to choose complex unitary frames for E of the form

$$\epsilon_i^\alpha = \frac{1}{\sqrt{2}} \left(e_i^\alpha - \sqrt{-1} J e_i^\alpha \right)$$

such that (e_i^α), $(i = 1, \ldots, n)$ is a g-orthonormal frame for L_α over U, $(\alpha = 0, 1)$. Such fields are called L_α-related unitary frames. This corresponds to a further reduction of the structure group to $O(n)$, the orthogonal group. There then exists a transformation between the L_0 and L_1-related frames of the form

$$\epsilon_1 = \epsilon_0 A_U$$

where $A_U : U \to U(n)$ is a differentiable mapping. A change of L_α-related basis in the above corresponds to multiplying A_U on the left and right by orthogonal matrices. Hence, given a trivialising cover $\{U\}$ of M, the $\{A_U\}$ glue together into a global mapping

$$A(L_0, L_1) : M \to U(n)/O(n).$$

This is essentially the same mapping as the mapping ψ defined in (11) above. As there, the pull-back of the generator of cohomology on S^1 under the composition $\det^2 \circ A(L_0, L_1)$ defines a one-dimensional integral cohomology class $m(L_0, L_1)$ on M called the Maslov class of (L_0, L_1) - see Definition 4.3.1 of [36]. Vaisman uses this notation because his main point is to show that the Maslov class is a fundamental obstruction to the transversality of the sub-bundles L_0 and L_1 - i.e. the vanishing of $m(L_0, L_1)$ is a necessary condition for the transversality of L_0 and L_1. For if they are transversal, then J can be chosen so that $L_1 = JL_0$. Then $A_U = \sqrt{-1} I_n$ and so $m(L_0, L_1) = 0$.

There is in fact a whole series of such transversality conditions arising from the existence of higher dimensional Maslov classes for the triple (E, L_0, L_1). The definition of these classes uses the above mentioned fact that the existence of a Lagrangian sub-bundle L_0 of E corresponds to a reduction of the structure group to $O(n)$. With such a reduction can be defined an L_0-orthogonal unitary connection (Definition 4.4.14 of [36]). In the same way that a single connection gives rise to the Chern or Pontryagin classes of a vector bundle, so Vaisman shows how a pair of connections (one on L_0 and one on L_1) give rise to the secondary characteristic classes or Maslov classes of the pair of Lagrangian sub-bundles L_0 and L_1 of E - see Definitions 4.2.6 and 4.4.24 of [36]. There is a class in each of the dimensions $(4h - 3)$ for $h \in \mathbf{N}$ and the one-dimensional class coincides with $m(L_0, L_1)$ defined above. The vanishing of each of them is a necessary condition for the transversality of L_0 and L_1 - Theorem 4.4.26. Furthermore, for the particular case considered here where M is an immersed Lagrangian submanifold of standard $2n$-dimensional phase space $\mathbf{R}^{2n}$, all of these classes can be computed in terms of just the second fundamental form of M as a submanifold of $\mathbf{R}^{2n}$ - Theorem 4.5.11 and Proposition 4.5.12.

The importance of transversality lies in the fact that if $L_0 = \{\text{planes } x_i = \text{const.}\}$ and $L_1 = TM$ where M is a Lagrangian manifold on which $H(x, p) = 0$

for all $(x,p) \in M$, then transversality of L_0 and L_1 is a necessary condition for the existence of a smooth solution S to the HJB equation $H(x, \partial S/\partial x) = 0$ - Proposition 1.1.7 of [36]. We illustrate how this might apply in a control theoretic setting with reference to some results of van der Schaft on H_∞ control. In Corollary 3.6 and Proposition 4.3 of [39] and in Proposition 12 and Theorem 13 of [38], he assumes the existence of a diffeomorphism π from the stable Lagrangian manifold $N^- \subset T^*M$ of a Hamiltonian system onto the state space manifold M. The Hamiltonian arises from an L_2 gain or H_∞ control problem and the existence of π implies the existence of a smooth solution to the respective problem. The existence of such a diffeomorphism π is equivalent to the transversality of N^- and the vertical foliation $\mathcal{V}(T^*M)$ consisting of the tangent spaces to the fibres of T^*M. As discussed above the vanishing of the Maslov classes is necessary for the transversality of these two Lagrangian subbundles. In particular the study of the Maslov classes of stable manifolds might provide a starting point for investigating the following points made by van der Schaft in [39] - after Proposition 3.8 he says that "the condition that N^- is parameterizable by the x-coordinates is in general not easily checkable "while before Example 4.6 he says there is a need for "a really non-linear analysis to tell us if and where problems with the parameterization of the stable manifold N^- by the x-coordinates arise ". There are of course significant gaps to fill in in such an analysis. In particular, stable manifolds are not generally expressible in terms of immersions into phase space. They are usually known about in terms of the evolution of the Hamiltonian trajectories lying on them. This means that the straightforward calculation of Maslov classes in terms of second fundamental forms is not obviously applicable. The other problem is that the analysis would only give necessary conditions. This is because the requirement of smooth solutions is too restrictive. We show in the next section how the vanishing of the Maslov index is actually sufficient for the existence of a viscosity solution corresponding to the stable manifold - at least in low dimensions.

To end this section we consider one more approach to defining the Maslov index - see for instance [29], Definition 7.4. This involves the canonical coordinate charts defined above in (1). Let U_I and U_J be two canonical coordinate neighbourhoods of M and denote

$$I_1 = I \cap J, \quad I_2 = I \backslash J, \quad I_3 = J \backslash I, \quad I_4 = \{1, \ldots, n\} \backslash (I \cup J).$$

Then the coordinates on U_I are $(x^{I_1}, x^{I_2}, p_{I_3}, p_{I_4})$ and on U_J are $(x^{I_1}, x^{I_3}, p_{I_2}, p_{I_4})$. From the matrix whose determinant is the Jacobian of the change of coordinates on $U_I \cap U_J$, it is then possible to extract a symmetric non-degenerate sub-matrix. This in turn gives rise to an integer

$$c_{IJ}(M) = \operatorname{sgn} \frac{\partial \left(x^{I_2}, -p_{I_3}\right)}{\partial \left(p_{I_2}, x^{I_3}\right)}$$

on each intersection $U_I \cap U_J$. The c_{IJ} are locally constant and they define an integral 1-cochain of the covering $\{U_I\}$ in the sense of Cech cohomology. Details can be found in Section 2.3 of [33]. It is shown there that $\{c_{IJ}\}$ is the pre-image of a universal integral 1-cocycle on the Lagrangian Grassmannian $\mathcal{L}(\mathbf{R}^{2n})$ given by a certain covering of the Grassmannian. The universal integral 1-cocycle in question turns out to be $4g$, where g is the generator of $H^1(\mathcal{L}(\mathbf{R}^{2n}), \mathbf{R})$. The class defined by $\mu = \{c_{IJ}\}$ on M is therefore $1/4$ of the classes $m(L_0, L_1)$ or g^* defined above.

This form of the definition appears in the second Maslov quantization condition

$$\mu \equiv 0 (\text{mod } 8). \tag{12}$$

The first quantization condition has already been introduced above in equation (6). These two conditions appear as obstructions to the existence of global asymptotic solutions to linear pseudodifferential equations of the form

$$ih\frac{\partial\psi}{\partial t} = H\left(x, -ih\frac{\partial}{\partial x}\right)\psi, \quad \psi(x,0) = \exp\left(\frac{i}{h}S_0(x)\right)\phi_0(x) \tag{13}$$

as $h \to 0$. The construction of a local asymptotic solution is obtained by the famous Wentzel, Kramers and Brillouin (WKB) method. This involves looking for a solution of the form

$$\psi(x,t) = \exp\left(\frac{i}{h}S(x,t)\right)\phi(x,t).$$

The asymptotic phase S is the solution of the characteristic HJB equation

$$\frac{\partial S}{\partial t} + H\left(x, \frac{\partial S}{\partial x}\right) = 0, \quad S(x,0) = S_0(x).$$

The function $\phi(x,t)$ is the solution of a transport equation. This representation only holds up to the first focal point of the Hamiltonian flow corresponding to H with respect to the initial manifold $M_0 = \{x, \partial S_0/\partial x\}$ in phase space.

It can however be extended to a global representation as follows. Consider another canonical coordinate chart U_I on the Lagrangian manifold M formed by the evolution of M_0 along the Hamiltonian flow given by H. Let S_I denote the generating function of M in U_I. Let $\hat{p} = -ih\partial/\partial x$. Then there is a $1/h$-Fourier transform in x and p on phase space that takes the pseudolinear equation

$$ih\partial\psi/\partial t = H(x, \hat{p})\psi$$

to one for which

$$\psi_I(x^I, p_{\bar{I}}, t) = \exp\left(\frac{i}{h}S_I(x^I, p_{\bar{I}}, t)\right)\phi_I(x^I, p_{\bar{I}}, t)$$

is the local WKB solution. This is transformed back to a solution in coordinates $\{x^1, \ldots, x^n, t\}$ by an inverse $1/h$-Fourier transform. The construction of a global asymptotic solution on $\mathbf{R}^n \times \mathbf{R}$ then involves showing that the various representations are asymtotically equal on intersections $U_I \cap U_J$. The integrals involved in the inverse $1/h$-Fourier transforms are asymptotically expanded using the method of stationary phase. The first Maslov quantization guarantees that the phases S_I and S_J glue together into a global asymptotic phase. The method of stationary phase involves the square root of the Jacobian of the change of coordinates between U_I and U_J. This is a complex number. The second Maslov quantization condition guarantees that the branches of the square root function can be chosen in such a way that the arguments of the square root cancel out globally on M. Details of the above can be found in, for instance, Section 4.1 of [33]. A very readable survey of the subject is given in the Introduction of the same reference. The whole proceedure can be formulated without reference to Cauchy problems in terms of the so-called Maslov canonical operator. This maps functions ψ defined on M_0 globally to $\mathbf{R}^n$ and gives an asymptotic Green's function for equation (13).

