CHAPTER 2. BOUNDARY VALUE PROBLEM FOR HAMILTONIAN SYSTEMS

1. Rapid course in calculus of variations

In this preliminary section we present in a compact form the basic facts of the calculus of variations which are relevant to the asymptotical methods developed further. Unlike most standard courses in calculus of variations, see e.g. [Ak], [ATF], [GH], we develop primarily the Hamiltonian formalism in order to include in the theory the case of degenerate Hamiltonians, whose Lagrangians are singular (everywhere discontinuous) and for which in consequence the usual method of obtaining the formulas for the first and second variations (which lead to the basic Euler-Lagrange equations) makes no sense. Moreover, we draw more attention to the absolute minimum, (and not only to local minima), which is usually discussed in the framework of the so called direct methods of the calculus of variations.

A) Hamiltonian formalism and the Weierstrass condition. Let $H = H(x,p)$ be a smooth real-valued function on $\mathcal{R}^{2n}$. By "smooth" we shall always mean existence of as many continuous derivatives as appears in formulas and conditions of theorems. For the main results of this section it is enough to consider H to be twice continuously differentiable. Let $X(t,x_0,p_0), P(t,x_0,p_0)$ denote the solution of the Hamiltonian system

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial p}(x,p) \\ \dot{p} = -\frac{\partial H}{\partial x}(x,p) \end{cases} \tag{1.1}$$

with initial conditions (x_0,p_0) at time zero. The projections on the x-space of the solutions of (1.1) are called *characteristics of the Hamiltonian* H, or *extremals*. Suppose for some x_0 and $t_0 > 0$, and all $t \in (0,t_0]$, there exists a neighbourhood of the origin in the p-space $\Omega_t(x_0) \in \mathcal{R}^n$ such that the mapping $p_0 \mapsto X(t,x_0,p_0)$ is a diffeomorphism from $\Omega_t(x_0)$ onto its image and, moreover, this image contains a fixed neighbourhood $D(x_0)$ of x_0 (not depending on t). Then the family $\Gamma(x_0)$ of solutions of (1.1) with initial data (x_0,p_0), $p_0 \in \Omega_t(x_0)$, will be called *the family (or field) of characteristics starting from* x_0 *and covering* $D(x_0)$ *in times* $t \leq t_0$. The discussion of the existence of this family $\Gamma(x_0)$ for different Hamiltonians is one of the main topics of this chapter and will be given in the following sections. Here, we shall suppose that the family exists, and therefore there exists a smooth function

$$p_0(t,x,x_0) : (0,t_0] \times D(x_0) \mapsto \Omega_t(x_0)$$

such that

$$X(t,x_0,p_0(t,x,x_0)) = x. \tag{1.2}$$

The family $\Gamma(x_0)$ defines two natural vector fields in $(0,t_0] \times D(x_0)$, namely, with each point of this set are associated the momentum and velocity vectors

$$p(t,x) = P(t,x_0,p_0(t,x,x_0)), \quad v(t,x) = \frac{\partial H}{\partial p}(x,p(t,x)) \tag{1.3}$$

of the solution of (1.1) joining x_0 and x in time t.

Furthermore, to each solution $X(t,x_0,p_0), P(t,x_0,p_0)$ of (1.1) corresponds the action function defined by the formula

$$\sigma(t,x_0,p_0)=\int_0^t (P(\tau,x_0,p_0)\dot{X}(\tau,x_0,p_0)-H(X(\tau,x_0,p_0),P(\tau,x_0,p_0)))\,d\tau. \tag{1.4}$$

Due to the properties of the field of characteristics $\Gamma(x_0)$, one can define locally the *two-point function* $S(t,x,x_0)$ as the action along the trajectory from $\Gamma(x_0)$ joining x_0 and x in time t, i.e.

$$S(t,x,x_0)=\sigma(t,x_0,p_0(t,x,x_0)). \tag{1.5}$$

Using the vector field $p(t,x)$ one can rewrite it in the equivalent form

$$S(t,x,x_0)=\int_0^t (p(\tau,x)\,dx-H(x,p(\tau,x))\,d\tau), \tag{1.6}$$

the curvilinear integral being taken along the characteristic $X(\tau,x_0,p_0(t,x;x_0))$.

The following statement is a central result of the classical calculus of variations.

Proposition 1.1. *As a function of* (t,x) *the function* $S(t,x,x_0)$ *satisfies the Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t}+H(x,\frac{\partial S}{\partial x})=0 \tag{1.7}$$

in the domain $(0,t_0]\times D(x_0)$*, and moreover*

$$\frac{\partial S}{\partial x}(t,x)=p(t,x). \tag{1.8}$$

Proof. First we prove (1.8). This equation can be rewritten as

$$P(t,x_0,p_0)=\frac{\partial S}{\partial x}(t,X(t,x_0,p_0))$$

or equivalently as

$$P(t,x_0,p_0)=\frac{\partial \sigma}{\partial p_0}(t,x_0,p_0(t,x,x_0))\frac{\partial p_0}{\partial x}(t,x,x_0).$$

Due to (1.2) the inverse matrix to $\frac{\partial p_0}{\partial x}(t,x,x_0)$ is $\frac{\partial X}{\partial p_0}(t,x_0,p_0(t,x,x_0))$. It follows that equation (1.8) written in terms of the variables (t,p_0) has the form

$$P(t,x_0,p_0)\frac{\partial X}{\partial p_0}(t,x_0,p_0)=\frac{\partial \sigma}{\partial p_0}(t,x_0,p_0). \tag{1.9}$$

This equality clearly holds at $t = 0$ (both parts vanish). Moreover, differentiating (1.9) with respect to t one gets using (1.1) (and omitting some arguments for brevity) that

$$-\frac{\partial H}{\partial x}\frac{\partial X}{\partial p_0} + P\frac{\partial^2 X}{\partial t \partial p_0} = \frac{\partial P}{\partial p_0}\frac{\partial H}{\partial p} + P\frac{\partial^2 X}{\partial t \partial p_0} - \frac{\partial H}{\partial p}\frac{\partial P}{\partial p_0} - \frac{\partial H}{\partial x}\frac{\partial X}{\partial p_0},$$

which clearly holds. Therefore, (1.9) holds for all t, which proves (1.8).

To prove (1.7), let us first rewrite it as

$$\frac{\partial \sigma}{\partial t} + \frac{\partial \sigma}{\partial p_0}\frac{\partial p_0}{\partial t}(t, x) + H(x, p(t, p_0(t, x))) = 0.$$

Substituting for $\frac{\partial \sigma}{\partial t}$ from (1.4) and for $\frac{\partial \sigma}{\partial p_0}$ from (1.9) yields

$$P(t, x_0, p_0)\dot{X}(t, x_0, p_0) + P(t, x_0, p_0)\frac{\partial X}{\partial p_0}(t, x_0, p_0)\frac{\partial p_0}{\partial t} = 0. \tag{1.10}$$

On the other hand, differentiating (1.2) with respect to t yields

$$\frac{\partial X}{\partial p_0}(t, x_0, p_0)\frac{\partial p_0}{\partial t} + \dot{X}(t, x_0, p_0) = 0,$$

which proves (1.10).

We now derive some consequences of Proposition 1.1 showing in particular what it yields for the theory of optimisation.

Corollary 1. *The integral in the r.h.s. of (1.6) does not depend on the path of integration, i.e. it has the same value for all smooth curves $x(\tau)$ joining x_0 and x in time t and lying completely in the domain $D(x_0)$.*

Proof. This is clear, because, by (1.7) and (1.8), this is the integral of a complete differential.

In the calculus of variations, the integral on the r.h.s. of (1.6) is called the invariant Hilbert integral and it plays the crucial role in this theory.

Let the Lagrange function $L(x, v)$ be defined as the Legendre transform of $H(x, p)$ in the variable p, i.e.

$$L(x, v) = \max_p (pv - H(x, p)), \tag{1.11}$$

and let us define the functional

$$I_t(y(.)) = \int_0^t L(y(\tau), \dot{y}(\tau))\, d\tau \tag{1.12}$$

on all piecewise-smooth curves (i.e. these curves are continuous and have continuous derivatives everywhere except for a finite number of points, where the left and right derivatives exist) joining x_0 and x in time t, i.e. such that $y(0) = x_0$

and $y(t) = x$. Together with the invariant Hilbert integral, an important role in the calculus of variations belongs to the so called Weierstrass function $W(x,q,p)$ defined (in the Hamiltonian picture) as

$$W(x,q,p) = H(x,q) - H(x,p) - (q-p, \frac{\partial H}{\partial p}(x,p)). \tag{1.13}$$

One says that the Weierstrass condition holds for a solution $(x(\tau), p(\tau))$ of system (1.1), if $W(x(\tau), q, p(\tau)) \geq 0$ for all τ and all $q \in \mathcal{R}^n$. Note that if the Hamiltonian H is convex (even non-strictly) in the variable p, then the Weierstrass function is non-negative for any choice of its arguments, thus in this case the Weierstrass condition holds trivially for all curves.

Corollary 2. (Weierstrass sufficient condition for a relative minimum). *If the Weierstrass condition holds on a trajectory $X(\tau, x_0, p_0), P(\tau, x_0, p_0)$ of the field $\Gamma(x_0)$ joining x_0 and x in time t (i.e. such that $X(t, x_0, p_0) = x$), then the characteristic $X(\tau, x_0, p_0)$ provides a minimum for the functional (1.12) over all curves lying completely in $D(x_0)$. Furthermore $S(t, x, x_0)$ is the corresponding minimal value.*

Proof. For any curve $y(\tau)$ joining x_0 and x in time t and lying in $D(x_0)$ one has (from (1.11):

$$I_t(y(.)) = \int_0^t L(y(\tau), \dot{y}(\tau))\, d\tau \geq \int_0^t (p(t, y(\tau))\dot{y}(\tau) - H(y(\tau), p(\tau, y(\tau)))\, d\tau. \tag{1.14}$$

By Corollary 1, the r.h.s. here is just $S(t, x, x_0)$. It remains to prove that $S(t, x, x_0)$ gives the value of I_t on the characteristic $X(\tau, x_0, p_0(t, x.x_0))$. It is enough to show that

$$P(\tau, x_0, p_0)\dot{X}(\tau, x_0, p_0) - H(X(\tau, x_0, p_0), P(\tau, x_0, p_0))$$

equals $L(X(\tau, x_0, p_0), \dot{X}(\tau, x_0, p_0))$, where $p_0 = p_0(t, x, x_0)$, i.e. that

$$P(\tau, x_0, p_0)\frac{\partial H}{\partial p}(X(\tau, x_0, p_0), P(\tau, x_0, p_0)) - H(X(\tau, x_0, p_0), P(\tau, x_0, p_0))$$

$$\geq q\frac{\partial H}{\partial p}(X(\tau, x_0, p_0), P(\tau, x_0, p_0)) - H(X(\tau, x_0, p_0), q)$$

for all q. But this inequality is just the Weierstrass condition, which completes the proof.

Remark. In the more usual Lagrangian picture, i.e. in terms of the variables x, v connected with the canonical variables x, p by the formula $v(x,p) = \frac{\partial H}{\partial p}(x,p)$, the Weierstrass function (1.13) takes its original form

$$W(x, v_0, v) = L(x, v) - L(x, v_0) - (v - v_0, \frac{\partial L}{\partial v}(x, v_0))$$

and the invariant Hilbert integral (1.6) in terms of the field of velocities (or slopes) $v(t,x)$ (see (1.3)) takes the form

$$\int \frac{\partial L}{\partial v}(x,v)\,dx - \left((v, \frac{\partial L}{\partial v}(x,v)) - L(x,v)\right)\,dt.$$

Before formulating the next result let us recall a fact from convex analysis: if H is convex (but not necessarily strictly) and smooth, and L is its Legendre transform (1.11), then H is in its turn the Legendre transform of L, i.e.

$$H(x,p) = \max_v (vp - L(x,v)), \tag{1.15}$$

moreover, the value of v furnishing maximum in this expression is unique and is given by $v = \frac{\partial H}{\partial p}$. The proof of this fact can be found e.g. in [Roc]. In fact, we use it either for strictly convex H (with $\frac{\partial^2 H}{\partial p^2} > 0$ everywhere), or for quadratic Hamiltonians, and for both these cases the proof is quite straightforward.