The typical equation to which the above method is applied is the Schrodinger equation

$$ih\frac{\partial\psi}{\partial t} = \left(-\frac{h^2}{2}\Delta + V(x)\right)\psi.$$

The search for the low level asymptotic eigenfunctions of the Schrodinger operator leads to the study of the large time asymptotics of the equation

$$h\frac{\partial u}{\partial t} = \left(\frac{h^2}{2}\Delta - V(x)\right)u, \tag{14}$$

i.e. to the study of asymptotic quantum tunnelling solutions - see [14], Section 1 or [30]. The logarithmic asymptotics of this equation also turn up in the study of large deviation problems in probability. The canonical operator giving the global asymptotic (as $h \to 0$) solution of (14) is called the tunnelling canonical operator. It gives asymptotic solutions in which the principal term is of the form $\exp(-S(x,t)/h)$ where the entropy S is the generalised solution to

$$\frac{\partial S}{\partial t} + \frac{1}{2}\left(\frac{\partial S}{\partial x}\right)^2 - V(x) = 0, \quad S(x,0) = S_0(x). \tag{15}$$

This is therefore also the generalised solution to an optimisation problem. We outline how this generalised solution is arrived at.

The tunnelling canonical operator differs from that described above in that, in the neighbourhood of a focal point, the transformation between canonical charts on the Lagrangian manifold M corresponding to (15) is achieved by translating along the Hamiltonian flow given by $H_I = (1/2)\Sigma_{j\in I}p_j^2$ where $I \subset \{1,\ldots,n\}$. The corresponding transformation of the solution to (14) is therefore obtained by applying the solving operator for the heat equation. The resulting asymptotic integral expansion uses the Laplace method. This is simpler that the method of stationary phase. At any given point x, it identifies the branch of M on which the generating function S takes its minimum value over all the branches projecting onto x, rather than summing the contribution from all the branches. The first Maslov quantization condition is still required to ensure the existence of a globally defined entropy S on M. The second quantization condition is not needed however as the Jacobian entering into Laplace's method is assumed to be positive and so there are no problems in taking its square root.

The logarithmic limit as $h \to 0$ of the asymptotic solution to (14) gives the idempotent generalised solution to (15). This takes the form of a resolving operator R_t such that $S(x,t) = R_tS_0(x)$ is the generalised solution to (15) where R_t is defined on the set of functions bounded from below by the formula

$$R_tS_0(x) = \inf_{\xi}(S_0(\xi) + S(t,x,\xi)) \tag{16}$$

and $S(t,x,\xi)$ is the value of the cost functional for the variational problem corresponding to (15) along the minimising extremal starting at ξ at time $-t$ and ending at x at time 0. The operator R_t is linear with respect to the arithmetic operations $(\min, +)$, i.e.

$$\begin{aligned} R_t(\min(S_1, S_2)) &= \min(R_t S_1, R_t S_2), \\ R_t(\lambda + S(x)) &= \lambda + R_t S(x), \qquad \lambda \in \mathbf{R}. \end{aligned}$$

This is what is meant by an idempotent solution to (15). The fact that it is obtained by a linear resolving operator over an appropriate space allows the usual apparatus of analysis, i.e. weak solutions, distributions, Green's functions, convolutions, etc. and the corresponding numerical approximation schemes, to also be carried over to this space in order to find generalised solutions to HJB equations.

To see how the linearity arises, consider just the 1-dimensional heat equation with a small parameter

$$\frac{\partial u}{\partial t} = \frac{h}{2}\frac{\partial^2 u}{\partial x^2}, \tag{17}$$

i.e. equation (14) with no potential term. Linear superposition holds for this equation - if u_1 and u_2 are solutions then so is any linear combination of them. To find a local asymptotic solution, substitute $u = \exp(-S(x,t)/h)$. Simple manipulation shows that S satisfies

$$\frac{\partial S}{\partial t} + \frac{1}{2}\left(\frac{\partial S}{\partial x}\right)^2 - \frac{h}{2}\frac{\partial^2 S}{\partial x^2} = 0. \tag{18}$$

Although this is non-linear, it follows from the above substitution that if S_1 and S_2 are solutions of (18), then so is the combination

$$S_3 = -h \log\left(e^{-S_1/h} + e^{-S_2/h}\right). \tag{19}$$

The asymptotic solution to (17) is obtained by taking the limit as $h \to 0$. The corresponding limit of (18) is the HJB equation

$$\frac{\partial S}{\partial t} + \frac{1}{2}\left(\frac{\partial S}{\partial x}\right)^2 = 0 \tag{20}$$

and the limit of (19), for positive S_1 and S_2, is the superposition law

$$S = \min(S_1, S_2) \tag{21}$$

for solutions to the HJB equation (20). In other words, this equation is linear when considered with respect to the arithmetic operations of min and + instead of the usual operations of + and $\times$. This example appears in [31].

The connection between idempotent solutions and Day's construction can now be stated. See Section 5 of [14] for more details. Suppose that in equation (15), the potential V is a smooth nonnegative function with a bounded matrix of second derivatives. Suppose also that $V(x) \to \infty$ as $|x| \to \infty$ and that $V(x)$ is zero at just a finite number of points $\xi_1, \ldots \xi_l$ at each of which the matrix of second derivatives of V is non-singular. As already mentioned above, for the variational problem

$$J(c(.)) = \int_{-t}^{0} \left(\frac{1}{2}\dot{c}^2(\tau) + V(c(\tau))\right) d\tau$$

corresponding to (15), $S(t,x,\xi)$ denotes the value of the cost functional along the minimising extremal starting at ξ at time $-t$ and ending at x at time 0. Standard arguments show that the cost functional takes its minimum value on a smooth curve. Now define

$$S_k(x) = \inf\{S(t,x,\xi_k) : t > 0\} = \lim_{t\to\infty} S(t,x,\xi_k)$$

and

$$S(x) = \min\{S_k(x) : k = 1,\ldots,l\}.$$

It is pointed out in Proposition 1 of [14] that $S_k(x)$ and $S(x)$ are stationary idempotent solutions of (15). Furthermore, it is shown in Proposition 3 that $S_k(x)$ is the generating function for the essential parts of the unstable Lagrangian manifold through $(\xi_k, 0)$ - a non-essential point r on a Lagrangian manifold being one for which there is another point on the Lagrangian manifold with the same projection onto state space and at which the generating function has a lower value than at r. Thus it can be seen that the construction (9) considered by Day is already known to give stationary idempotent solutions and to relate them to the corresponding unstable manifold for the Hamiltonian flow when V satisfies the above assumptions.

In the same way as mentioned in the Introduction for viscosity solutions, the Lipschitz continuity of finite time idempotent solutions to the Cauchy problem (15) follows from the resolving operator representation (16) and the Lipschitz continuity of the initial condition S_0. However, this argument again fails to work for stationary idempotent solutions to (15). If we can show that Day's construction (9) is Lipschitz continuous then it will follow that, for the particular form of potential V considered here, stationary idempotent solutions to (15) coincide with stationary viscosity solutions.

We end this discussion of canonical tunnelling operators and idempotent solutions with the following remark. Recall from above that in the definition of the standard canonical tunnelling operator, the transformation between different canonical charts on the Lagrangian manifold defining the entropy is achieved with heat transforms. Day's construction (9) can be considered as the logarithmic limit of a canonical tunnelling operator. However in this case the transition between local representations for the entropy is achieved via the relation (6). This can be viewed as the logarithmic limit of a $1/h$-Laplace transform. In other words, the corresponding canonical tunnelling operator is constructed using Laplace transforms instead of heat transforms. The need to take account of the one-sidedness of the Laplace tansform manifests itself in the discussion in Section 1.4 and Section 5 of [12] concerning time-reversals, 'forward-backward' transformations and the question of whether the corresponding variational problem involves an inf or a sup type cost functional. We return to this point in the next section.

4 Lipschitz Continuity

We now show that, for low dimensional Lagrangian manifolds, Day's construction gives a viscosity solution if the corresponding Lagrangian manifold has zero Maslov index on closed curves. Recall from Section 2 that he makes the following six hypotheses.

Hypotheses 1 *1. M is a Lagrangian submanifold of $\mathbf{R}^{2n}$.*

2. H is a C^2 real valued function on $\mathbf{R}^{2n}$ with $H(x,p) = 0$ for all $(x,p) \in M$.

3. pdx is globally exact on M.

4. M is embedded in $\mathbf{R}^{2n}$.

5. M is locally bounded.

6. M covers an open region Ω of state space $\mathbf{R}^n$ and has no boundary points over Ω.

He then considers the function defined as

$$W(x) = \inf\{S(x,p) : p \text{ such that } (x,p) \in M\} \tag{22}$$

where $S(x,p)$ is the smooth function defined globally on M and satisfying $dS = pdx$ whose existence is implied by the above hypotheses. As has already been stated, condition (5) ensures that the infimum in (22) is achieved for every $x \in \Omega$. Theorem 3 of [12] then says that

Theorem 1 *If $H(x,p)$ is convex in p for each x and if $W(x)$ is (locally) Lipschitz continuous in Ω, then W is a viscosity solution of $H(x, \partial W(x)/\partial x) = 0$ in Ω.*

To show that W is Lipschitz in Ω, it is not necessary for us to consider every point in Ω. Recall that a point $x \in \Omega$ is called a caustic point if there exists a point $(x,p) \in M$ at which the projection π onto state space is singular. The definition of an essential point on M has already been made at the end of the last section. In the light of Day's hypotheses, an essential caustic point $x \in \Omega$ can be defined to be a caustic point for which π is singular at every $(x,p) \in M$ at which S achieves its minimum over x. Denote the set of essential caustic points by C_*. This is of course a subset of the set of caustics in Ω. Then it is shown in Theorem 2 of [12] that

Theorem 2 *W is continuous in $\Omega \backslash C_*$ and locally Lipschitz in the interior of $\Omega \backslash C_*$.*

For Lagrangian manifolds of dimension less than 6 in general position, the caustics or, more precisely, the singularities of the projection onto state space have been classified by Arnold - see Section 11 of [2] or Appendix 12 of [3]. General position can be achieved by an arbitrarily small perturbation in the class of Lagrangian manifolds. The approach of the rest of this section is therefore to go through the list of caustics and show that some cannot be essential, that the remaining set of essential caustics is closed and that W is Lipschitz at each such caustic. To achieve this we need the following hypotheses, in addition to those assumed by Day and listed above in Hypotheses 1.

Hypotheses 2 *1. The Maslov index of any closed curve on M is zero.*

2. There exists a neighbourhood U on M on which the projection π onto state space is non-singular and such that $\pi(U) \subset \Omega$.

3. For any $x \in \Omega$, there exists a curve lying on M which has Maslov index zero and which connects the point $(x, p^*) \in M$ to the neighbourhood $U \subset M$, where (x, p^*) is the point at which $S(x, .)$ achieves its minimum over all p such that $(x, p) \in M$.

4. Any other point (x, p) lying over $x \in \Omega$ at which $S(x, .)$ does not achieve its minimum can be connected to $U \subset M$ via a curve lying on M which passes through at most one singular point for the projection π of M onto state space.

5. For any $x \in \Omega$ and $p, q \in \mathbf{R}^n$ such that $(x, p) \in M$ and $(x, q) \in M$, then there exist n curves $\gamma_1, ..., \gamma_n$ lying on M such that each curve γ_i connects (x, p) to (x, q), doesn't go through the neighbourhood U and lies totally in the phase subplane (x^i, p_i), i.e. x^j and p_j are all constant on γ_i for $j \neq i$.

Some comments about the above are in order. As with Day's hypotheses, they are attempts to extract the geometrical essence of the problem independent of any variational interpretation. The applications given in the next section are all to infinite time variational problems for which the variational approach leads to difficult limiting arguments. The geometrical approach may prove useful. We take each of the hypotheses in turn.

Note first that Hypothesis (1) is equivalent to the vanishing of the first Maslov class $g^* = m(L_0, L_1) \in H^1(M, \mathbf{R})$. This will clearly be satisfied if M is simply connected. For finite time variational problems, if M_0 denotes the Lagrangian manifold corresponding to the initial or final cost term, then the $(n+1)$-dimensional Lagrangian manifold M traced out in time by the evolution of M_0 along the Hamiltonian flow is simply connected. As already mentioned, the Lipschitz property is however easily established in the finite time case via other arguments. Our interest will primarily be in infinite time variational problems and stationary HJB solutions corresponding to stable or unstable Lagrangian manifolds for hyperbolic equilibrium points. Stable and unstable manifolds are not in general simply connected. However, it is easy to construct large portions of them which are. This will be dealt with in the next section.

Hypothesis (2) is also natural in the context of finite time variational problems. With the same notation as the previous paragraph, U would be given by a small time neighbourhood of M_0 on M. Again, however, our interest will be in infinite time problems. In this case U would be a small neighbourhood of the equilibrium point on the stable or unstable manifold. Its existence is related to the existence of a smooth solution to the local linearised problem at the equilibrium point. The existence of this local smooth solution itself follows from the same assumptions on the linearised problem which guarantee that the equilibrium point is hyperbolic in the first place. Examples of such assumptions will be given in the next section.

Hypothesis (3) is again an attempt to generalise what is natural in the finite time context. The curve in question in the finite time case would be the characteristic curve for the Hamiltonian H which starts on M_0 and passes through the minimising point (x, p^*). The zero Maslov index condition follows from the differentiability of the value function along an optimal trajectory. This is well known for the finite time case provided the optimal trajectory is unique - see for instance [7, 5]. Points at which two optimal trajectories meet correspond to the formation of shocks in the evolution of M viewed as the graph in phase space of the vector function $p = p(t, x)$ - see Section 4.2 of [12] and [6]. At such points, the value function is no longer differentiable and the minimising point (x, p^*) in formula (22) jumps from one branch of M to another. The proofs below show that the jump cannot occur at a singular point in the projection of M onto state space. In the infinite time case, the candidate curve would again be the Hamiltonian characteristic through (x, p^*) lying on the stable or unstable manifold for the equilibrium point. However, we now have to assume that there are no singularities in the projection of this curve onto state space, or at least that they cancel out to give zero Maslov index.

Hypothesis (4) restricts attention to Lagrangian manifolds on which the non-optimal Hamiltonian characteristics have only passed through one singularity in the projection of M onto state space. It would appear possible to create very complicated patterns of caustics if the connecting curves are allowed to evolve far enough to pass through more than one singularity on M. The arguments presented here cannot deal with such cases.

Hypothesis (5) also restricts attention to Lagrangian manifolds on which the various branches evolve uniformly in some sense which will become clearer in the proofs of the results below. In the infinite time case, any two points on a stable or unstable manifold which project onto the same state space point can be pulled back along the Hamiltonian dynamics to the smooth neighbourhood U of the equilibrium point described in Hypothesis (2). The pulled back points can be connected via paths lying in U. Suppose paths can be found in U such that their images under the Hamiltonian flow satisfy Hypothesis (5). Then this is the sense in which the branches have to evolve uniformly. This issue doesn't arise in the finite time case as any two points on M projecting to the same point in $(n+1)$-dimensional state space are simultaneous and so are connected via curves lying in the image of M_0 under the Hamiltonian flow.

It is also, at this point, worth returning briefly to the issue of time-reversal mentioned at the end of the last section. Day's arguments apply equally well to show that for Hamiltonians $\hat{H}(x, p)$ which are concave in p, the function

$$\hat{W}(x) = \sup\{S(x, p) : p \text{ such that } (x, p) \in M\} \tag{23}$$

is a viscosity solution of $\hat{H}(x, \partial\hat{W}(x)/\partial x) = 0$ provided $\hat{W}$ is Lipschitz. This situation is normally associated with variational problems in which the cost functional involves a supremum rather than an infimum. A good example is the L_2-gain problem in non-linear systems theory as compared to the infinite time optimal control problem. These are both special cases of the H_∞ control problem described, for instance, in [35]. If the sign conventions for both problems are taken to be consistent with the overall sign convention for the min-max type H_∞ cost functional given in [35], then the L_2-gain cost functional involves a supremum over disturbances while the optimal control cost functional involves an infimum over controls. For both problems, the Hamiltonians are respectively concave and convex in the adjoint variable. Formulae (23) and (22) respectively give the solution to the HJB equation if the Lipschitz condition is satisfied. We will discuss the Lipschitz condition for these two examples in the next section.

The solutions to both the L_2-gain and the infinite time optimal control problem are usually arrived at by taking, respectively, the supremum or the infimum over all final times $T > 0$ of the corresponding family of finite time variational problems with fixed initial point and a final cost term at time T.

For Cauchy problems such as these with final data rather than initial data, the adjoint or momentum variable is related to the value function via $p = -\partial\hat{W}/\partial x$. This is well known. It is used by Day in Section 5 of his paper as the basis for a change of variables $(x,p) \to (x,-p)$ in phase space in order to convert the L_2-gain supremum cost functional and concave Hamiltonian into an infimum cost functional and convex Hamiltonian to which the infimum version of his construction (22) can be applied.

It also holds the key to removing the upstream discontinuity which he observes when applying his construction in Example 4.1.1 of his paper. In that example, he considers the variational problem

$$W(t,x) = \inf_{x(t)=x} \left\{ \Phi(x_0) + \int_0^t L(\dot{x}(s))ds \right\}$$

and applies his construction (22) on the $(n+1)$-dimensional Lagrangian manifold M formed by the evolution of the initial manifold $(x_0, \partial\Phi/\partial x_0)$ under the Hamiltonian flow. This gives a viscosity solution on the portion downstream from the initial manifold. He observes that (22) is discontinuous on the upstream portion. The convex Hamiltonian on M is given by

$$H^+(t,x,\sigma,p) = \sigma + H(p)$$

where H is the Legendre transformation of the convex integrand L and (t,σ) are the extra canonical coordinates required to extend $\mathbf{R}^{2n}$ phase space to $\mathbf{R}^{2n+2}$, as described after equation (5) above. On M, $H^+ = 0$, i.e. $\sigma = -H$.

On the upstream portion of M, the corresponding variational problem is

$$\begin{aligned} W(-t,x) &= \inf_{x(-t)=x} \left\{ \Phi(x_0) + \int_0^{-t} L(\dot{x}(s))ds \right\} \\ &= -\sup_{x(-t)=x} \left\{ \int_{-t}^{0} L(\dot{x}(s))ds - \Phi(x_0) \right\} \end{aligned}$$

where $t > 0$. So the cost functional is of supremum type, the value function to solve for is $\hat{W} = -W$ and it satisfies the final condition $\hat{W}(0,x_0) = -W(0,x_0) = -\Phi(x_0)$, which is the same as the initial condition for the down-stream portion. Also,

$$p = \partial W/\partial x = -\partial\hat{W}/\partial x$$

and

$$\sigma = -H = \partial W/\partial t = -\partial\hat{W}/\partial t.$$

Since W is a viscosity solution of $H^+(t,x,\partial W/\partial t, \partial W/\partial x) = 0$ on the downstream portion of M, it follows that $-W$ is a viscosity solution of

$$-H^+(t,x,-\partial W/\partial t, -\partial W/\partial x) = 0$$

on the same portion of M - see Remark 1.4 of [9]. Thus on the upstream portion of M, the correct and consistent Hamiltonian with respect to which $\hat{W}$ can be expected to be a viscosity solution is $-H^+(t,x,-\sigma,-p) = 0$. This is concave with respect to σ and p. It is clear from the plots given by Day that formula (23) is Lipschitz on the upstream portion of M (this also follows from the variational interpretation on this portion of M). Thus (23) gives a viscosity solution on the upstream portion.

So by using the appropriate construction (22) or (23) depending on whether the corresponding Hamiltonian is convex or concave, we obtain a viscosity solution on the whole of M and avoid the discontinuities observed by Day when only one form of the construction is applied. As mentioned above, this is related to which side of the $1/h$-Laplace transform is applied in the corresponding canonical tunnelling operator. We conjecture that Hypotheses 2 are sufficient to ensure that the appropriate side of the transform or the appropriate form of either (22) and (23) can be chosen in a consistent way on all of M.