Corollary 3. *If H is (possibly non-strictly) convex, then the characteristic of the family Γ joining x_0 and x in time t is the unique curve minimising the functional I_t (again in the class of curves lying in $D(x_0)$).*

Proof. From the fact of the convex analysis mentioned above, the inequality in (1.14) will be strict whenever $\dot{y}(\tau) \neq v(\tau, y)$ (the field of velocities v was defined in (1.3)), which proves the uniqueness of the minimum.

B) Conjugate points and Jacobi's theory. The system in variations corresponding to a solution $x(\tau), p(\tau)$ of (1.1) is by definition the linear (non-homogeneous) system

$$\begin{cases} \dot{v} = \frac{\partial^2 H}{\partial p \partial x}(x(\tau),p(\tau))v + \frac{\partial^2 H}{\partial p^2}(x(\tau),p(\tau))w, \\ \dot{w} = -\frac{\partial^2 H}{\partial x^2}(x(\tau),p(\tau))v - \frac{\partial^2 H}{\partial x \partial p}(x(\tau),p(\tau))w. \end{cases} \tag{1.16}$$

This equation holds clearly for the derivatives of the solution with respect to any parameter, for instance, for the characteristics from the family $\Gamma(x_0)$ it is satisfied by the matrices

$$v = \frac{\partial X}{\partial p_0}(\tau, x_0, p_0), \quad w = \frac{\partial P}{\partial p_0}(\tau, x_0, p_0).$$

The system (1.16) is called the Jacobi equation (in Hamiltonian form). One sees directly that (1.16) is itself a Hamiltonian system corresponding to the quadratic inhomogeneous Hamiltonian

$$\frac{1}{2}\left(\frac{\partial^2 H}{\partial x^2}(x(\tau),p(\tau))v, v\right) + \left(\frac{\partial^2 H}{\partial p \partial x}(x(\tau),p(\tau))v, w\right) + \frac{1}{2}\left(\frac{\partial^2 H}{\partial p^2}(x(\tau),p(\tau))w, w\right) \tag{1.17}$$

Two points $x(t_1), x(t_2)$ on a characteristic are called conjugate, if there exists a solution of (1.16) on the interval $[t_1, t_2]$ such that $v(t_1) = v(t_2) = 0$ and v does not vanish identically on $[t_1, t_2]$.

Proposition 1.2 (Jacobi condition in Hamiltonian form). *Suppose the Hamiltonian H is strictly convex and smooth. If a characteristic $x(\tau)$ contains two conjugate points $x(t_1), x(t_2)$, then for any $\delta > 0$, its interval $[x(t_1), x(t_2+\delta)]$ does not yield even a local minimum for the functional (1.12) among the curves joining $x(t_1)$ and $x(t_2+\delta)$ in time $t_2 - t_1 + \delta$.*

The standard proof of this statement (see any textbook in the calculus of variations, for example [Ak]) uses the Lagrangian formalism and will be sketched at the end of subsection D) in a more general situation. In Sect. 3, we shall present a Hamiltonian version of this proof, which can be used for various classes of degenerate Hamiltonians.

C) Connections between the field of extremals and the two-point function. These connections are systematically used in the construction of WKB-type asymptotics for pseudo-differential equations.

Let us first write the derivatives of $S(t, x, x_0)$ with respect to x_0.

Proposition 1.3. *Let the assumptions of Proposition 1.1 hold for all x_0 in a certain domain. Then*

$$\frac{\partial S}{\partial x_0}(t, x, x_0) = -p_0(t, x, x_0). \tag{1.18}$$

Moreover, as a function of t, x_0, the function $S(t, x, x_0)$ satisfies the Hamilton-Jacobi equation corresponding to the Hamiltonian $\tilde{H}(x, p) = H(x, -p)$.

Proof. If the curve $(x(\tau), p(\tau))$ is a solution of (1.1) joining x_0 and x in time t, then the curve $(\tilde{x}(\tau) = x(t-\tau),\ \tilde{p}(\tau) = -p(t-\tau))$ is the solution of the Hamiltonian system with Hamiltonian $\tilde{H}$ joining the points x and x_0 in time t. Both statements of the Proposition follow directly from this observation and Proposition 1.1.

Corollary 1. *If $\frac{\partial X}{\partial p_0}(t, x_0, p_0)$ is a non-degenerate matrix, then*

$$\frac{\partial^2 S}{\partial x^2}(t, x, x_0) = \frac{\partial P}{\partial p_0}(t, x_0, p_0)\left(\frac{\partial X}{\partial p_0}(t, x_0, p_0)\right)^{-1}, \tag{1.19}$$

$$\frac{\partial^2 S}{\partial x_0^2}(t, x, x_0) = \left(\frac{\partial X}{\partial p_0}(t, x_0, p_0)\right)^{-1}\frac{\partial X}{\partial x_0}(t, x_0, p_0), \tag{1.20}$$

$$\frac{\partial^2 S}{\partial x_0 \partial x}(t, x, x_0) = -\left(\frac{\partial X}{\partial p_0}(t, x_0, p_0)\right)^{-1}. \tag{1.21}$$

Proof. This follows from (1.2), (1.8), and (1.18) by differentiating.

Formula (1.19) combined with the following result, which is a consequence of (1.8) and Taylor's formula, can be used for the asymptotic calculations of S.

Corollary 2. *Let $\tilde{x}(t, x_0), \tilde{p}(t, x_0)$ denote the solution of (1.1) with initial data $(x_0, 0)$. Then*

$$S(t, x, x_0) = S(t, \tilde{x}, x_0) + (\tilde{p}(t, x_0), x - \tilde{x})$$

$$+\int_0^1 (1-\theta)\left(\frac{\partial^2 S}{\partial x^2}(t,\tilde{x}+\theta(x-\tilde{x}),x_0)(x-\tilde{x}),x-\tilde{x}\right)\,d\theta. \tag{1.22}$$

Finally let us mention here the formula for the "composition of Jacobians" after splitting of a characteristic. The function $J(t,x,x_0)=\det\frac{\partial X}{\partial p_0}(t,x_0,p_0)$ is called the Jacobian (corresponding to the family $\Gamma(x_0)$).

Proposition 1.4. *Under the assumptions of Proposition 1.3 let* $t_1+t_2\le t_0$ *and define*

$$f(\eta)=S(t_1,x,\eta)+S(t_2,\eta,x_0).$$

Denote $q=p_0(t_1+t_2,x;x_0)$ *and*

$$\tilde{\eta}=X(t_2,x_0,q),\quad \tilde{p}=P(t_2,x_0,q).$$

Then

$$\frac{\partial X}{\partial p_0}(t_1+t_2,x_0,q)=\frac{\partial X}{\partial p_0}(t_1,\tilde{\eta},\tilde{p})\frac{\partial^2 f}{\partial \eta^2}(\tilde{\eta})\frac{\partial X}{\partial p_0}(t_2,x_0,q).$$

In particular,

$$\det\frac{\partial^2 f}{\partial \eta^2}(\tilde{\eta})=J(t_1+t_2,x,x_0)J^{-1}(t_1,x,\tilde{\eta})J^{-1}(t_2,\tilde{\eta},x_0). \tag{1.23}$$

Proof. Let us represent the map $X(t_1+t_2,x_0,p_0)$ as the composition of two maps

$$(x_0,p_0)\mapsto(\eta=X(t_2,x_0,p_0),p_\eta=P(t_2,x_0,p_0))$$

and $(\eta,p_\eta)\mapsto X(t_1,\eta,p_\eta)$. Then

$$\frac{\partial X}{\partial p_0}(t_1+t_2,x_0,p_0)=\frac{\partial X}{\partial x_0}(t_1,\eta,p_\eta)\frac{\partial X}{\partial p_0}(t_2,x_0,p_0)+\frac{\partial X}{\partial p_0}(t_1,\eta,p_\eta)\frac{\partial P}{\partial p_0}(t_2,x_0,p_0)$$

For $p_0=q$ we have $\eta=\tilde{\eta},p_\eta=\tilde{p}_\eta$. Substituting in the last formula for the derivatives $\frac{\partial X}{\partial p_0}$ and $\frac{\partial X}{\partial x_0}$ in terms of the second derivatives of the two-point function by means of (1.19)-(1.21), yields the Proposition.

D) Lagrangian formalism. We discuss now the Lagrangian approach to the calculus of variations studying directly the minimisation problem for functionals depending on the derivatives of any order. Let the function L on $\mathcal{R}^{n(\nu+2)}$, the Lagrangian, be given, and assume that it has continuous derivatives of order up to and including $\nu+2$ in all its arguments. Consider the integral functional

$$I_t(y(.))=\int_0^t L(y(\tau),\dot{y}(\tau),...,y^{(\nu+1)}(\tau)\,d\tau) \tag{1.24}$$

on the class of functions y on $[0,t]$ having continuous derivatives up to and including order ν, having a piecewise-continuous derivative of order $\nu+1$ and satisfying the boundary conditions

$$\begin{cases} y(0)=a_0, & \dot{y}(0)=a_1, & ..., & y^{(\nu)}(0)=a_\nu, \\ y(t)=b_0, & \dot{y}(t)=b_1, & ..., & y^{(\nu)}(t)=b_\nu. \end{cases} \tag{1.25}$$

Such functions will be called admissible. Suppose now that there exists an admissible function $\bar{y}(\tau)$ minimising the functional (1.24). It follows that for any function $\eta(\tau)$ on $[0,t]$ having continuous derivatives of order up to and including ν, having a piecewise-continuous derivative of order $\nu+1$, and vanishing together with all its derivatives up to and including order ν at the end points $0, t$, the function

$$f(\epsilon) = I_t(y(.) + \epsilon\eta(.))$$

has a minimum at $\epsilon = 0$ and therefore its derivative at this point vanishes:

$$f'(0) = \int_0^t \left(\frac{\partial \bar{L}}{\partial y}\eta + \frac{\partial \bar{L}}{\partial \dot{y}}\dot{\eta} + ... + \frac{\partial \bar{L}}{\partial y^{(\nu+1)}}\eta^{(\nu+1)} \right) = 0,$$

where we denote $\frac{\partial \bar{L}}{\partial y^{(j)}} = \frac{\partial L}{\partial y^{(j)}}(\bar{y}(\tau), \dot{\bar{y}}(\tau), ..., \bar{y}^{(\nu+1)}(\tau))$. Integrating by parts and using boundary conditions for η, yields

$$\int_0^t \left[\left(\frac{\partial \bar{L}}{\partial \dot{y}} - \int_0^\tau \frac{\partial \bar{L}}{\partial y}\, ds \right) \dot{\eta}(\tau) + \frac{\partial \bar{L}}{\partial \ddot{y}}\ddot{\eta}(\tau) + ... + \frac{\partial \bar{L}}{\partial y^{(\nu+1)}}\eta^{(\nu+1)}(\tau) \right] d\tau = 0.$$

Continuing this process of integrating by parts one obtains

$$\int_0^t g(\tau)\eta^{(\nu+1)}(\tau)\, d\tau = 0, \tag{1.26}$$

where

$$g(\tau) = \frac{\partial \bar{L}}{\partial y^{(\nu+1)}}(\tau) - \int_0^\tau \frac{\partial \bar{L}}{\partial y^{(\nu)}}(\tau_1)\, d\tau_1 + \int_0^\tau \left(\int_0^{\tau_1} \frac{\partial \bar{L}}{\partial y^{(\nu-1)}}(\tau_2)\, d\tau_2 \right) d\tau_1 + ...$$

$$+(-1)^{\nu+1} \int_0^\tau ... \left(\int_0^{\tau_\nu} \frac{\partial \bar{L}}{\partial y}(\tau_{\nu+1})\, d\tau_{\nu+1} \right) ...d\tau_1. \tag{1.27}$$

Proposition 1.5 (Second lemma of the calculus of variations). *If a continuous function g on $[0,t]$ satisfies (1.26) for all η in the class described above, then g is a polynomial of order ν, i.e. there exist constants $c_0, c_1, ..., c_\nu$ such that*

$$g(\tau) = c_0 + c_1\tau + ... + c_\nu\tau^\nu. \tag{1.28}$$

Proof. There exist constants $c_0, c_1, ..., c_\nu$ such that

$$\int_0^t (g(\tau) - c_0 - c_1\tau - ... - c_\nu\tau^\nu)\tau^j\, d\tau = 0 \tag{1.29}$$

for all $j = 0, 1, ..., \nu$. In fact, (1.29) can be written in the form

$$c_0\frac{t^{j+1}}{j+1} + c_1\frac{t^{j+2}}{j+2} + ... + c_\nu\frac{t^{j+\nu+1}}{j+\nu+1} = \int_0^t g(\tau)\tau^j\, d\tau, \quad j = 0, ..., \nu.$$

The system of linear equations for $c_0, ..., c_\nu$ has a unique solution, because its matrix is of form (1.2.8) and has non-vanishing determinant given by (1.2.9). Let us set

$$\bar{\eta}(\tau) = \int_0^\tau \frac{(\tau - s)^\nu}{\nu!}[g(s) - c_0 - c_1 s - ... - c_\nu s^\nu]\, ds.$$

It follows that

$$\bar{\eta}^{(j)}(\tau) = \int_0^\tau \frac{(\tau - s)^{\nu - j}}{(\nu - j)!}[g(s) - c_0 - c_1 s - ... - c_\nu s^\nu]\, ds, \quad j = 0, ..., \nu,$$

and therefore, by (1.28), $\bar{\eta}$ satisfies the required conditions and one can take $\eta = \bar{\eta}$ in (1.26), which gives

$$\int_0^t (g(\tau) - c_0 - c_1\tau - ... - c_\nu \tau^\nu) g(\tau)\, d\tau = 0.$$

Using (1.29) one can rewrite this equation in the form

$$\int_0^t (g(\tau) - c_0 - c_1\tau - ... - c_\nu \tau^\nu)^2\, d\tau = 0,$$

which implies (1.28).

Equation (1.28), with $g(\tau)$ given by (1.27), is called the Euler equation in the integral form. Solutions of this equation are called extremals of functional (1.24). The following simple but important result was first stated by Hilbert in the case of the standard problem, i.e. when $\nu = 0$.

Proposition 1.6 (Hilbert's theorem on the regularity of extremals). Let $\bar{y}(\tau)$ be an admissible curve for problem (1.24) satisfying equation (1.27), (1.28), and let the matrix

$$\frac{\partial^2 \bar{L}}{(\partial y^{(\nu+1)})^2}(\tau) = \frac{\partial^2 L}{(\partial y^{(\nu+1)})^2}(\bar{y}(\tau), \dot{\bar{y}}(\tau), ..., \bar{y}^{(\nu+1)}(\tau))$$

be positive-definite for all $\tau \in [0, t]$. Then $\bar{y}(\tau)$ has continuous derivatives of order up to and including $2(\nu + 1)$; moreover, it satisfies the Euler differential equation

$$\frac{\partial L}{\partial y} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{y}} + \frac{d^2}{d^2\tau}\frac{\partial L}{\partial \ddot{y}} - ... + (-1)^{\nu+1}\frac{d^{\nu+1}}{d\tau^{\nu+1}}\frac{\partial L}{\partial y^{(\nu+1)}} = 0. \tag{1.30}$$

Proof. By the assumption of the Proposition and the implicit function theorem, one can solve equation (1.27),(1.28) (at least locally, in a neighbourhood of any point) for the last derivative $y^{(\nu+1)}$. This implies that $\bar{y}^{(\nu+1)}$ is a continuously differentiable function, moreover, differentiating (with respect to τ) the formula for $\bar{y}^{(\nu+1)}$ thus obtaines, one gets by induction the existence of the required number of derivatives. (In fact, if L is infinitely differentiable, one gets by this method that $\bar{y}$ is also infinitely differentiable.) Once the regularity of $\bar{y}$ is proved, one obtains (1.30) differentiating (1.27),(1.28) $\nu+1$ times.

Remark. The standard Euler-Lagrange equations correspond to (1.27), (1.28) and (1.30) with $\nu = 0$. It is worth mentioning also that if one assumes from the beginning that a minimising curve has $2(\nu+1)$ derivatives one can easily obtain differential equations (1.30) directly, without using the (intuitively not much appealing) integral equations (1.27), (1.28).

We now present the Hamiltonian form of the Euler-Lagrange equation (1.30) thus giving the connection between the Lagrangian formalism of the calculus of variations and the Hamiltonian theory developed above. For this purpose, let us introduce the canonical variables $x = (x_0, ..., x_\nu)$ by $x_0 = y, x_1 = \dot{y}, ..., x_\nu = y^{(\nu)}$ and $p = (p_0, ..., p_\nu)$ by the equations

$$\begin{cases} p_\nu = \frac{\partial L}{\partial y^{(\nu+1)}}(y, \dot{y}, ..., y^{(\nu)}, y^{(\nu+1)}), \\ p_{\nu-1} = \frac{\partial L}{\partial y^{(\nu)}} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial y^{(\nu+1)}}\right), \\ ... \\ p_0 = \frac{\partial L}{\partial \dot{y}} - \frac{d}{d\tau}\left(\frac{\partial L}{\partial \ddot{y}}\right) + ... + (-1)^\nu \frac{d^\nu}{d\tau^\nu}\left(\frac{\partial L}{\partial u^{(\nu+1)}}\right). \end{cases} \tag{1.31}$$

Proposition 1.7. Let $\frac{\partial^2 L}{(\partial y^{(\nu+1)})^2} \geq \delta$ everywhere for some $\delta > 0$. Then the equations (1.31) can be solved for $y^{(\nu+1)}, ..., y^{(2\nu+1)}$. Moreover, $y^{(\nu+l+1)}$ does not depend on $p_0, ..., p_{\nu-l-1}$, i.e. for all l:

$$y^{(\nu+l+1)} = f_l(x, p_\nu, p_{\nu-1}, ..., p_{\nu-l}). \tag{1.32}$$

Proof. Due to the assumptions of the Proposition, the first equation in (1.31) can be solved for $y^{(\nu+1)}$:

$$y^{(\nu+1)} = f_0(x_0, ..., x_\nu, p_\nu). \tag{1.33}$$

The second equation in (1.31) takes the form

$$p_{\nu-1} = \frac{\partial L}{\partial y^{(\nu)}}(x, y^{(\nu+1)}) - \frac{\partial^2 L}{\partial y^{(\nu+1)} \partial y} x_1 - ... - \frac{\partial^2 L}{\partial y^{(\nu+1)} \partial y^{(\nu-1)}} x_\nu$$

$$- \frac{\partial^2 L}{\partial y^{(\nu+1)} \partial y^{(\nu)}} y^{(\nu+1)} - \frac{\partial^2 L}{(\partial y^{(\nu+1)})^2} y^{(\nu+2)}.$$

One solves this equation with respect to $y^{(\nu+2)}$ and proceeding in the same way one obtains (1.32) for all l by induction.

The following fundamental result can be checked by direct calculations that we omit (see e.g. [DNF]).

Proposition 1.8 (Ostrogradski's theorem). *Under the assumptions of Proposition 1.7, the Lagrangian equations (1.30) are equivalent to the Hamiltonian system (1.1) for the canonical variables* $x = (x_0, ..., x_\nu), p = (p_0, ..., p_\nu)$, *with the Hamiltonian*

$$H = x_1 p_0 + ... + x_\nu p_{\nu-1} + f_0(x, p_\nu) p_\nu - L(x_0, ..., x_\nu, f_0(x, p_\nu)), \tag{1.34}$$

where f_0 is defined by (1.33).

The most important example of a Lagrangian satisfying the assumptions of Proposition 1.7 is when L is a quadratic form with respect to the last argument, i.e.

$$L(x_0, ..., x_\nu, z) = \frac{1}{2}(g(x)(z+\alpha(x)), z+\alpha(x)) + V(x). \tag{1.35}$$

The corresponding Hamiltonian (1.34) has the form

$$H = x_1 p_0 + ... + x_\nu p_{\nu-1} + \frac{1}{2}(g^{-1}(x)p_\nu, p_\nu) - (\alpha(x), p_\nu) - V(x). \tag{1.36}$$

To conclude, let us sketch the proof of Jacobi's condition, Prop. 1.2, for Hamiltonians of form (1.34) (which are degenerate for $\nu > 0$) corresponding to functionals depending on higher derivatives. One verifies similarly to Proposition 1.8 that the Jacobi equations (1.16), being the Hamiltonian equations for the quadratic approximation of the Hamiltonian (1.34), are equivalent to the Lagrangian equation for the quadratic approximation $\tilde{L}$ of the Lagrangian L around the characteristic $x(.)$. Moreover, this (explicitly time- dependent) Lagrangian $\tilde{L}(\eta, \eta', ..., \eta^{(\nu+1)})$ turns out to be the Lagrangian of the second variation of (1.24), i.e. of the functional

$$\frac{d^2}{d\epsilon^2}|_{\epsilon=0}(I_t(x(.)+\epsilon\eta(.)).$$

For this functional $\eta = 0$ clearly furnishes a minimum. However, if the point $x(s), s \in (0,t)$ is conjugate to $x(0)$, then a continuous curve $\bar{\eta}$ equal to a nontrivial solution of Jacobi's equation on the interval $[0,s]$ and vanishing on $[s,t]$ provides a broken minimum (with derivative discontinuous at s) to this functional, which is impossible by Proposition 1.6. Notice that we have only developed the theory for time-independent Lagrangians, but one sees that including an explicit dependence on time t does not affect the theory.

2. Boundary value problem for non-degenerate Hamiltonians

This section is devoted to the boundary value problem for the system (1.1) with the Hamiltonian

$$H(x,p) = \frac{1}{2}(G(x)p, p) - (A(x), p) - V(x), \tag{2.1}$$

where $G(x)$ is a uniformly strictly positive matrix, i.e. $G(x)^{-1}$ exists for all x and is uniformly bounded. The arguments used for this simple model (where the main results are known) are given in a form convenient for generalisations to more complex models discussed later. We first prove the existence of the field of characteristics $\Gamma(x_0)$, i.e. the uniqueness and existence for the local boundary value problem for system (1.1), and then the existence of the global minimum for functional (1.12), which also gives the global existence for the boundary value problem. Finally the asymptotic formulas for solutions are given. Before proceeding with the boundary value problem, one needs some estimates on the solutions of the Cauchy problem for the Hamiltonian system (1.1). For the case of the Hamiltonian (2.1), (1.1) takes the form

$$\left\{ \begin{array}{l} \dot{x} = G(x)p - A(x) \\ \dot{p}_i = -\frac{1}{2}(\frac{\partial G}{\partial x_i}p, p) + (\frac{\partial A}{\partial x_i}, p) + \frac{\partial V}{\partial x_i}, \quad i = 1, ..., m, \end{array} \right. \tag{2.2}$$

where we have written the second equation for each coordinate separately for clarity.

Lemma 2.1. *For an arbitrary $x_0 \in \mathcal{R}^m$ and an arbitrary open bounded neighbourhood $U(x_0)$ of x_0, there exist positive constants t_0, c_0, C such that if $t \in (0, t_0]$, $c \in (0, c_0]$ and $p_0 \in B_{c/t}$, then the solution $X(s, x_0, p_0), P(s, x_0, p_0)$ of (2.2) with initial data (x_0, p_0) exists on the interval $[0, t]$, and for all $s \in [0, t]$,*

$$X(s, x_0, p_0) \in U(x_0), \quad \|P(s, x_0, p_0)\| < C(\|p_0\| + t). \tag{2.3}$$

Proof. Let $T(t)$ be the time of exit of the solution from the domain $U(x_0)$, namely

$$T(t) = \min\left(t, \sup\{s : X(s, x_0, p_0) \in U(x_0),\ P(s, x_0, p_0) < \infty\}\right).$$

Since G, A and their derivatives in x are continuous, it follows that for $s \le T(t)$ the growth of $\|X(s, x_0, p_0) - x_0\|$, $\|P(s, x_0, p_0)\|$ is bounded by the solution of the system

$$\left\{ \begin{array}{l} \dot{x} = K(p + 1) \\ \dot{p} = K(p^2 + 1) \end{array} \right.$$

with the initial conditions $x(0) = 0, p(0) = \|p_0\|$ and some constant K. The solution of the second equation is

$$p(s) = \tan(Ks + \arctan p(0)) = \frac{p(0) + \tan Ks}{1 - p(0)\tan Ks}.$$

Therefore, if $\|p_0\| \le c/t$ with $c \le c_0 < 1/\tilde{K}$, where $\tilde{K}$ is chosen in such a way that $\tan Ks \le \tilde{K}s$ for $s \le t_0$, then

$$1 - \|p_0\| \tan Ks > 1 - \|p_0\|\tilde{K}s \ge 1 - c_0\tilde{K}$$

for all $s \le T(t)$. Consequently, for such s,

$$\|P(s, x_0, p_0)\| \le \frac{\|p_0\| + \tan Ks}{1 - c_0\tilde{K}}, \quad \|X(s, x_0, p_0) - x_0\| \le Ks + K\frac{c + s\tan Ks}{1 - c_0\tilde{K}}.$$

We have proved the required estimate but only for $T(t)$ instead of t. However, if one chooses t_0, c_0 in such a way that the last inequality implies $X(s, x_0, p_0) \in U(x_0)$, it will follow that $T(t) = t$. Indeed, if $T(t) < t$, then either $X(T(t), x_0, p_0)$

belongs to the boundary of $U(x_0)$ or $P(T(t),x_0,p_0)=\infty$, which contradicts the last inequalities. The lemma is proved.