However, the issue is not settled by the above discussion, as will become apparent in the proofs below. Our analysis in these proofs simply looks at the patterns of Lagrangian singularites which can develop given the above hypotheses. It would appear possible to construct patterns in which there exist multiple branches with the same state space projection and on which some of the branches have evolved from regions of M on which the min version of the construction applies, while others have evolved from regions on which the max version applies. At the point where these branches have come to overlap, there clearly exists a Lipschtiz viscosity solution, but its analytic definition requires a more subtle construction than either (22) or (23) - one involving the min over some branches and the max over others. We have not computed any specific examples, but it would seem to involve taking into account inflection points on M, i.e. the third partial derivatives of S, in the definition of W. We leave this for others with more insight and simply present some pictures below.

The following results apply to either form (22) or (23) of Day's construction. In the case of formula (23), the point (x, p^*) referred to in Hypothesis 2(3) is the point at which $S(x, .)$ achieves its maximum over all p such that $(x,p) \in M$, the point (x,p) referred to in Hypothesis 2(4) is any point at which $S(x, .)$ fails to achieve its maximum over x and an essential caustic is a caustic x for which π is singular at every $(x,p) \in M$ at which S achieves its maximum over x.

We now proceed to discuss the Lipschitz property for W given by (22) (or $\hat{W}$ given by (23) as appropriate) at all the caustic points corresponding to Lagrangian manifolds M in general position of dimension less than 6. Let n denote the dimension of M and assume in the following that Hypotheses 1 and 2 hold. The first case is trivial.

Proposition 3 *If $n = 1$ then W (or $\hat{W}$) is locally Lipschitz in the whole of Ω.*

Proof Hypothesis 2(3) restricts Ω to equal $\pi(U)$ and so there are no caustic points in Ω. C^* is thus empty. The result follows from Theorem 2. ∎

The next case is more involved.

Theorem 4 *If $n = 2$ then W (or $\hat{W}$) is locally Lipschitz in the whole of Ω.*

Proof There are two types of singularity on a Lagrangian manifold M in general position of dimension 2. Arnold denotes these as A_2 and A_3. They are of co-dimension 1 and 2 respectively. The corresponding caustics in 2-dimensional state space are curves on which the only singularities are semi-cubical cusps and points of transversal self-intersection. The cusps are the images of points on M of type A_3. All the other caustic points are images of points of type A_2. The self-intersection points are the image of more than one point of type A_2. There are no generic intersections of caustics corresponding to singularities of type A_2 and A_3 or A_3 and A_3 - i.e. there are no caustic points which are images of a point of type A_2 and one of type A_3 nor which are images of more than one point of type A_3.

We consider first singularities of type A_2. Suppose that the corresponding caustic is not a point of transversal self-intersection of caustic curves. Then such a caustic cannot appear in the set of essential caustics for M. This follows from the next theorem.

Theorem 5 *Let M be a Lagrangian manifold of dimension 2. Suppose (x,p) is a singular point of type A_2 on M and that this is the only singular point projecting onto x in state space. Then (x,p) cannot be a minimising or maximising point for $S(x,.)$ in formula (22) or (23) respectively.*

Proof In a neighbourhood on M of a singularity of type A_2 , coordinates (p_1, x_2) can be chosen on M (as described above in equation (1)) so that it has the generating function

$$G = \pm p_1^3 + \phi(x_2)$$

for some smooth function ϕ. From equation (3) it then follows that M is represented in phase space by the relations

$$x_1 = -\frac{\partial G}{\partial p_1} = \mp 3p_1^2 \quad p_2 = \frac{\partial G}{\partial x_2} = \frac{\partial \phi}{\partial x_2}. \tag{24}$$

Note this is the opposite of the sign convention used by Arnold, but it agrees with most of the other references quoted on symplectic geometry. The projection from M onto state space is thus of the form

$$(p_1, x_2) \to (\mp 3p_1^2, x_2).$$

The function S appearing in formula (22) or (23) is then given by equation (7)

$$S = G + x_1 p_1 = \pm p_1^3 + x_1 p_1 + \phi(x_2). \tag{25}$$

The singular points lie on the curve on M on which $x_1 = p_1 = 0$. We pick one such point, i.e. fix the value of x_2. We can thus ignore for the moment the smooth part ϕ of the function S as it doesn't affect the following. Using the relation $x_1 = \mp 3p_1^2$ given by the projection, we obtain the following relation between $S = \pm p_1^3 + x_1 p_1$ and x_1

$$S^2 = \mp \frac{4}{27} x_1^3.$$

We thus get the following picture Figure 1 for the multi-valued function S of $x_1$where the corresponding value of p_1 on each branch of S is indicated.

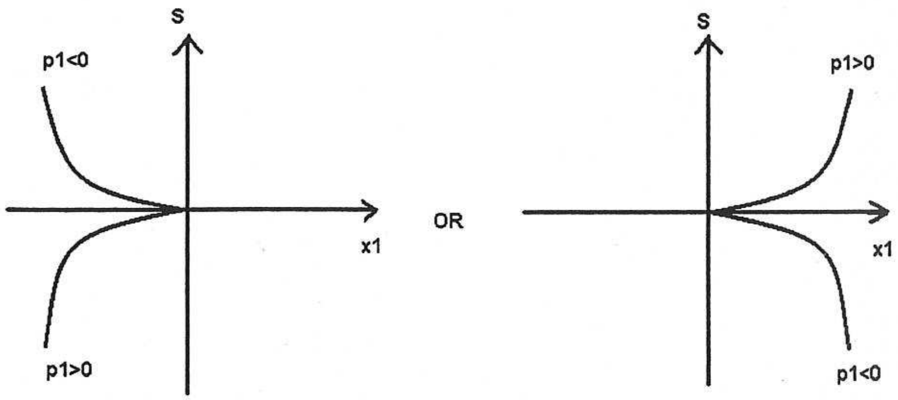

Figure 1

The associated Lagrangian manifold M therefore looks as in Figure 2 at the singularity.

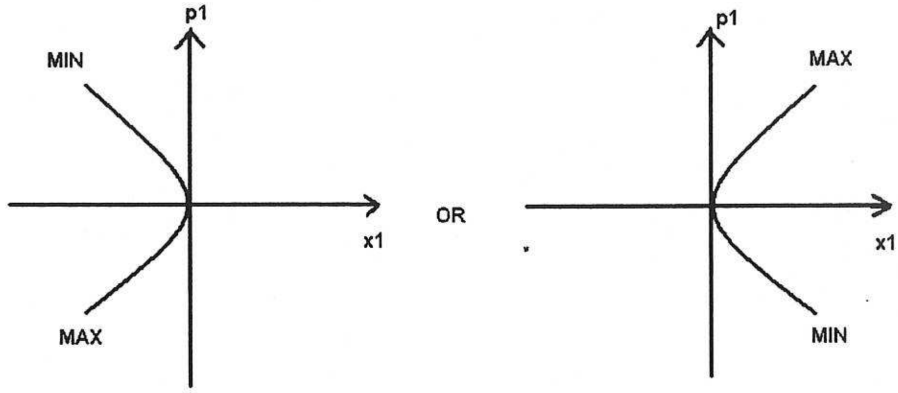

Figure 2

Depending on whether we are applying formula (22) or (23) forces us onto one or other branch of M as indicated in the picture. Suppose we are applying formula (22) and the singularity is of the form represented by the left hand picture in Figure 2 (the argument works equally well for the other possible choices). Then over the point $x_1 = x$, the minimising point for S on M is the point A shown in Figure 3. We show that the minimum for S cannot occur at $E1$.

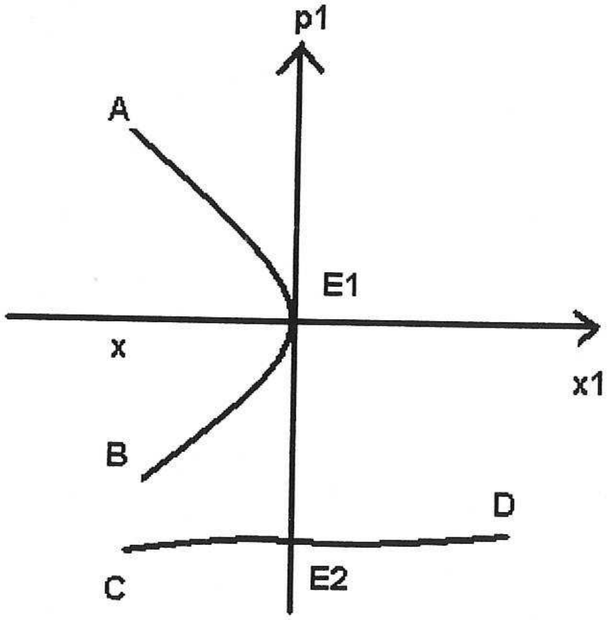

Figure 3

As we are assuming that the region Ω is open, there must exist another branch of M on which the minimising point for S lies for $x_1 > 0$. The projection of this branch onto state space must, by the hypotheses, be non-singular at $x_1 = 0$ and therefore in a neighbourhood round $x_1 = 0$. Its intersection with the (x_1, p_1) coordinate plane could lie above or below the branch $A \to E1 \to B$ in Figure 3. We draw it below for the moment and discuss the other choices later. Denote this branch $C \to E2 \to D$. Now the connecting curves, referred to in Hypothesis 2(3), from the minimising points A and D back to the neighbourhood U on M have Maslov index zero. The end points of these curves can in turn be joined by a curve lying in U of Maslov index zero since U has non-singular projection onto state space. Also, the points B and C can be chosen to have the same x_1 coordinate. So by Hypothesis 2(5) there is a curve γ from B to C lying on M, not passing through U and wholly contained within the (x_1, p_1) plane. Thus we have a closed path on M. Its Maslov index must be zero. Therefore the segment of this closed path made up of the union of the paths $A \to E1 \to B$ and γ must have index zero. The segment $A \to E1 \to B$ has index $+1$ if traversed in that direction - see formula (10). Thus γ must have index -1 when traversed in the same direction.

Further, suppose B' is any point on γ between B and C and distinct from both. Then the initial arc of γ from B to B' can never have Maslov index greater than zero. This is because the connecting curve from B' back to U given by either Hypothesis 2(3) or (4) passes through at most one singularity on M. If it passes through such a singularity, then it must be of type either A_2 or A_3. The Maslov index of a curve passing transversely through either type of singularity changes by ± 1, depending on the direction of travel. The total closed curve from U through $A \to E1 \to B$ and then B' back to U has index zero. The initial arc of γ from B to B' thus has index 0, -1 or -2 depending on whether the connecting curve from B' back to U has index -1, 0 or $+1$. The initial arc of γ from B to B' cannot therefore spiral in a clockwise direction as this leads to a positive index.

It is now a case of examining all possible paths γ from B to C of total index -1 on which any initial arc from B to B' has index 0, -1 or -2. This has to take into account the fact that γ could include the initial arc $B \to E1 \to A$ and could include the final arc $D \to E2 \to C$. It also has to take into account the case where the branch $C \to E2 \to D$ lies above the branch $A \to E1 \to B$.

We go through in detail the case where γ includes neither of the above mentioned initial or final arcs and where $C \to E2 \to D$ is as drawn in Figure 3. In this case all possible choices for γ give the same result as the curve shown in Figure 4.

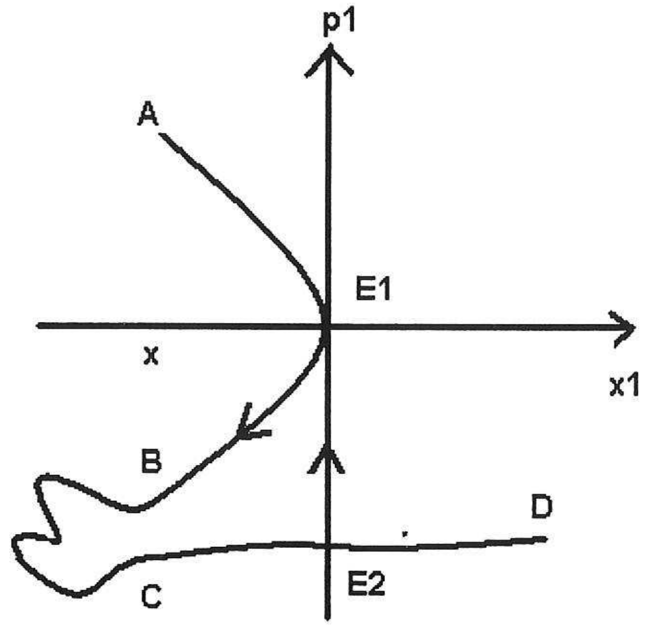

Figure 4

Now consider again the full function $S = p_1^3 + x_1 p_1 + \phi(x_2)$ and compare its value at $E1$ to that at $E2$. From the generating relations (24) and (25) above we see that

$$\begin{aligned} S(E2) - S(E1) &= \int_{(E1\to B)\cup\gamma\cup(C\to E2)} \sum_{i=1}^{2} p_i dx_i \\ &= \int_{(E1\to B)\cup\gamma\cup(C\to E2)} p_1 dx_1 + \int_{(E1\to B)\cup\gamma\cup(C\to E2)} p_2 dx_2. \end{aligned}$$

Now x_2 is constant on the curve $(E1 \to B) \cup \gamma \cup (C \to E2)$. So the second integral above is equal to zero.

To evaluate the first integral we let $E2 \to E1$ denote the path lying on the p_1-axis in Figure 4. Then since x_1 is constant on this path,

$$\int_{(E1\to B)\cup\gamma\cup(C\to E2)} p_1 dx_1 = \int_{(E1\to B)\cup\gamma\cup(C\to E2)} p_1 dx_1 + \int_{(E2\to E1)} p_1 dx_1$$

$$= \oint_{(E1 \to E1)} p_1 dx_1.$$

The integral round the closed curve is taken in the anti-clockwise direction as indicated by the arrows in Figure 4 and therefore is negative. It thus follows that

$$S(E2) < S(E1)$$

and so the minimising point for S on M cannot occur at the singularity $E1$. As we go to the right from $x_1 = x$ in Figure 4, the minimising point has to jump from the branch $A \to E1$ of M to another branch of M before we reach the caustic point at $x_1 = 0$.

It can be checked that the remaining cases where γ includes the initial arc $B \to E1 \to A$, where it includes the final arc $D \to E2 \to C$ and where the branch $C \to E2 \to D$ lies above the branch $A \to E1 \to B$ all lead to the same result, namely an anti-clockwise evaluation of the integral of pdx from the singularity $E1$ to some other point on M with the same x_1 coordinate. Thus $E1$ cannot be a minimising point for S on M in these cases either.

The only difficulty lies in the last case where $C \to E2 \to D$ lies above $A \to E1 \to B$. An example of a valid connecting curve γ satisfying the Maslov index restrictions described above is shown in Figure 5.

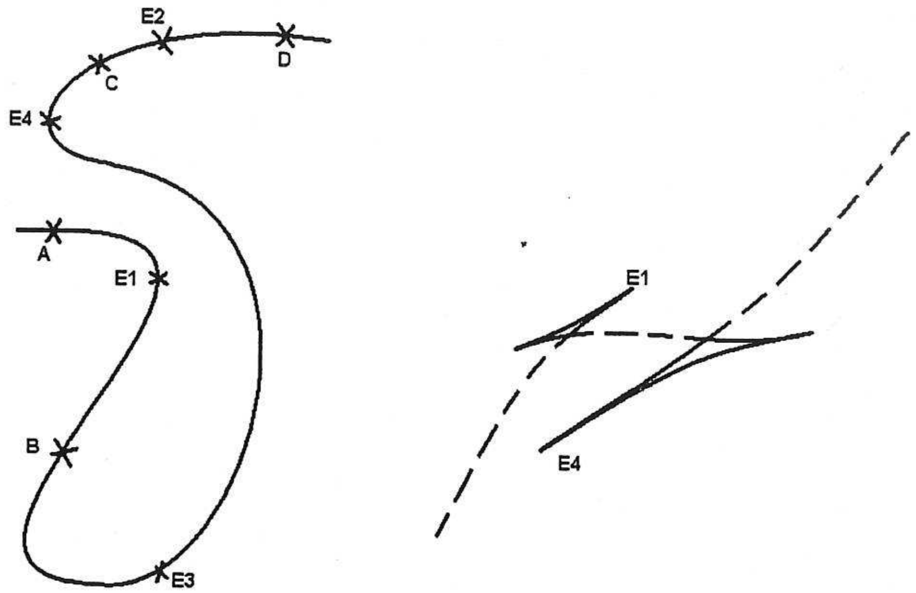

Figure 5

The above argument shows that the value of S is less at the point $E3$ on M than it is at the singularity $E1$, i.e. the minimum of S cannot occur at $E1$. This is what we set out to prove above. However, in this case, unlike in Figure 4, the max version of the construction (23) applies on the branch $C \to E2 \to D$ on which the viscosity solution continues for $x_1 > 0$. This is because the branch $A \to E1 \to E3$ has evolved from a region on which formula (22) applies while branch $E3 \to C \to D$ has evolved from a region where (23) applies. This is confirmed by looking at the multi-valued plot of S against x_1. It is clearly possible to pick out a Lipschitz viscosity solution as indicated by the dashed line in Figure 5. However, as mentioned above, defining it analytically requires a more subtle construction than either formula (22) or (23). We content ourselves here with having shown that, if you approach $E1$ along a curve lying on M on which the minimising points for S lie, then the minimising point has to jump to another branch before $E1$ is reached. Similarly, if you approach $E4$ along a curve lying on M on which the maximising points for S lie, then the maximising point has to jump to another branch before $E4$ is reached. ∎

Proof (of Theorem 4 resumed) Thus the set of essential caustics C^* can contain only points of transversal self-intersection corresponding to images of multiple points of type A_2 and semi-cubical cusp points corresponding to images of single points of type A_3. The set of such points is closed and so $\Omega\backslash C^*$ is open. Thus by Theorem 2, W (or $\hat{W}$) is locally Lipschitz in $\Omega\backslash C^*$. It thus remains to show that W (or $\hat{W}$) is locally Lipschitz in C^*.

We show next that the points of transversal self-intersection do not belong to C^*. We suppose that $E1$ is a singular point of type A_2 and that there exists at least one other such singularity with the same projection onto state space. Suppose, as above, that A is a minimising point for S over all points on M which project onto the same point in state space. We show that the minimum for S cannot occur at $E1$. An argument similar to that given above for the case of a single singular point of type A_2, shows that over the state space projection of $E1$, the intersection of M with the the (x_1, p_1)-plane looks as shown in Figure 6.

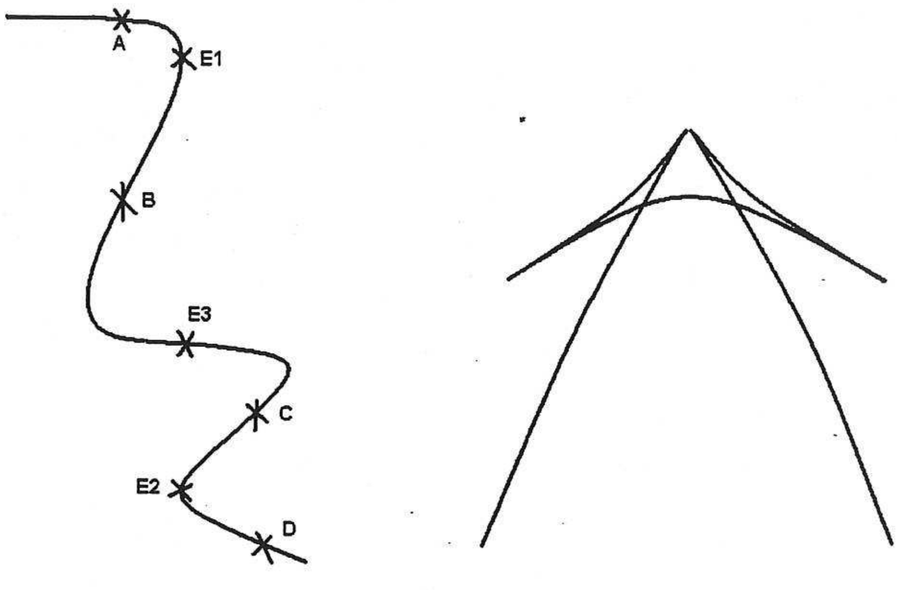

Figure 6

The branch $C \to E2 \to D$ can again go above or below the branch $A \to E1 \to B$. If it goes below, then an index counting argument shows the existence of a branch as shown on which there is a non-singular point $E3$ with the same projection as $E1$. Then, evaluating the integral of pdx in an anti-clockwise direction, as above, shows that $S(E3) < S(E1)$, as required. The corresponding multi-valued plot of S against x_1 is also shown in Figure 6.