Lemma 2.2 *There exist* $t_0>0$ *and* $c_0>0$ *such that if* $t\in(0,t_0]$, $c\in(0,c_0]$, $p_0\in B_{c/t}$, *then*

$$\frac{1}{s}\frac{\partial X}{\partial p_0}(s,x_0,p_0)=G(x_0)+O(c+t),\quad \frac{\partial P}{\partial p_0}(s,x_0,p_0)=1+O(c+t) \tag{2.4}$$

uniformly for all $s\in(0,t]$.

Proof. Differentiating the first equation in (2.2) yields

$$\ddot{x}_i=\frac{\partial G_{ik}}{\partial x_l}(x)\dot{x}_l p_k+G_{ik}(x)\dot{p}_k-\frac{\partial A_i}{\partial x_l}\dot{x}_l$$

$$=\left(\frac{\partial G_{ik}}{\partial x_l}p_k-\frac{\partial A_i}{\partial x_l}\right)(G_{lj}p_j-A_l)+G_{ik}\left(-\frac{1}{2}(\frac{\partial G}{\partial x_k}p,p)+(\frac{\partial A}{\partial x_k},p)+\frac{\partial V}{\partial x_k}\right). \tag{2.5}$$

Consequently, differentiating the Taylor expansion

$$x(s)=x_0+\dot{x}(0)s+\int_0^s(s-\tau)\ddot{x}(\tau)\,d\tau \tag{2.6}$$

with respect to the initial momentum p_0 and using (2.3) one gets

$$\frac{\partial X}{\partial p_0}(s,x_0,p_0)=G(x_0)s$$

$$+\int_0^s\left(O(1+\|p_0\|^2)\frac{\partial X}{\partial p_0}(\tau,x_0,p_0)+O(1+\|p_0\|)\frac{\partial P}{\partial p_0}(\tau,x_0,p_0)\right)(s-\tau)\,d\tau. \tag{2.7}$$

Similarly differentiating $p(s)=p_0+\int_0^s\dot{p}(\tau)\,d\tau$ one gets

$$\frac{\partial P}{\partial p_0}(s,x_0,p_0)$$

$$=1+\int_0^s\left(O(1+\|p_0\|^2)\frac{\partial X}{\partial p_0}(\tau,x_0,p_0)+O(1+\|p_0\|)\frac{\partial P}{\partial p_0}(\tau,x_0,p_0)\right)d\tau. \tag{2.8}$$

Let us now regard the matrices $v(s)=\frac{1}{s}\frac{\partial X}{\partial p_0}(s,x_0,p_0)$ and $u(s)=\frac{\partial P}{\partial p_0}(s,x_0,p_0)$ as elements of the Banach space $M_m[0,t]$ of continuous $m\times m$-matrix-valued functions $M(s)$ on $[0,t]$ with norm $\sup\{\|M(s)\|:s\in[0,t]\}$. Then one can write equations (2.7), (2.8) in abstract form

$$v=G(x_0)+L_1v+\tilde{L}_1u,\quad u=1+L_2v+\tilde{L}_2u,$$

where $L_1,L_2,\tilde{L}_1,\tilde{L}_2$ are linear operators in $M_m[0,t]$ with norms $\|L_i\|=O(c^2+t^2)$ and $\|\tilde{L}_i\|=O(c+t)$. This implies (2.4) for c and t small enough. In fact, from the second equation we get $u = 1 + O(c+t) + O(c^2+t^2)v$, substituting this equality in the first equation yields $v = G(x_0) + O(c+t) + O(c^2+t^2)v$, and solving this equation with respect to v we obtain the first equation in (1.4).

Now we are ready to prove the main result of this section, namely the existence of the family $\Gamma(x_0)$ of the characteristics of system (2.2) starting from x_0 and covering a neighbourhood of x_0 in times $t \le t_0$.

Theorem 2.1. *(i) For each $x_0 \in \mathcal{R}^m$ there exist c and t_0 such that for all $t \le t_0$ the mapping $p_0 \mapsto X(t, x_0, p_0)$ defined on the ball $B_{c/t}$ is a diffeomorphism onto its image.*

(ii) For an arbitrary small enough c there exist positive $r = O(c)$ and $t_0 = O(c)$ such that the image of this diffeomorphism contains the ball $B_r(x_0)$ for all $t \le t_0$.

Proof. (i) Note first that, by Lemma 2.2, the mapping $p_0 \mapsto X(t, x_0, p_0)$ is a local diffeomorphism for all $t \le t_0$. Furthermore, if $p_0, q_0 \in B_{c/t}$, then

$$X(t, x_0, p_0) - X(t, x_0, q_0) = \int_0^1 \frac{\partial X}{\partial p_0}(t, x_0, q_0 + s(p_0 - q_0))\, ds\, (p_0 - q_0)$$

$$= t(G(x_0) + O(c+t))(p_0 - q_0) \tag{2.9}$$

Therefore, for c and t sufficiently small, the r.h.s. of (2.9) cannot vanish if $p_0 - q_0 \ne 0$.

(ii), (iii) We must prove that for $x \in B_r(x_0)$ there exists $p_0 \in B_{c/t}$ such that $x = X(t, x_0, p_0)$, or equivalently, that

$$p_0 = p_0 + \frac{1}{t} G(x_0)^{-1}(x - X(t, x_0, p_0)).$$

In other words the mapping

$$F_x : p_0 \mapsto p_0 + \frac{1}{t} G(x_0)^{-1}(x - X(t, x_0, p_0)) \tag{2.10}$$

has a fixed point in the ball $B_{c/t}$. Since every continuous mapping from a ball to itself has a fixed point, it is enough to prove that F_x takes the ball $B_{c/t}$ in itself, i.e. that

$$\|F_x(p_0)\| \le c/t \tag{2.11}$$

whenever $x \in B_r(x_0)$ and $\|p_0\| \le c/t$. By (2.3), (2.5) and (2.6)

$$X(t, x_0, p_0) = x_0 + t(G(x_0)p_0 - A(x_0)) + O(c^2 + t^2),$$

and therefore it follows from (2.10) that (2.11) is equivalent to

$$\|G(x_0)^{-1}(x - x_0) + O(t + c^2 + t^2)\| \le c,$$

which certainly holds for $t \le t_0$, $|x - x_0| \le r$ and sufficiently small r, t_0 whenever c is chosen small enough.

Corollary. *If (i) either A, V, G and their derivatives are uniformly bounded, or (ii) if G is a constant and A and V'' are uniformly bounded together with their derivatives, then there exist positive r, c, t_0 such that for any $t \in (0, t_0]$ and any x_1, x_2 such that $|x_1 - x_2| \leq r$ there exists a solution of system (2.2) with the boundary conditions*

$$x(0) = x_1, \quad x(t) = x_2.$$

Moreover, this solution is unique under the additional assumption that $\|p(0)\| \leq c/t$.

Proof. The case (i) follows directly from Theorem (2.1). Under assumptions (ii), to get the analog of Lemma 2.1 one should take in its proof the system

$$\begin{cases} \dot{x} = K(p+1) \\ \dot{p} = K(1+p+x) \end{cases}$$

as a bound for the solution of the Hamiltonian system. This system is linear (here the asumption that G is constant plays the role) and its solutions can be easily estimated. the rest of the proof remains the same.

The proof of the existence of the boundary value problem given above is not constructive. However, when the well-posedness is given, it is easy to construct approximate solutions up to any order in small t for smooth enough Hamiltonians. Again one begins with the construction of the asymptotic solution for the Cauchy problem.

Proposition 2.1. *If the functions G, A, V in (2.1) have $k+1$ continuous bounded derivatives, then for the solution of the Cauchy problem for equation (2.2) with initial data $x(0) = x_0$, $p(0) = p_0$ one has the asymptotic formulas*

$$X(t, x_0, p_0) = x_0 + tG(x_0)p_0 - A(x_0)t + \sum_{j=2}^{k} Q_j(t, tp_0) + O(c+t)^{k+1}, \quad (2.12)$$

$$P(t, x_0, p_0) = p_0 + \frac{1}{t}\left[\sum_{j=2}^{k} P_j(t, tp_0) + O(c+t)^{k+1}\right], \quad (2.13)$$

where $Q_j(t,q) = Q_j(t, q^1, ..., q^m)$, $P_j(t,q) = P_j(t, q^1, ..., q^m)$ are homogeneous polynomials of degree j with respect to all their arguments with coefficients depending on the values of G, A, V and their derivatives up to order j at the point x_0. Moreover, one has the following expansion for the derivatives with respect to initial momentum

$$\frac{1}{t}\frac{\partial X}{\partial p_0} = G(x_0) + \sum_{j=1}^{k} \tilde{Q}_j(t, tp_0) + O(c+t)^{k+1}, \quad (2.14)$$

$$\frac{\partial P}{\partial p_0} = 1 + \sum_{j=1}^{k} \tilde{P}_j(t, tp_0) + O(c+t)^{k+1}, \quad (2.15)$$

where $\tilde{Q}_j, \tilde{P}_j$ are again homogeneous polynomials of degree j, but now they are matrix-valued.

Proof. This follows directly by differentiating equations (2.2), then using the Taylor expansion for its solution up to k-th order and estimating the remainder using Lemma 2.1.

Proposition 2.2. *Under the hypotheses of Proposition 2.1, the function $p_0(t,x,x_0)$ (defined by (1.2)), which, by Theorem 2.1, is well-defined and smooth in $B_R(x_0)$, can be expended in the form*

$$p_0(t,x,x_0) = \frac{1}{t}G(x_0)^{-1}\left[(x-x_0) + A(x_0)t + \sum_{j=2}^{k} P_j(t, x-x_0) + O(c+t)^{k+1}\right], \tag{2.16}$$

where $P_j(t,x-x_0)$ are certain homogeneous polynomials of degree j in all their arguments.

Proof. It follows from (2.12) that $x - x_0$ can be expressed as an asymptotic power series in the variable $(p_0 t)$ with coefficients that have asymptotic expansions in powers of t. This implies the existence and uniqueness of the formal power series of form (2.16) solving equation (2.12) with respect to p_0. The well-posedness of this equation (which follows from Theorem 2.1) completes the proof.

Proposition 2.3. *Under the assumptions of Proposition 2.1, the two-point function $S(t,x,x_0)$ defined in (1.5), can be expended in the form*

$$\begin{aligned} S(t,x,x_0) &= \frac{1}{2t}(x-x_0+A(x_0)t, G(x_0)^{-1}(x-x_0+A(x_0)t)) \\ &+\frac{1}{t}(V(x_0)t^2 + \sum_{j=3}^{k} P_j(t,x-x_0) + O(c+t)^{k+1}), \end{aligned} \tag{2.17}$$

where the P_j are again polynomials in t and $x - x_0$ of degree j (and the term quadratic in $x - x_0$ is written explicitly).

Proof. One first finds the asymptotic expansion for the action $\sigma(t,x_0,p_0)$ defined in (1.4). For Hamiltonian (2.1) one gets that $\sigma(t,x_0,p_0)$ equals

$$\int_0^t \left[\frac{1}{2}\left(G(X(\tau,x_0,p_0))P(\tau,x_0,p_0), P(\tau,x_0,p_0)\right) + V(X(\tau,x_0,p_0))\right] d\tau,$$

and using (2.12), (2.13) one obtains

$$\sigma(t,x_0,p_0) = \frac{1}{t}\left[\frac{1}{2}(p_0 t, G(x_0)p_0 t) + V(x_0)t^2 + \sum_{j=3}^{k} P_j(t, tp_0) + O(c+t)^{k+1}\right],$$

where P_j are polynomials of degree $\leq j$ in p_0. Inserting the asymptotic expansion (2.16) for $p_0(t,x,x_0)$ in this formula yields (2.17).