Again there is a difficulty in the case where $C \to E2 \to D$ lies above $A \to E1 \to B$ as shown in Figure 7.

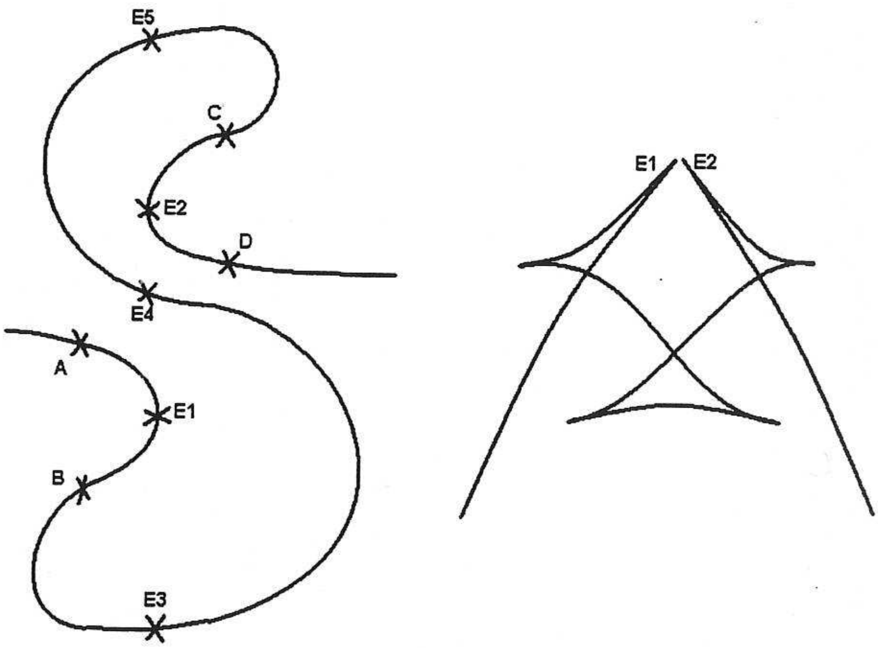

Figure 7

An index counting argument shows the existence of a connecting curve as shown. Then an anti-clockwise evaluation of the integral of pdx from $E1$ to $E3$ shows that, as $E1$ is approached from A, the minimum of S has to jump to another branch before $E1$ is reached, as required. However, the branch $E3 \to E4 \to E5$ has clearly evolved from one on which the maximum version (23) applies, while the branch $E5 \to C \to E2 \to D$ has evolved from one on which the minimum version (22) applies. This can be seen by looking at the multi-valued plot of S against x_1 shown in Figure 7. Clearly, a viscosity solution exists in the region where these branches overlap. As above, the definition of a version of (22) and (23) subtle enough to pick it out is left for the future.

Thus the set of essential caustics C^* can contain only semi-cubical cusp points corresponding to images of single points of type A_3. We show lastly that W (or $\hat{W}$) is locally Lipschitz at such points. The generating function for the Lagrangian manifold in a neighbourhood of such a singularity is given in the above cited references by Arnold - Section 11 of [2] or Appendix 12 of [3]. The singularity takes the form of a tuck. It differs from A_2 type singularities in that every point in a neighbourhood of the caustic in state space is the image of at least one point in a neighbourhood of the singularity on M.

To start with, we show that the minimising (or maximising) point for S cannot jump to another branch at an essential caustic point of type A_3. Suppose the contrary. The following argument is adapted from the proof of Theorem 2(a) of [12]. Let (x_0, p_0) denote the singularity of type A_3 and suppose that S achieves its minimum at (x_0, p_0) - the argument is the same for maximising points. Let M_0 denote a neighbourhood on M of the point (x_0, p_0) and let S_0 denote the value of S in M_0. Thus $W(x_0) = S_0(x_0, p_0)$. As there can be no other singularity projecting onto x_0, there must exist another branch of M, with non-singular state space projection at x_0, onto which the minimising point jumps. Let M_1 denote this branch, (x_0, p_1) denote the point on this branch which lies over x_0 and let S_1 denote the value of S in M_1. As M_1 has non-singular projection onto state space, S_1 is in fact a smooth function of x in a neighbourhood of x_0. Denote it by $G_1(x)$. Now consider a sequence of points $x_n \to x_0$ and $p_n \to p_1$ such that $(x_n, p_n) \in M_1$ and $W(x_n) = S_1(x_n, p_n) = G_1(x_n)$, i.e. approach x_0 along a curve in state space over which the minimising point for S lies on the non-singular branch M_1. As G_1 is a smooth function of x,

$$\lim_{x_n \to x_0} G_1(x_n) = G_1(x_0) = S_1(x_0, p_1).$$

Let $(x_n, p'_n) \in M_0$ be the corresponding sequence of points lying on the same branch of M as the singularity (x_0, p_0). Inspection of the A_3 generating function shows that $p'_n \to p_0$. Now it is pointed out in Theorem 1 of [12] that W is lower semi-continuous - this follows easily from Hypotheses 1. So then

$$W(x_0) \leq \liminf_{x_n \to x_0} W(x_n) = \lim_{x_n \to x_0} G_1(x_n) \leq \lim_{(x_n, p'_n) \to (x_0, p_0)} S_0(x_n, p'_n) = S_0(x_0, p_0).$$

where the last equality follows since S_0 is a smooth function on M. Thus

$$W(x_0) = S_0(x_0, p_0) = S_1(x_0, p_1),$$

i.e. the minimum value of S over x_0 is also achieved at the non-singular point $(x_0, p_1) \in M$. This contradicts the fact that x_0 is an essential caustic.

Thus, over a neighbourhood in state space of an essential caustic of type A_3, the minimising point for S lies on the same branch of M as the singularity itself. The generating function for this branch of M is given by Arnold. Simple inspection shows that the corresponding W is Lipschitz at the caustic. This completes the proof. ∎

We do not have time to go through the remaining cases of dimensions 3,4 and 5. We simply note here that the singularities that appear in dimensions 3 and 5 are similar to the singularity of type A_2 dealt with above in that the minimising point, if it occurs at the singularity, must jump to another branch in some directions of travel away from the corresponding caustic in state space. We conjecture that similar index counting arguments to the above would show that the minimising point cannot actually occur at the singularity and so these caustics cannot be essential. In the same way, the singularities appearing in dimension 4 are similar to A_3 in that the minimising point doesn't jump to another branch if it occurs at the singularity and so the Lipschitz property follows from inspection of the normal form of the relevant generating function as given by Arnold.

We end this section by noting that in Theorem 1 of [12], the supersolution property for formula (22) is proven using the assumption of convexity of the Hamiltonian $H(x, .)$. Similarly, the subsolution property for formula (23) follows if H is concave. However, as the above arguments in dimension 2 show that

W (or $\hat{W}$) is continuous in the whole of Ω, then the supersolution property (or respectively the subsolution property) follows directly from Proposition 1.3 of [10] without the assumption of convexity (respectively concavity). This leaves open the question of whether these assumptions can be removed from the other half of Day's argument showing that W (or $\hat{W}$) is a viscosity solution.

5 Proposed applications and further work

In this section we briefly outline how the above ideas might be applied to construct viscosity solutions to L_2-gain problems and infinite time optimal control problems. We also consider how viscosity solutions are related to quasi-potentials in small noise asymptotics and to conditional probability density functions in non-linear filtering. Lastly we indicate how they can be used to show the connection between stationary viscosity solutions and stationary idempotent solutions.

The main point is that all the above listed problems are some type of infinite time variational problem. The value function in each case is a stationary solution to the Hamilton-Jacobi-Bellman equation for the corresponding finite time variational problem. In each case there is an equilibrium point for the Hamiltonian dynamics. Given appropriate assumptions, this equilibrium point is hyperbolic. There then exist stable and unstable manifolds for the dynamics at this equilibrium point. These are Lagrangian manifolds. Construction (22) or (23) applied on the relevant manifold then gives a viscosity solution to the HJB equation. This solution is the value function. The key assumptions to be satisfied in applying (22) or (23) are Hypothesis 1(3) and 2(1) - the global exactness of pdx and the vanishing of the Maslov index on closed curves. This is done by restricting attention to that portion of the stable or unstable manifold which can be connected to a small neighbourhood of the equilibrium point by a Hamiltonian trajectory lying on the stable or unstable manifold. This portion can be pulled back to the small neighbourhood of the equilibrium point. As this is simply connected, then so is the portion of the stable or unstable manifold under consideration. Hence Hypothesis 1(3) and 2(1) are satisfied. A more precise definition of this submanifold of the stable or unstable manifold is given in hypotheses (M1), (M2) and (M3) of Section 5 of [12].

The first application area of L_2-gain problems in non-linear systems theory has been dealt with in detail by Day in [12]. As already mentioned, he assumes the Lipschitz property. We merely note here that Theorem 4 indicates how this follows from the fact that the Lagrangian manifold he uses is simply connected and so the Maslov index on closed curves on this manifold is zero.

The second area of infinite time optimal control involves problems of the form

$$V(\xi) = \inf_{u(.)\in L_2(0,\infty)} \int_0^\infty \frac{1}{2} \left(x(t)^T q(x(t)) x(t) + u(t)^T r(x(t)) u(t) \right) dt \qquad (26)$$

subject to $\dot{x} = f(x) + g(x)u$, $x(0) = \xi$, $\lim_{t\to\infty} x(t) = 0$ where $x \in \mathbf{R}^n$, $u \in \mathbf{R}^m$ and f, g, q and r are analytic functions of the appropriate dimensions. We assume there is an equilibrium at the origin, i.e. $f(0) = 0$, and that $q(x)$ is positive definite for $x \neq 0$ and $r(x)$ is positive definite for all x.

If we assume that the linearisation of (26) at the equilibrium is stabilizable and detectable, then the equilibrium is hyperbolic (Lemma 3 of [15]) . Thus there exists a stable Lagrangian manifold M on which formula (22) can be applied to obtain a Lipschitz viscosity solution to the HJB equation

$$\max_u \left\{ -\frac{\partial V}{\partial x}(f(x) + g(x)u) - \frac{1}{2}x^T q(x)x - \frac{1}{2}u^T r(x)u \right\} = 0. \tag{27}$$

Then similar arguments to those used by Day in Section 5 of [12] show that the solution so constructed is the value function V for the problem (26). This uses the subsolution property.

There remains the issue of synthesing the optimal control. We conjecture that it is given by $u = r^{-1}(x)g^T(x)\hat{y}$ where $(x, \hat{y})$ is the point on M at which $S(x, y)$ takes its minimum value over x. The argument is incomplete but we outline the following steps in it.

Note first that V is a supersolution of (27). So for all $p \in D^-V$, the subdifferential of V, we have

$$-p^T f + \frac{1}{2}p^T gr^{-1}g^T p - \frac{1}{2}x^T qx \geq 0$$

where, for a given $p \in D^-V$, the maximum in (27) is achieved by $u = -r^{-1}g^T p$. Then, if we consider the control $u = r^{-1}(x)g^T(x)\hat{y}$, the 'subderivative'of V along trajectories of the resulting controlled dynamics is

$$-p(f + gr^{-1}g^T\hat{y}) \geq -\frac{1}{2}pgr^{-1}g^T p + \frac{1}{2}x^T qx - pgr^{-1}g^T\hat{y}. \tag{28}$$

By Theorem I.14 of [9], V is strictly decreasing along trajectories of $\dot{x} = f + gr^{-1}g^T\hat{y}$ provided (28) is strictly positive for all $p \in D^-V$. Since $q(x)$ is positive definite for $x \neq 0$, this will follow provided

$$\frac{1}{2}pgr^{-1}g^T p + pgr^{-1}g^T\hat{y} = \frac{1}{2}(p + \hat{y})^T gr^{-1}g^T(p + \hat{y}) - \frac{1}{2}\hat{y}gr^{-1}g^T\hat{y} \leq 0 \tag{29}$$

for all $p \in D^-V$. Since V is Lipschitz, $D^-V \subset \partial V$, the generalised gradient (see [19]). Further, it is shown in the proof of Theorem 3 of [12] that for $V(x)$ of the form $\min\{S(x, y) : y \text{ such that } (x, y) \in M\}$, $\partial V \subset \text{co}\{-y : (x, y) \in M\}$ where co denotes the convex hull. Since the expression for p in (29) is convex, (29) will follow if

$$\frac{1}{2}(-y + \hat{y})^T gr^{-1}g^T(-y + \hat{y}) - \frac{1}{2}\hat{y}gr^{-1}g^T\hat{y} \leq 0.$$

for all y such that $(x, y) \in M$. If this inequality can be proved, then V is a Lyapunov function for the controlled dynamics, i.e. the controlled dynamics given by $u = r^{-1}(x)g^T(x)\hat{y}$ are asymptotically stable. A similar argument, with the same missing step at the end, shows that the cost function evaluated as $t \to \infty$ along the controlled dynamics is less than or equal to the value function at the initial point on the trajectory, i.e. the control is optimal. We leave the missing step as future work.

The third application area is to quasipotential functions arising in the study of small Brownian perturbations of dynamical systems. Section 1.5 of [12] gives a brief introduction showing that the problem can be put in the same form as the L_2-gain problem and defining the associated Hamiltonian. In [13], Day and Darden show that the Wentzel-Freidlin quasipotential can be constructed using (22) on the unstable Lagrangian manifold for the equilibrium point of the Hamiltonian system. As noted above, the vanishing Maslov class condition is easy to show on the unstable manifold for an equilibrium point. An interesting question for the future is whether the Maslov class vanishes on unstable manifolds associated with limit cycles. The relevance of this is that Day has shown in [11] that a quasipotential can be constructed using (22) applied on such an unstable manifold in the case where the basic dynamical system is the Van der Pol oscillator.

The fourth application area is to nonlinear filtering. Suppose we have a nonlinear, autonomous system driven by white noise with white noise corrupted observations

$$\begin{aligned} \dot{x}(t) &= f[x(t)] + G[x(t)]w(t) \\ z(t) &= h[x(t)] + v(t) \end{aligned} \tag{30}$$

where $w(t)$ and $v(t)$ are zero-mean, white and gaussian and uncorrelated with themselves and with $x(t_0)$ such that for $t > t_0$,

$$\begin{aligned} \text{cov}\{w(t), w(\tau)\} &= Q\delta(t-\tau) \\ \text{cov}\{v(t), v(\tau)\} &= R\delta(t-\tau). \end{aligned}$$

Define the processes $d\omega(t) = w(t)dt$, $d\nu(t) = v(t)dt$ and $dy(t) = z(t)dt$. Then the above system can be written more properly as the following Ito sense stochastic differential equations

$$\begin{aligned} dx(t) &= f[x(t)]dt + G[x(t)]d\omega(t) \\ dy(t) &= h[x(t)]dt + d\nu(t) \end{aligned}$$

where $\omega(t)$ and $\nu(t)$ are independent Browian motions uncorrelated with $x(t_0)$ such that

$$\begin{aligned} \text{cov}\{\omega(t), \omega(\tau)\} &= Q\text{min}(t,\tau). \\ \text{cov}\{\nu(t), \nu(\tau)\} &= R\text{min}(t,\tau). \end{aligned}$$

Let $Y(t) = \{y(\tau) : t_0 \le \tau \le t\}$ denote the observations up to time t. Let $\hat{x}(t) = \mathcal{E}\{x(t)|Y(t)\}$ denote the conditional mean, i.e. the minimum variance optimal estimate, and $V(t) = \text{var}\{x(t) - \hat{x}(t)|Y(t)\}$ denote the conditional error variance. Then to a first order approximation the solution to the filtering problem is given by the extended Kalman filter

$$\begin{aligned} d\hat{x}(t) &= f[\hat{x}(t)]dt + V(t)\frac{\partial h^T[\hat{x}(t)]}{\partial \hat{x}(t)}R^{-1}\{dy(t) - h[\hat{x}(t)]\}\,dt \\ dV(t) &= \{\frac{\partial f[\hat{x}(t)]}{\partial \hat{x}(t)}V(t) + V(t)\frac{\partial f^T[\hat{x}(t)]}{\partial \hat{x}(t)} + G[\hat{x}(t)]QG^T[\hat{x}(t)] \\ &\quad -V(t)\frac{\partial h^T[\hat{x}(t)]}{\partial \hat{x}(t)}R^{-1}\frac{\partial h[\hat{x}(t)]}{\partial \hat{x}(t)}V(t)\}dt. \end{aligned} \tag{31}$$

The initial conditions for the extended Kalman filter are $\hat{x}(t_0) = \mathcal{E}\{x(t_0)\}$ and $V(t_0) = \text{var}\{x(t_0)|Y(t_0)\}$.

Suppose now that $f(0) = 0$ and $h(0) = 0$, i.e. there is an equilibrium at $x = 0$. If the linearised system at the origin is completely controllable and completely observable, then the linear filter for the local problem is asymptotically stable with Lyapunov function

$$S(\hat{x}, t) = \frac{1}{2}\hat{x}^T V^{-1}(t)\hat{x}. \tag{32}$$

One of the key questions concerning the extended Kalman filter is under what conditions is it stable, i.e. under what conditions is the homogeneous part of the above dynamical system

$$d\hat{x}(t) = f[\hat{x}(t)]dt - V(t)\frac{\partial h^T[\hat{x}(t)]}{\partial \hat{x}(t)} R^{-1} h[\hat{x}(t)]dt$$

asymptotically stable? We put forward some observations based on the above ideas concerning viscosity solutions and Lagrangian manifolds.

First recall that the equations for the extended Kalman filter are derived from a first order approximate solution to the modified Fokker-Plank equation. This describes the evolution of the conditional probability density of $x(t)$. There is another approach based on formulating a least squares version of the problem of estimating $x(t)$ given the observations $z(t)$. This involves minimising the error function

$$\begin{aligned} J &= \frac{1}{2}(x(t_0) - \bar{x}(t_0))^T \bar{V}^{-1}(t_0)(x(t_0) - \bar{x}(t_0)) \\ &\quad + \frac{1}{2}\int_{t_0}^{t} \left[(z(\tau) - h(x(\tau)))^T R^{-1}(z(\tau) - h(x(\tau))) + w^T(\tau) Q^{-1} w(\tau) \right] d\tau \end{aligned}$$

subject to the dynamic constraint

$$\dot{x}(t) = f(x(t)) + G(x(t))w(t)$$

where $\bar{x}(t_0) = \mathcal{E}\{x(t_0)\}$ and $\bar{V}(t_0) = \text{var}\{x(t_0)\}$. This can be thought of as attempting to determine (estimate) $x(\tau)$ for $t_0 \le \tau \le t$ so that, simultaneously, the errors in the dynamical system and in the observations are small. In view of the constraint it is enough to minimise J with respect to $x(t_0)$ and $w(\tau)$ since $x(\tau)$ is then determined for $t_0 \le \tau \le t$. Let

$$S(x, t) = \min_{w(\tau)} J.$$

Then we get the dynamic programming equation

$$-\frac{\partial S}{\partial t} = \frac{\partial S}{\partial x}^T f(x) + \frac{1}{2}\frac{\partial S}{\partial x}^T G(x) Q G^T(x) \frac{\partial S}{\partial x} - \frac{1}{2}(z - h(x))^T R^{-1}(z - h(x)). \tag{33}$$

It can be shown that $-S(x,t)$ is the exponent of the conditional probability density function of $x(t)$. Thus $-S(x,t)$ can be interpreted as the likelihood of the state trajectory passing through x at time t, given the observations z made up to time t. The value of x which maximises $-S$ is thus the maximum likelihood estimate denoted $\hat{x}(t)$. If the dynamics are linear, e.g. for the local problem at the origin, this estimate coincides with the minimum variance estimate given by the Kalman filter. Further, the linear Kalman filter equations can be derived from (33) - see [23] Section 5.3 and Examples 7.11 and 7.12.