Remark. One can calculate the coefficients of the expansion (2.17) directly from the Hamilton-Jacobi equation without solving the boundary value problem for (2.2) (as we shall do in the next chapter). The theory presented above explains why the asymptotic expansion has such a form and justifies the formal calculation of its coefficients by means of, for example, the method of undetermined coefficients.

By Theorem 2.1, all the assumptions of Proposition 1.1 hold for the Hamiltonian (2.1); moreover, for all x_0, the domain $D(x_0)$ can be chosen as the ball $B_r(x_0)$. It was proved in Corollary 2 of Proposition 1.1 that the two-point function $S(t,x,x_0)$, defined by (1.5),(1.6) in a neighbourhood of x_0, is equal to the minimum of the functional (1.12) over all curves lying in the domain $B_r(x_0)$. We shall show first that (at least for r sufficiently small) it is in fact the global minimum (i.e. among all curves, not only those lying in $B_r(x_0)$).

Proposition 2.4. *Let the potential V be uniformly bounded from below. Then there exists $r_1 \le r$ such that for $x \in B_{r_1}(x_0)$ the function $S(t,x,x_0)$ defined by (1.5) gives the global minimum of the functional (1.12).*

Proof. It follows from asymptotic representation (2.17) (in fact, from only its first two terms) that there exist $r_1 \le r$ and $t_1 \le t_0$ such that for a $\delta > -t_0 \inf V$

$$\max_{\|x-x_0\|=r_1} S(t_1,x,x_0) \le \min_{t\le t_1} \min_{\|y-x_0\|=r} S(t,y,x_0) - \delta. \tag{2.18}$$

Because $L(x,\dot{x}) - V(x)$ is positive, the result of Proposition 2.4 follows from (2.18) using the fact that, on the one hand, the functional (1.12) depends additively on the curve, and on the other hand, if a continuous curve $y(\tau)$ joining x_0 and $x \in B_{r_1}(x_0)$ in time t_1 is not completely contained in $B_r(x_0)$, there exists $t_2 < t_1$ such that $|y(t_2) - x_0| = r$ and $y(\tau) \in B_r(x_0)$ for $t \le t_2$.

The following result is a consequence of the Hilbert regularity theorem (Proposition 1.6). However we shall give another proof, which is independent of the Lagrangian formalism and which will be convenient for some other situations.

Proposition 2.5. *Suppose that the absolute minimum of the functional (1.12) for the Hamiltonian (2.1) is attained by a piecewise smooth curve. Then this curve is in fact a smooth characteristic, i.e. a solution of (2.2).*

Proof. Suppose that $y(\tau)$ gives a minimum for I_t and at $\tau = s$ its derivative is discontinuous (or it is not a solution of (2.2) in a neighbourhood of $\tau = s$). Then for δ sufficiently small, $|y(s+\delta) - y(s-\delta)| < r_1$, where r_1 is defined in Proposition 2.4, and replacing the segment $[y(s-\delta), y(s+\delta)]$ of the curve $y(\tau)$ by the characteristic joining $y(s-\delta)$ and $y(s+\delta)$ in time 2δ one gets a trajectory, whose action is be less then that of $y(\tau)$. This contradiction completes the proof.

Proposition 2.6. (Tonelli's theorem). *Under the assumptions of Proposition 2.4, for arbitrary $t > 0$ and x_0, x, there exists a solution $(x(\tau), p(\tau))$ of (2.2) with the boundary conditions $x(0) = x_0, x(t) = x$ such that the characteristic $x(\tau)$ attains the global minimum for the corresponding functional (1.12).*

Proof. Suppose first that the functions G, V, A are uniformly bounded together with all their derivatives. Let $t \le t_0$ and suppose that $kr_1 < |x -$

$x_0| \leq (k+1)r_1$ for some natural number k. Then there exists a piecewise-smooth curve $y(s)$ joining x_0 and x in time t such that it has not more than k points $y(s_1), ..., y(s_k)$, where the derivatives $\dot{y}(s)$ is discontinuous, and for all j, $|y(s_j) - y(s_{j-1})| \leq r_1$ and which is an extremal on the interval $[s_{j-1}, s_j]$. The existence of such a curve implies in particular that $J = \inf I_t$ is finite. Now let $y_n(s)$ be a minimising sequence for (1.12), i.e. a sequence of piecewise- smooth curves joining x_0 and x in time t such that $\lim_{n\to\infty} I_t(y_n(.)) = J$. Comparing the action along $y_n(s)$ with the action along $y(s)$ using (2.18) one concludes that, for $n > n_0$ with some n_0, all curves $y_n(s)$ lie entirely in $B_{kr}(x_)) \cup B_r(x_0)$. Consequently, one can define a finite sequence of points $y_n(t_j(n)), j \leq k$, on the curve y_n recursively by

$$t_j(n) = \sup\{t > t_{j-1} : |y_n(t) - y_n(t_{j-1}(n))| < r\}.$$

Since all $y_n(t_j(n))$ belong to a compact set, it follows that there exists a subsequence, which will again be denoted by y_n such that the number of the $\{t_j\}$ does not depend on n and the limits $t_j = \lim_{n\to\infty} t_j(n)$ and $y_j = \lim_{n\to\infty} y_n(t_j(n))$ exist for all j. Consider now the sequence of curves $\tilde{y}_n(s)$ constructed by the following rule: on each interval $[t_{j-1}(n), t_j(n)]$ the curve $\tilde{y}_n$ is the extremal joining $y_n(t_{j-1}(n))$ and $y_n(t_j(n))$ in time $t_j(n) - t_{j-1}(n)$. By corollary 2 of Proposition 1.1, the limit of the actions along $\tilde{y}_n$ is also J. But, clearly, the sequence $\tilde{y}_n$ tends to a broken extremal $\tilde{y}$ (whose derivatives may be discontinuous only in points t_j) with the action J, i.e. $\tilde{y}(s)$ gives a minimum for (1.12). By Proposition 2.5, this broken extremal is in fact everywhere smooth. Finally, one proves the result of Proposition 2.6 for all $t > 0$ by a similar procedure using the splitting of the curves of a minimising sequence into parts with the time length less than t_0 and replacing these parts of curves by extremals. The case when G, A, V are not uniformly bounded, is proved by a localisation argument. Namely, any two points can be placed in a large ball, where everything is bounded.

We shall describe now the set of regular points for a variational problem with fixed ends. As we shall see in the next chapter, these are the points for which the WKB-type asymptotics of the corresponding diffusion takes the simplest form.

Let us fix a point x_0. We say that the pair (t, x) is a regular point (with respect to x_0), if there exists a unique characteristic of (2.2) joining x_0 and x in time t and furnishing a minimum to the corresponding functional (1.12), which is not degenerate in the sense that the end point $x(t)$ is not conjugate to x_0 along this characteristic, which implies in particular that $\frac{\partial X}{\partial p_0}(t, x_0, p_0)$ is not degenerate.

Proposition 2.7. *For arbitrary x_0, the set $Reg(x_0)$ of regular points is an open connected and everywhere dense set in $\mathcal{R}_+ \times \mathcal{R}^m$. Moreover, for arbitrary (t,x), all pairs $(\tau, x(\tau))$ with $\tau < t$ lying on any minimising characteristic joining x_0 and x in time t are regular. For any fixed t, the set $\{x : (t,x) \in Reg(x_0)\}$ is open and everywhere dense in $\mathcal{R}^m$.*

Proof. The second statement is a direct consequence of Proposition 2.5. In its turn, this statement, together with Proposition 2.6, implies immediately that the set $Reg(x_0)$ is everywhere dense and connected. In order to prove that this set is open, suppose that (t, x) is regular, and therefore $\frac{\partial X}{\partial p_0}(t, x_0, p_0)$ is non-degenerate. By the inverse function theorem this implies the existence of a continuous family of characteristics emanating of x_0 and coming to any point in a neighbourhood of (t, x). Then by the argument of the proof of Corollary 2 to Proposition 1.1 one proves that each such characteristic furnishes a local minimum to (1.12). Since at (t, x) this local minimum is in fact global, one easily gets that the same holds for neighbouring points. The last statement of the Proposition is a consequence of the others.

At the beginning of section 1.1 we defined the two-point function $S(t, x, x_0)$ locally as the action along the unique characteristic joining x_0 and x in time t. Then we proved that it gives a local minimum, and then that it gives even a global minimum for the functional (1.12), when the distance between x_0 and x is small enough. As a consequence of the last propositions, one can claim this connection to be global.

Proposition 2.8. *Let us define the two-point function $S(t, x, x_0)$ for all $t > 0, x, x_0$ as the global minimum of the functional (1.12). Then, in the case of the Hamiltonian (2.1), $S(t, x, x_0)$ is an everywhere finite and continuous function, which for all all (t, x, x_0) is equal to the action along a minimising characteristic joining x_0 and x in time t. Moreover, on the set $Reg(x_0)$ of regular points, $S(t, x, x_0)$ is smooth and satisfies the Hamilton-Jacobi equation (1.7).*

Remark 1. As stated in Proposition 2.8, the two-point function $S(t, x, x_0)$ is almost everywhere smooth and almost everywhere satisfies the Hamilton-Jacobi equation. In the theory of generalised solutions of Hamilton-Jacobi-Bellman equation (see e.g. [KM1], [KM2]) one proves that $S(t, x, x_0)$ is in fact the generalised solution of the Cauchy problem for equation (1.7) with discontinuous initial data: $S(0, x_0, x_0) = 0$ and $S(0, x, x_0) = +\infty$ for $x \neq x_0$.

Remark 2. An important particular case of the situation considered in this section is the case of a purely quadratic Hamiltonian, namely when $H = (G(x)p, p)$. The solutions of the corresponding system (1.1) (or more precisely, their projections on x-space) are called geodesics defined by the Riemanian metric given by the matrix $g(x) = G^{-1}(x)$. For this case, theorem 2.1 reduces to the well known existence and uniqueness of minimising geodesics joining points with sufficiently small distance between them. The proofs for this special case are essentially simpler, because geodesics enjoy the following homogeneity property. If $(x(\tau), p(\tau))$ is a solution of the corresponding Hamiltonian system, then the pair $(x(\epsilon t), \epsilon p(\epsilon t))$ for any $\epsilon > 0$ is a solution as well. Therefore, having the local diffeomorphism for some $t_0 > 0$ one automatically gets the results for all $t \leq t_0$.

Remark 3. There seems to be no reasonable general criterion for uniqueness of the solution of the boundary value problem, as is shown even by the case of geodesic flows (where uniqueness holds only under the assumption of negative curvature). Bernstein's theorem (see, e.g. [Ak]) is one of the examples of (very restrictive) conditions that assure global uniqueness. Another example is provided by the case of constant diffusion and constant drift, which we shall now discuss.

Proposition 2.9 [M6, DKM1]. *Suppose that in the Hamiltonian (2.1) $G(x) = G$ and $A(x) = A$ are constant and the matrix of second derivatives of V is uniformly bounded. Then the estimates (2.4) for the derivatives are global. More precisely,*

$$\frac{1}{t}\frac{\partial X}{\partial p_0}(t, x_0, p_0) = G(x_0) + O(t^2), \quad \frac{\partial P}{\partial p_0}(s, x_0, p_0) = 1 + O(t^2), \quad (2.19)$$

$$\frac{\partial^2 X}{\partial p_0^2} = O(t^4), \quad \frac{\partial^3 X}{\partial p_0^3} = O(t^4) \quad (2.20)$$

and therefore

$$\frac{\partial^2 S}{\partial x^2} = \frac{1}{tG}(1 + O(t^2)), \quad \frac{\partial^2 S}{\partial x_0^2} = \frac{1}{tG}(1 + O(t^2)), \quad \frac{\partial^2 S}{\partial x \partial x_0} = -\frac{1}{tG}(1 + O(t^2)) \quad (2.21)$$

uniformly for all $t \leq t_0$ and all p_0. Moreover, for some $t_0 > 0$ the mapping $p_0 \mapsto X(t, x_0, p_0)$ is a global diffeomorphism $\mathcal{R}^n \mapsto \mathcal{R}^n$ for all $t \leq t_0$, and thus the boundary value problem for the corresponding Hamiltonian system has a unique solution for small times and arbitrary end points x_0, x such that this solution provides an absolute minimum in the corresponding problem of the calculus of variations.

Sketch of the proof. It follows from the assumptions that the Jacobi equation (1.16) has uniformly bounded coefficients. This implies the required estimates for the derivatives of the solutions to the Hamiltonian system with respect to p_0. This, in turn, implies the uniqueness of the boundary value problem for small times and arbitrary end points x_0, x. The corresponding arguments are given in detail in a more general (stochastic) situation in Section 7.

Similar global uniqueness holds for the model of the next section, namely for the Hamiltonian (3.4), if g is a constant matrix, $\alpha(x, y) = y$ and $b(x, y) = b(x), V(x, y) = V(x)$ do not depend on y and are uniformly bounded together with their derivatives of first and second order; see details in [KM2].

3. Regular degenerate Hamiltonians of the first rank

We are turning now to the main topic of this chapter, to the investigation of the boundary value problem for degenerate Hamiltonians. As in the first chapter, we shall suppose that the coordinates (previously denoted by x) are divided into two parts, $x \in \mathcal{R}^n$ and $y \in \mathcal{R}^k$ with corresponding momenta $p \in \mathcal{R}^n$ and $q \in \mathcal{R}^k$ respectively, and that H is non-degenerate with respect to q. More precisely, H has the form

$$H(x, y, p, q) = \frac{1}{2}(g(x, y)q, q) - (a(x, y), p) - (b(x, y), q) - V(x, y), \quad (3.1)$$

where $g(x, y)$ is a non-singular positive-definite $(k \times k)$-matrix such that

$$\Lambda^{-1} \leq g(x, y) \leq \Lambda \quad (3.2)$$

for all x, y and some positive Λ. It is natural to try to classify the Hamiltonians of form (3.1) in a neighbourhood of a point (x_0, y_0) by their quadratic approximations

$$\tilde{H}_{x_0,y_0}(x,y,p,q) = -\left(a(x_0,y_0) + \frac{\partial a}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial a}{\partial y}(x_0,y_0)(y-y_0), p\right)$$

$$+\frac{1}{2}(g(x_0,y_0)q,q) - \tilde{V}_{x_0,y_0}(x,y)$$

$$-\left[b(x_0,y_0) + \frac{\partial b}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial b}{\partial y}(x_0,y_0)(y-y_0), q\right], \tag{3.3}$$

where $\tilde{V}_{x_0,y_0}$ is the quadratic term in the Taylor expansion of $V(x,y)$ near x_0, y_0. For the Hamiltonian (3.3), the Hamiltonian system (1.1) is linear and its solutions can be investigated by means of linear algebra. However, it turns out that the qualitative properties of the solutions of the boundary value problem for H are similar to those of its approximation (3.3) only for rather restrictive class of Hamiltonians, which will be called regular Hamiltonians. In the next section we give a complete description of this class and further we shall present an example showing that for non-regular Hamiltonians the solution of the boundary value problem may not exist even locally, even though it does exist for its quadratic approximation. In this section we investigate in detail the simplest and the most important examples of regular Hamiltonians, which correspond in the quadratic approximation to the case, when $k \geq n$ and the matrix $\frac{\partial a}{\partial y}(x_0, y_0)$ has maximal rank. For this type of Hamiltonian we shall make a special notational convention, namely we shall label the coordinates of the variables x and q by upper indices, and those of the variables y and p by low indices. The sense of this convention will be clear in the next chapter when considering the invariant diffusions corresponding to these Hamiltonians.

Definition. *The Hamiltonian of form (3.1) is called regular of the first rank of degeneracy, if $k \geq n$, g does not depend on y, the functions a, b, V are polynomials in y of degrees not exceeding 1,2 and 4 respectively with uniformly bounded coefficients depending on x, the polynomial V is bounded from below, and the rank of $\frac{\partial a}{\partial y}(x,y)$ everywhere equals to its maximal possible value n.*

Such a Hamiltonian can be written in the form

$$H = \frac{1}{2}(g(x)q,q) - (a(x) + \alpha(x)y, p) - (b(x) + \beta(x)y + \frac{1}{2}(\gamma(x)y,y), q) - V(x,y), \tag{3.4}$$

or more precisely as

$$H = \frac{1}{2}g_{ij}(x)q^iq^j - (a^i(x) + \alpha^{ij}y_j)p_i - (b_i(x) + \beta_i^j(x)y_j + \frac{1}{2}\gamma_i^{jl}y_jy_l)q^i - V(x,y), \tag{3.4'}$$

where $V(x,y)$ is a polynomial in y of degree ≤ 4, bounded from below, and *rank* $\alpha(x) = n$. The Hamiltonian system (1.1) for this Hamiltonian has the form

$$\begin{cases} \dot{x}^i = -(a^i(x) + \alpha^{ij}(x)y_j) \\ \dot{y}_i = -(b_i(x) + \beta_i^j(x)y_j + \frac{1}{2}\gamma_i^{jl}(x)y_jy_m) + g_{ij}(x)q^j \\ \dot{q}^i = \alpha^{ji}(x)p_j + (\beta_j^i(x) + \gamma_j^{il}(x)y_l)q^j + \frac{\partial V}{\partial y_i}(x,y) \\ \dot{p}_i = (\frac{\partial a^j}{\partial x^i} + \frac{\partial \alpha^{jl}}{\partial x^i}y_l)p_j + \left(\frac{\partial b_j}{\partial x^i} + \frac{\partial \beta_j^l}{\partial x^i}y_l + \frac{1}{2}\frac{\partial \gamma_j^{lm}}{\partial x^i}y_ly_m\right)q^j - \frac{1}{2}\frac{\partial g_{jl}}{\partial x^i}q^jq^l + \frac{\partial V}{\partial x^i}, \end{cases} \tag{3.5}$$

where for brevity we have omitted the dependence on x of the coefficients in the last equation.

Proposition 3.1. *There exist constants K, t_0, and c_0 such that for all $c \in (0, c_0]$ and $t \in (0, t_0]$, the solution of the system (3.5) with initial data $(x(0), y(0), p(0), q(0))$ exists on the interval $[0,t]$ whenever*

$$|y(0)| \leq \frac{c}{t}, \quad |q(0)| \leq \frac{c^2}{t^2}, \quad |p(0)| \leq \frac{c^3}{t^3}, \tag{3.6}$$

and on this interval

$$|x - x(0)| \leq Kt(1 + \frac{c}{t}), \quad |y - y(0)| \leq Kt(1 + \frac{c^2}{t^2}), \tag{3.7}$$

$$|q - q(0)| \leq Kt(1 + \frac{c^3}{t^3}), \quad |p - p(0)| \leq Kt(1 + \frac{c^4}{t^4}). \tag{3.8}$$

Proof. Estimating the derivatives of the magnitudes $|y|, |q|, |p|$ from (3.5), one sees that their growths do not exceed the growths of the solutions of the system

$$\begin{cases} \dot{y} = \sigma(1 + q + y^2) \\ \dot{q} = \sigma(1 + p + yq + y^3) \\ \dot{p} = \sigma(1 + yp + q^2 + y^4) \end{cases} \tag{3.9}$$

for some constant σ, and with initial values $y_0 = |y(0)|, q_0 = |q(0)|, p_0 = |p(0)|$. Suppose (3.6) holds. We claim that the solution of (3.9) exists on the interval $[0,t]$ as a convergent series in $\tau \in [0,t]$. For example, let us estimate the terms of the series

$$p(t) = p_0 + t\dot{p}_0 + \frac{1}{2}t^2\ddot{p}_0 + ...,$$

where the $p_0^{(j)}$ are calculated from (3.9). The main observation is the following. If one allocates the degrees 1,2 and 3 respectively to the variables y, q, and p, then the right hand sides of (3.9) have degrees 2, 3 and 4 respectively. Moreover, $p_0^{(j)}$ is a polynomial of degree $j+3$. Therefore, one can estimate

$$p_0^{(j)} \leq \nu_j \sigma^j \left(1 + \left(\frac{c}{t}\right)^{j+3}\right)$$

where ν_j are natural numbers depending only on combinatorics. A rough estimate for ν_j is

$$\nu_j \le 4 \cdot (4 \cdot 4) \cdot (4 \cdot 5) ... (4 \cdot (k+2)) = \frac{(j+2)!}{6} 4^j,$$

because each monomial of degree $\le d$ (with coefficient 1) can produce after differentiation in t not more than $4d$ new monomials of degree $\le (d+1)$ (again with coefficient 1). Consequently, the series for $p(t)$ is dominated by

$$p_0 + \sum_{j=1}^{\infty} \frac{t^j}{j!} \frac{(j+2)!}{6} (4\sigma)^j \left(1 + \left(\frac{c}{t}\right)^{j+3}\right) \le K \left(1 + \left(\frac{c}{t}\right)^3\right),$$

for some K, if $4\sigma c_0 < 1$ and $4\sigma t_0 < 1$. Using similar estimates for q and y we prove the existence of the solution of (3.5) on the interval $[0,t]$ with the estimates

$$|y| \le K \left(1 + \frac{c}{t}\right), \quad |q| \le K \left(1 + \left(\frac{c}{t}\right)^2\right), \quad |p| \le K \left(1 + \left(\frac{c}{t}\right)^3\right), \tag{3.10}$$

which directly imply (3.7).(3.8).

Using Proposition 3.1 we shall obtain now more precise formulae for the solution of (3.5) for small times. We shall need the development of the solutions of the Cauchy problem for (3.5) in Taylor series up to orders 1,2,3 and 4 respectively for p, q, y, and x. The initial values of the variables will be denoted by x_0, y^0, p^0, q_0. In order to simplify the forms of rather long expressions that appear after differentiating the equations of (3.5), it is convenient to use the following pithy notation: μ will denote an arbitrary (uniformly) bounded function in x and expressions such as μy^j will denote polynomials in $y = (y_1, ... y_k)$ of degree j with coefficients of the type $O(\mu)$. Therefore, writing $p(t) = p(0) + \int_0^t \dot p(\tau)\, d\tau$ and using the last equation from (3.5) one gets

$$p = p(0) + \int_0^t [(\mu + \mu y)p + (\mu + \mu y + \mu y^2)q + \mu q^2 + \mu + \mu y + \mu y^2 + \mu y^3 + \mu y^4]\, d\tau. \tag{3.11}$$

Differentiating the third equation in (3.5) yields

$$\ddot q^i = \frac{\partial \alpha^{ji}}{\partial x^m} \dot x^m p_j + \alpha^{ji} \dot p_j + \left(\frac{\partial \beta^i_j}{\partial x^m} \dot x^m + \frac{1}{2} \frac{\partial \gamma^{im}_j}{\partial x^l} \dot x^l y_m + \gamma^{im}_j \dot y_m \right) q^j$$

$$+ (\beta^i_j + \gamma^{im}_j y_m) \dot q^j + \frac{\partial^2 V}{\partial y_i \partial x^m} \dot x^m + \frac{\partial^2 V}{\partial y_i \partial y_m} \dot y_m, \tag{3.12}$$

and from the representation $q(t) = q(0) + t\dot q(0) + \int_0^t (t-\tau) \ddot q(\tau)\, d\tau$ one gets

$$q^i = q^i_0 + \left[\alpha^{ji}(x_0) p^0_j + (\beta^i_j(x_0) + \gamma^{im}_j(x_0) y^0_m) q^j_0 + \frac{\partial V}{\partial y_i}(x_0, y^0) \right] t$$

$$+\int_0^t (t-\tau)[(\mu+\mu y)p+(\mu+\mu y+\mu y^2)q+\mu q^2+\mu+\mu y+\mu y^2+\mu y^3+\mu y^4]\,d\tau. \quad (3.13)$$