In general, for nonlinear dynamics, the maximum likelihood estimate does not coincide with the minimum variance estimate. Also it is not possible to derive an exact equation for $\hat{x}(t)$ from (33). However it is possible to derive a first order approximate equation for $\hat{x}(t)$ by expanding the various terms in (33) in Taylor series around $\hat{x}(t)$ in terms up to order 1. It turns out that these equations are the same as those for the extended Kalman filter - details are given in [32]. So the extended Kalman filter dynamics can be considered as a first order approximation to the nonlinear Hamiltonian system associated with the Hamilton-Jacobi-Bellman equation (33).

Now the assumptions on the equilibrium point at the origin imply that it is hyperbolic, so there exists an unstable manifold for the Hamiltonian dynamics. The tangent plane to this manifold at the origin is generated by the Lyapunov function for the stable local linear filter. (The switch between stable homogeneous filter dynamics and unstable Hamiltonian dynamics arises because the filter feedback is given by $V(t)$ while the Lyapunov function, and hence the Hamiltonian dynamics, are related to $V^{-1}(t)$.) This local Lyapunov function is a stationary solution to the linearisation of (33) at the origin. The corresponding stationary solution to (33) itself can be constructed by applying formula (22) on the unstable manifold to obtain a viscosity solution for the exponent of the conditional probability density.

The last step in the argument is to show that viscosity solutions to (33) which start at finite t_0 all converge to the stationary viscosity solution corresponding to the unstable manifold as $t_0 \to -\infty$. First order approximations to these solutions then provide Lyapunov functions for the homogeneous part of the extended Kalman filter dynamics. Convergence of the finite time viscosity solutions to the stationary solution requires conditions on the Hamiltonian corresponding to (33). The type of argument required is that given in Proposition 2 and the Corollary in Section 5 of [14] for a specific class of Hamiltonian.

This brings us to the last application area, namely the link between stationary viscosity solutions and stationary idempotent solutions. Results on general connections between viscosity and idempotent solutions have been announced in [21]. We simply repeat the observation already made at the end of Section 3 that the construction (22) has been shown to provide a stationary idempotent solution to a HJB equation for a specific class of potential functions - see Section 5 of [14]. Furthermore, this solution is related to the unstable manifolds of a set of hyperbolic equilibrium points. The arguments given in [12] show that the same construction gives a viscosity solution provided it is locally Lipschitz. This property will follow by restricting attention to simply connected submanifolds as then the Maslov index will vanish on closed curves on such submanifolds.

References

[1] V.I. Arnold, "Characteristic class entering in quantization conditions", *Functional Analysis and its Applications*, vol 1(1), pp 1–13, 1967.

[2] V.I. Arnold, "Normal forms for functions near degenerate critical points, the Weyl groups of A_k, D_k, E_k and Lagrangian singularities", *Functional*

Analysis and its Applications, vol 6(4), pp 3–25, 1972.

[3] V.I. Arnold, *Mathematical Methods of Classical Mechanics,* 2nd Ed., Springer-Verlag, New York, 1989.

[4] P. Brunovsky, "On the optimal stabilization of nonlinear systems", *Czech. Math. J.*,vol 18, pp 278–293, 1968.

[5] C. Byrnes and H. Frankowska, "Unicite des solutions optimales et absence de chocs pour les equations d'Hamilton-Jacobi-Bellman et de Riccati", *Comptes Rendues Acad. Sci. Paris*, vol 315, Serie I, pp 427–431, 1992.

[6] C. Byrnes, "On the Riccati partial differential equation for nonlinear Bolza and Lagrange problems", *Journal of Mathematical Systems, Estimation and Control,* vol 8, no 1, pp 1-54 , 1998.

[7] P. Cannarsa and H. Frankowska, "Some characterizations of optimal trajectories in control theory", *SIAM J. Control & Optimisation* , vol 29 No. 6, pp 1322–1347, 1991.

[8] N. Caroff and H. Frankowska, "Conjugate points and shocks in nonlinear optimal control", *Transactions of the American Mathematical Society*, vol 348(8), pp 3133–3153, 1996.

[9] M.G. Crandall and P.L. Lions, "Viscosity solutions of Hamilton-Jacobi equations", *Transactions of the American Mathematical Society*, vol 277(1), pp 1–42, 1983.

[10] M.G. Crandall, L.C. Evans and P.L. Lions, "Some properties of viscosity solutions of Hamilton-Jacobi equations", *Transactions of the American Mathematical Society*, vol 282(2), pp 487–502, 1984.

[11] M.V. Day, "Exit cycling for the Van der Pol oscillator and quasipotential calculations", to appear in *J. Dynamics and Differential Equations.*

[12] M.V. Day, "On Lagrange manifolds and viscosity solutions", to appear in *Journal of Mathematical Systems, Estimation and Control*, vol 8, 1998.

[13] M.V. Day and T.A. Darden, "Some regularity results on Ventcel-Freidlin quasi-potential function", *Applied Maths & Optimisation*, vol 13, pp 259–282, 1985.

[14] S.Yu. Dobrokhotov, V.N. Kolokoltsov and V.P. Maslov, "Quantization of the Bellman equation, exponential asymptotics and tunneling", in *Advances in Soviet Mathematics*, eds. V.P. Maslov and S.N. Samborskii, vol 13, pp 1–46, American Mathematical Society, Providence, Rhode Island, 1992.

[15] J.C. Doyle et al., "State space solutions to standard H_2 and H_∞ control problems", *IEEE Transactions on Automatic Control*, vol AC-34(8), pp 831–847, 1989.

[16] L.C. Evans, "Some min-max methods for the Hamlton-Jacobi equation", *Indiana Uni. Math. J.*, vol 33, pp 31-50, 1984.

[17] W.H. Fleming and R.W. Rishel, *Deterministic and stochastic optimal control,* Springer-Verlag, New York ,1975 .

[18] W.H Fleming and H.M. Soner, *Controlled markov processes and viscosity solutions,* Springer-Verlag, New York, 1993.

[19] H. Frankowska, "Hamilton-Jacobi equations: viscosity solutions and generalised gradients", *Journal of Mathematical Analysis & Applications*, vol 141, pp 21–26, 1989.

[20] H. Frankowska, "Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equations", *SIAM J. Control & Optimisation* , vol 31 No. 1, pp 257–272, 1993.

[21] M. Gondran, "Analyse MINPLUS", *Comptes Rendues Acad. Sci. Paris*, vol 323, Serie I, pp 371–375, 1996.

[22] J. Gunawardena (ed.), *Idempotency*, Publications of the Newton Institute, Cambridge University Press, Cambridge, UK, 1998.

[23] A.H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970.

[24] V.N. Kolokoltsov and V.P. Maslov, "Idempotent analysis as a tool of control theory and optimal synthesis I", *Functional Analysis & its Applications*, vol 23(1), pp 1–11, 1989.

[25] V.N. Kolokoltsov and V.P. Maslov, *Idempotent analysis and its applications*, Kluwer, Dordrecht, Holland, 1998.

[26] P.L. Lions, *Generalised solutions of Hamilton-Jacobi equations*, (Research Notes in Mathematics; 69), Pitman, London, 1982.

[27] D.L. Lukes, "Optimal regulation of nonlinear dynamical systems", *SIAM Journal on Control*, vol 7(1), pp 75–100, 1969.

[28] E. Hopf, "Generalized solutions of non-linear equations of first order", *J. Mathematics & Mechanics*, vol 14, pp 951–973, 1965.

[29] V.P. Maslov and M.V. Fedoriuk, *Semi-classical approximation in quantum mechanics*, D. Reidel, Dordrecht, Holland, 1981.

[30] V.P. Maslov, "Global exponential asymptotics of solutions of tunnel-type equations", *Soviet Math. Dokl.*, vol 24(3), pp 655-659, 1981.

[31] V.P. Maslov, "On a new principle of superposition for optimisation problems", *Russian Math. Surveys*, vol 42(3), pp43–54, 1987.

[32] D. McCaffrey, R. Harrison and S.P. Banks, "Asymptotically optimal nonlinear filtering", University of Sheffield, Department of Automatic Control and Systems Engineering, Research Report Number 734.

[33] A.S. Mishchenko, V.E. Shatalov and B.Yu. Sternin, *Lagrangian Manifolds and the maslov operator*, Springer-Verlag, Berlin, 1990.

[34] D.L. Russell, *Mathematics of Finite Dimensional Control Systems*, Dekker, New York, 1979.

[35] P. Soravia, "H_∞ control of nonlinear systems: differential games and viscosity solutions", *SIAM Journal on Control and Optimisation*, vol 34(3), pp 1071–1097, 1996.

[36] I. Vaisman, *Symplectic geometry and secondary characteristic classes*, (Progress in mathematics, vol 72), Birkhauser, Boston, 1987.

[37] A.J. van der Schaft, "On a state space approach to nonlinear H_∞ control", *Systems and Control Letters*, vol 16, pp 1–8, 1991.

[38] A.J. van der Schaft, "L_2 gain analysis of nonlinear systems and nonlinear state feedback H_∞ control", *IEEE Transactions on Automatic Control*, vol AC-37(6), pp 770–784, 1992.

[39] A.J. van der Schaft, "Nonlinear state space H_∞ control theory", in *Essays on control: perspectives in the theory and its applications*, eds. H.L. Trentelman and J.C. Willems, Birkhauser, Boston, 1993.