Furthermore,

$$\ddot{y}_i = \left(\frac{\partial b_i}{\partial x^m} + \frac{\partial \beta_i^j}{\partial x^m} y_j + \frac{1}{2}\frac{\partial \gamma_i^{jl}}{\partial x^m} y_j y_l\right)(a^m + \alpha^{mk} y_k)$$

$$-(\beta_i^j + \gamma_i^{jl} y_l)\left(g_{jm} q^m - (b_j + \beta_j^m y_m + \frac{1}{2}\gamma_j^{mk} y_m y_k)\right)$$

$$-\frac{\partial g_{ij}}{\partial x^m}(a^m + \alpha^{ml} y_l) q^j + \left(\alpha^{lj} p_l + (\beta_l^j + \gamma_l^{jm} y_m) q^l + \frac{\partial V}{\partial y_j}\right), \quad (3.14)$$

or, in concise notation,

$$\ddot{y} = \mu + \mu y + \mu y^2 + \mu y^3 + (\mu + \mu y) q + \mu p.$$

It follows that

$$y^{(3)} = (\mu + \mu y + \mu y^2 + \mu y^3 + \mu y^4) + (\mu + \mu y + \mu y^2) q + \mu q^2 + (\mu + \mu y) p. \quad (3.15)$$

Let $(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) = (\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q})(t, x_0, y^0, 0, 0)$ denote the solution of (3.5) with initial condition $(x_0, y^0, 0, 0)$. From (3.14), (3.15) one obtains

$$y = \tilde{y} + t g(x_0) q_0 + \frac{1}{2} t^2 \left[g(x_0)\alpha(x_0) p^0 + (\omega + (\Omega y^0)) q_0\right]$$

$$+\int_0^t (t-\tau)^2 [(\mu+\mu y+\mu y^2+\mu y^3+\mu y^4)+(\mu+\mu y+\mu y^2)q+\mu q^2+(\mu+\mu y)p]\,d\tau, \quad (3.16)$$

where the matrices ω and Ωy^0 are given by

$$\omega_{il} = g_{ij}(x_0)\beta_l^j(x_0) - \beta_i^j(x_0) g_{jl}(x_0) - \frac{\partial g_{il}}{\partial x^j}(x_0) a_j(x_0), \quad (3.17)$$

$$(\Omega y^0)_{il} = \left(g_{ij}(x_0)\gamma_l^{jm}(x_0) g_{jl}(x_0) - \frac{\partial g_{il}}{\partial x^j}(x_0)\alpha_{jm}(x_0)\right) y_m^0. \quad (3.18)$$

Differentiating the first equation in (3.5) yields

$$\ddot{x}^i = \left(\frac{\partial a^i}{\partial x^m} + \frac{\partial \alpha^{ij}}{\partial x^m} y_j\right)(a^m + \alpha^{ml} y_l) - \alpha_{ij} \dot{y}_j,$$

and therefore

$$(x^{(3)})^i = \left[\left(2\frac{\partial \alpha^{ij}}{\partial x^m}\alpha^{ml} + \frac{\partial \alpha^{il}}{\partial x^m}\alpha^{mj}\right) y_l + 2\frac{\partial \alpha^{ij}}{\partial x^m} a^m + \frac{\partial a^i}{\partial x^m}\alpha^{mj}\right] \dot{y}_j$$

$$-\alpha^{ij}\ddot{y}_j + \mu + \mu y + \mu y^2 + \mu y^3.$$

In particular, the consice formula for $x^{(3)}$ is the same as for $\ddot{y}$ and for $\dot{q}$. Let us write now the formula for x which one gets by keeping three terms of the Taylor series:

$$x = \tilde{x} - \frac{1}{2}t^2\alpha(x_0)g(x_0)q_0 - \frac{1}{6}t^3[(\alpha g\alpha')(x_0)p^0 + (\omega' + (\Omega' y^0))q_0]$$

$$+\int_0^t (t-\tau)^3[(\mu+\mu y+\mu y^2+\mu y^3+\mu y^4)+(\mu+\mu y+\mu y^2)q+\mu q^2+(\mu+\mu y)p]\,d\tau, \tag{3.19}$$

where the entries of the matrices ω' and $\Omega' y^0$ are given by

$$(\omega')^i_k = \alpha^{ij}(x_0)\omega_{jk} - \left(2\frac{\partial\alpha^{ij}}{\partial x^m}a^m + \frac{\partial a^i}{\partial x^m}\alpha^{mj}\right)(x_0)g_{jk}(x_0), \tag{3.20}$$

$$(\Omega' y^0)^i_k = \alpha^{ij}(x_0)(\Omega y^0)_{jk} - \left(2\frac{\partial\alpha^{ij}}{\partial x^m}\alpha^{ml} + \frac{\partial\alpha^{il}}{\partial x^m}\alpha^{mj}\right)(x_0)g_{jk}(x_0)y^0_l. \tag{3.21}$$

One can now obtain the asymptotic representation for the solutions of (3.5) and their derivatives with respect to the initial momenta in the form needed to prove the main result on the well-posedness of the boundary value problem. In the following formulas, δ will denote an arbitrary function of order $O(t+c)$, and α_0 will denote the matrix $\alpha(x_0)$, with similar notation for other matrices at initial point x_0. Expanding the notations of Sect. 2.1 we shall denote by $(X, Y, P, Q)(t, x_0, y^0, p^0, q_0)$ the solution of the Cauchy problem for (3.5).

Proposition 3.2. *Under the assumptions of Proposition 3.1, for the solutions of (3.5) one has*

$$X = \tilde{x} - \frac{1}{2}t^2\alpha_0 g_0 q_0 - \frac{1}{6}t^3[\alpha_0 g_0\alpha'_0 p^0 + (\omega' + (\Omega' y^0))q_0] + \delta^4, \tag{3.22}$$

$$Y = \tilde{y} + tg_0q_0 + \frac{1}{2}t^2[g_0\alpha'_0p^0 + (\omega + (\Omega y^0))q_0] + \frac{1}{t}\delta^4, \tag{3.23}$$

where the matrices $\omega, \omega', (\Omega y^0), (\Omega' y^0)$ are defined by (3.17),(3.18),(3.20),(3.21), and the solution $(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q})$ of (3.5) with initial conditions $(x_0, y^0, 0, 0)$ is given by

$$\tilde{x} = x_0 - (a_0 + \alpha_0 y^0)t + O(t^2), \quad \tilde{y} = y^0 + O(t), \tag{3.24}$$

$$\tilde{p} = O(t), \quad \tilde{q} = O(t). \tag{3.25}$$

Moreover,

$$\frac{\partial X}{\partial p^0} = -\frac{1}{6}t^3(\alpha_0 g_0\alpha'_0 + \delta), \quad \frac{\partial X}{\partial q_0} = -\frac{1}{2}t^2(\alpha_0 g_0 + \frac{t}{3}(\omega' + \Omega' y^0) + \delta^2), \tag{3.26}$$

$$\frac{\partial Y}{\partial p^0} = \frac{1}{2}t^2(g_0\alpha'_0 + \delta), \quad \frac{\partial Y}{\partial q_0} = t(g_0 + \frac{t}{2}(\omega + \Omega y^0) + \delta^2), \tag{3.27}$$

$$\frac{\partial P}{\partial p^0} = 1 + \delta, \quad \frac{\partial P}{\partial q_0} = \frac{\delta^2}{t}, \tag{3.28}$$

$$\frac{\partial Q}{\partial p^0} = t(\alpha_0' + \delta), \quad \frac{\partial Q}{\partial q_0} = 1 + t(\beta_0' + \gamma' y^0) + \delta^2, \tag{3.29}$$

where $(\beta_0' + \gamma_0' y^0)_j^i = \beta_j^i(x_0) + \gamma_j^{im}(x_0) y_m^0$.

Proof. (3.22)-(3.25) follow directly from (3.16), (3.19) and the estimates (3.10). They imply also that the matrices

$$v_1 = \frac{1}{t^3}\frac{\partial X}{\partial p^0}, \quad u_1 = \frac{1}{t^2}\frac{\partial Y}{\partial p^0}, \quad v_2 = \frac{1}{t^2}\frac{\partial X}{\partial q_0}, \quad u_2 = \frac{1}{t}\frac{\partial Y}{\partial q_0}, \tag{3.30}$$

are bounded (on the time interval defined by Proposition 3.1). Let us consider them as elements of the Banach space of continuous matrix-valued functions on $[0, t]$. Differentiating (3.11) with respect to p^0, q^0 and using (3.10) yields

$$\frac{\partial P}{\partial p^0} = 1 + v_1 O(tc^3 + c^4 + t^2c^2) + u_1 O(c^3 + tc^2) + \frac{\partial P}{\partial p^0} O(t + c) + \frac{\partial Q}{\partial p^0} O(t + c + \frac{c^2}{t}), \tag{3.31}$$

$$\frac{\partial P}{\partial q_0} = v_2 O(c^3 + \frac{c^4}{t} + tc^2) + u_2 O(\frac{c^3}{t} + c^2) + \frac{\partial P}{\partial q_0} O(t + c) + \frac{\partial Q}{\partial q_0} O(t + c + \frac{c^2}{t}). \tag{3.32}$$

Similarly, from (3.13)

$$\frac{\partial Q}{\partial p^0} = t\alpha_0 + v_1 O(t^2c^3 + tc^4 + t^3c^2) + u_1 O(tc^3 + t^2c^2) + \frac{\partial P}{\partial p^0} O(t^2 + ct) + \frac{\partial Q}{\partial p^0} O(t^2 + c^2), \tag{3.33}$$

$$\frac{\partial Q}{\partial q_0} = 1 + t(\beta_0' + \gamma_0' y^0) + v_2 O(tc^3 + c^4 + t^2c^2) + u_2 O(c^3 + tc^2)$$
$$+ \frac{\partial P}{\partial q_0} O(t^2 + tc) + \frac{\partial Q}{\partial q^0} O(t^2 + tc + c^2). \tag{3.34}$$

From (3.33) one has

$$\frac{\partial Q}{\partial p^0} = (1 + \delta^2)\left[t\delta \frac{\partial P}{\partial p^0} + v_1 t\delta^4 + u_1 t\delta^3 + t\alpha_0\right]. \tag{3.35}$$

Inserting this in (3.31) yields

$$\frac{\partial P}{\partial p^0} = 1 + \delta \frac{\partial P}{\partial p^0} + v_1\delta^4 + u_1\delta^3,$$

and therefore

$$\frac{\partial P}{\partial p^0} = 1 + \delta + v_1\delta^4 + u_1\delta^3, \tag{3.36}$$

$$\frac{\partial Q}{\partial p^0} = t(\alpha_0 + \delta + v_1\delta^4 + u_1\delta^3). \tag{3.37}$$

Similarly, from (3.32),(3.34) one gets

$$\frac{\partial P}{\partial q_0} = \frac{1}{t}(\delta^2 + v_2\delta^4 + u_2\delta^3), \tag{3.38}$$

$$\frac{\partial Q}{\partial q_0} = 1 + t(\beta_0' + \gamma_0' y^0) + \delta^2 + v_2\delta^4 + u_2\delta^3. \tag{3.39}$$

Furthermore, differentiating (3.16) with respect to p_0, q_0 yields

$$u_1 = \frac{1}{2} g_0 \alpha_0' + v_1\delta^4 + u_1\delta^3 + \frac{\partial P}{\partial p_0}\delta + \frac{\partial Q}{\partial p_0}\frac{\delta^2}{t},$$

$$u_2 = g_0 + \frac{t}{2}(\omega + \Omega y^0) + v_2\delta^4 + u_2\delta^3 + \frac{\partial P}{\partial q_0}\delta t + \frac{\partial Q}{\partial q_0}\delta^2.$$

By (3.36)-(3.39), this implies

$$u_1 = \frac{1}{2} g_0 \alpha_0' + \delta + v_1\delta^4 + u_1\delta^3,$$

$$u_2 = g_0 + \frac{t}{2}(\omega + \Omega y^0) + \delta^2 + v_2\delta^4 + u_2\delta^3.$$

Similarly, differentiating (3.19) yields

$$v_1 = -\frac{1}{6}\alpha_0 g_0 \alpha_0' + \delta + v_1\delta^4 + u_1\delta^3,$$

$$v_2 = -\frac{1}{2}\alpha_0 g_0 - \frac{t}{6}(\omega' + \Omega' y^0) + \delta^2 + v_2\delta^4 + u_2\delta^3.$$

From the last 4 equations one easily obtains

$$u_1 = \frac{1}{2} g_0 \alpha_0' + \delta, \quad u_2 = g_0 + \frac{t}{2}(\omega + \Omega y^0) + \delta^2,$$

$$v_1 = -\frac{1}{6}\alpha_0 g_0 \alpha_0' + \delta, \quad v_2 = -\frac{1}{2}\alpha_0 g_0 - \frac{t}{6}(\omega' + \Omega' y^0) + \delta^2,$$

which is equivalent to (3.26)-(3.29). Formulas (3.30), (3.31) then follow from (3.36)-(3.39).

We shall prove now the main result of this section.

Theorem 3.1. *(i) There exist positive real numbers c and t_0 (depending only on x_0) such that for all $t \leq t_0$ and $\|y\| \leq c/t$, the mapping $(p^0, q_0) \mapsto (X, Y)(t, x_0, y^0, p^0, q_0)$ defined on the polydisc $B_{c^3/t^3} \times B_{c^2/t^2}$ is a diffeomorphism onto its image.*

(ii) There exists $r > 0$ such that by reducing c and t_0 if necessary, one can assume that the image of this diffeomorphism contains the polydisc $B_{r/t}(\tilde{y}) \times B_r(\tilde{x})$. These c, t_0, r can be chosen smaller than an arbitrary positive number δ.

(iii) Assume the matrix $(\alpha(x)g(x)\alpha'(x))^{-1}$ is uniformly bounded. Then c, t_0, r do not depend on x_0.

Proof. (i) From Proposition 3.2, one gets

$$\begin{pmatrix} X \\ Y \end{pmatrix}(t, x_0, y^0, p^0, q_0) - \begin{pmatrix} X \\ Y \end{pmatrix}(t, x_0, y^0, \pi^0, \xi_0) = tG_\delta(t, x_0)\begin{pmatrix} p^0 - \pi^0 \\ q_0 - \xi_0 \end{pmatrix},$$

where

$$G_\delta(t, x_0) = \begin{pmatrix} -\frac{1}{6}t^2(\alpha_0 g_0 \alpha_0' + \delta) & -\frac{1}{2}t(\alpha_0 g_0 + \delta) \\ \frac{1}{2}t(g_0\alpha_0' + \delta) & g_0 + \delta \end{pmatrix}. \tag{3.40}$$

To prove (i), it is therefore enough to prove that the matrix (3.40) is invertible. We shall show that its determinant does not vanish. Using the formula

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det d \cdot \det(a - bd^{-1}c) \tag{3.41}$$

(which holds for arbitrary matrices a, b, c, d with invertible d) one gets

$$\det G_\delta(t, x_0) = \det(g_0 + \delta)$$

$$\times \det\left[-\frac{1}{6}t^2(\alpha_0 g_0 \alpha_0' + \delta) + \frac{1}{4}t^2(\alpha_0 g_0 + \delta)(g_0^{-1} + \delta)(g_0\alpha_0' + \delta)\right].$$

The second factor is proportional to $\det(\alpha_0 g_0 \alpha_0' + \delta)$ and therefore, neither factor vanishes for small δ.

(ii) As in the proof of Theorem 2.1, one notes that the existence of (x, y) such that $(x, y) = (X, Y)(t, x_0, y^0, p^0, q_0)$ is equivalent to the existence of a fixed point of the map

$$F_{x,y}\begin{pmatrix} p^0 \\ q_0 \end{pmatrix} = \begin{pmatrix} p^0 \\ q_0 \end{pmatrix} + t^{-1}G(t, x_0)^{-1}\begin{pmatrix} x - X(t, x_0, y^0, p^0, q_0) \\ y - Y(t, x_0, y^0, p^0, q_0) \end{pmatrix},$$

where

$$G(t, x_0) = \begin{pmatrix} -\frac{1}{6}t^2\alpha_0 g_0 \alpha_0' & -\frac{1}{2}t(\alpha_0 g_0 + \frac{1}{3}t(\omega' + \Omega' y^0)) \\ \frac{1}{2}tg_0\alpha_0' & g_0 + \frac{t}{2}(\omega + \Omega y^0) \end{pmatrix}.$$

By (3.22), (3.23),

$$F_{x,y}\begin{pmatrix} p^0 \\ q_0 \end{pmatrix} = t^{-1}G(t, x_0)^{-1}\begin{pmatrix} x - \tilde{x} + \delta^4 \\ y - \tilde{y} + \delta^4/t \end{pmatrix}.$$

To estimate $G(t, x_0)^{-1}$ notice first that $G(t, x_0)$ is of form (3.40) and therefore it is invertible. Moreover, by the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} (a - bd^{-1}c)^{-1} & -a^{-1}b(d - ca^{-1}b)^{-1} \\ -d^{-1}c(a - bd^{-1}c)^{-1} & (d - ca^{-1}b)^{-1} \end{pmatrix} \tag{3.42}$$

(which holds for an invertible matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with invertible blocks a, d), one gets that $G(t, x_0)$ has the form

$$\begin{pmatrix} O(t^{-2}) & O(t^{-1}) \\ O(t^{-1}) & O(1) \end{pmatrix}.$$

Therefore, to prove that $F_{x,y}(p^0, q_0) \in B_{c^3/t^3} \times B_{c^2/t^2}$ one must show that

$$O(\|x - \tilde{x}\| + t\|y - \tilde{y}\| + \delta^4) \leq c^3,$$

which is certainly satisfied for small c, t_0 and r.

(iii) Follows directly from the above proof, because all parameters depend on the estimates for the inverse of the matrix (3.40).

Since $\tilde{x} - x_0 = O(t)$ for any fixed y^0, it follows that for arbitrary (x_0, y^0) and sufficiently small t, the polydisc $B_{r/2}(x_0) \times B_{r/2t}(y_0)$ belongs to the polydisc $B_r(\tilde{x}) \times B_{r/t}(\tilde{y})$. Therefore, due to Theorem 3.1, all the assumptions of Proposition 1.1 are satisfied for Hamiltonians (3.4), and consequently, all the results of Proposition 1 and its corollaries hold for these Hamiltonians. We shall prove now, following the same line of arguments as in the previous section, that the two-point function $S(t, x, y; x_0, y^0)$ in fact gives the global minimum for the corresponding functional (1.12). To get the necessary estimates of S, we need a more precise formula for the inverse of the matrix (3.40).

Let us express $G_\delta(t, x_0)$ in the form

$$G_\delta(t, x_0) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{g_0} \end{pmatrix} \begin{pmatrix} -\frac{1}{6}t^2(A_0A_0' + \delta) & -\frac{t}{2}(A_0 + \delta) \\ \frac{1}{2}t(A_0' + \delta) & 1 + \delta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{g_0} \end{pmatrix},$$

where $A_0 = \alpha_0\sqrt{g_0}$. Denoting the matrix $g_0^{-1/2}$ by J and using (3.42) yields that $G_\delta(t, x_0)^{-1}$ equals

$$\begin{pmatrix} 1 & 0 \\ 0 & J \end{pmatrix} \begin{pmatrix} \frac{12}{t^2}(A_0A_0')^{-1}(1 + \delta) & \frac{3}{t}(A_0A_0')^{-1}A_0\beta^{-1}(1 + \delta) \\ -\frac{6}{t}A_0'(A_0A_0')^{-1}(1 + \delta) & -\beta^{-1} + \delta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & J \end{pmatrix}, \tag{3.43}$$

where $-\beta = 1 - \frac{3}{2}A_0'(A_0A_0')^{-1}A_0$. In the simplest case, when $k = n$, i.e. when the matrix $\alpha(x)$ is square non-degenerate, (3.43) reduces to

$$G_\delta(t, x_0)^{-1} = \begin{pmatrix} \frac{12}{t^2}(\alpha_0 g_0 \alpha_0')^{-1}(1 + \delta) & \frac{6}{t}(g_0\alpha_0')^{-1}(1 + \delta) \\ -\frac{6}{t}(\alpha_0 g_0)^{-1}(1 + \delta) & -2g_0^{-1} + \delta \end{pmatrix}. \tag{3.44}$$

To write down the formula for the general case, let us decompose $Y = \mathcal{R}^k$ as the orthogonal sum $Y = Y_1 \oplus Y_2$, where $Y_2 = Ker\, A_0$. Let us denote again by A_0 the restriction of A_0 to Y_1. Then A_0^{-1} exists and $\beta|_{Y_1} = -\frac{1}{2}$, $\beta|_{Y_2} = 1$. With respect to this decomposition,

$$G_\delta(t, x)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & J_{11} & J_{12} \\ 0 & J_{21} & J_{22} \end{pmatrix}$$

$$\times \begin{pmatrix} \frac{12}{t^2}(A_0A_0')^{-1}(1+\delta) & \frac{6}{t}(A_0')^{-1}(1+\delta) & \frac{\delta}{t} \\ -\frac{6}{t}A_0^{-1}(1+\delta) & -2+\delta & \delta \\ \frac{\delta}{t} & \delta & 1+\delta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & J_{11} & J_{12} \\ 0 & J_{21} & J_{22} \end{pmatrix}. \quad (3.45)$$

From (1.19), (3.30), (3.31) one gets

$$\frac{\partial^2 S}{\partial(x,y)^2}(t,x,y,x_0,y^0) = \frac{1}{t}\begin{pmatrix} 1+\delta & \delta^2/t \\ t(\alpha_0'+\delta & 1+\delta \end{pmatrix} G_\delta(t,x_0)^{-1}. \quad (3.46)$$

In the case $n=k$, this takes the form

$$\frac{\partial^2 S}{\partial(x,y)^2}(t,x,y,x_0,y^0) = \begin{pmatrix} \frac{12}{t^3}(\alpha_0 g_0\alpha_0')^{-1}(1+\delta) & \frac{6}{t^2}(g_0\alpha_0')^{-1}(1+\delta) \\ \frac{6}{t^2}(\alpha_0 g_0)^{-1}(1+\delta) & \frac{4}{t}g_0^{-1}+\delta \end{pmatrix}, \quad (3.47)$$

and in the general case, from (3.45)-(3.46) one gets (omitting some arguments for brevity) that $\partial^2 S/\partial(x,y)^2$ equals

$$= \begin{pmatrix} \frac{12}{t^3}((A_0A_0')^{-1}+\delta) & \frac{6}{t^2}((A_0')^{-1}J_{11}+\delta) & \frac{6}{t^2}((A_0')^{-1}J_{12}+\delta) \\ \frac{6}{t^2}(J_{11}A_0^{-1}+\delta) & \frac{1}{t}(4J_{11}^2+J_{12}J_{21}+\delta) & \frac{1}{t}(4J_{11}J_{12}+J_{12}J_{22}+\delta) \\ \frac{6}{t^2}(J_{21}A_0^{-1}+\delta) & \frac{1}{t}(4J_{21}J_{11}+J_{22}J_{21}+\delta) & \frac{1}{t}(4J_{21}J_{12}+J_{22}^2+\delta) \end{pmatrix}. \quad (3.48)$$

Therefore, from (1.22) one gets the following.

Proposition 3.3. *For small t and (x,y) in a neighbourhood of (x_0,y^0) the two-point function $S(t,x,y;x_0,y^0)$ is equal to*

$$\frac{6}{t^3}\left((A_0^{-1}+\delta)(x-\tilde{x}), A_0^{-1}(x-\tilde{x})\right) + \frac{6}{t^2}\left((J_{11}+\delta)(y_1-\tilde{y}_1), (A_0^{-1}+\delta)(x-\tilde{x})\right)$$
$$+\frac{6}{t^2}\left((J_{12}+\delta)(y_2-\tilde{y}_2), (A_0^{-1}+\delta)(x-\tilde{x})\right)$$
$$+\frac{1}{2t}\left(y_1-\tilde{y}_1, (4J_{11}^2+J_{12}J_{21}+\delta)(y_1-\tilde{y}_1)\right)$$
$$+\frac{1}{2t}\left(y_2-\tilde{y}_2, (4J_{21}J_{12}+J_{22}^2+\delta)(y_2-\tilde{y}_2)\right)$$
$$+\frac{1}{t}\left(y_1-\tilde{y}_1, (4J_{11}J_{12}+J_{21}J_{22}+\delta)(y_2-\tilde{y}_2)\right). \quad (3.49)$$

Expanding $\tilde{x},\tilde{y}$ in t we can present this formula in terms of x_0,y^0. Let us write down the main term (when $\delta=0$), for brevity only in the case $k=n$:

$$S_0(t,x,y,x_0,y^0) = \frac{6}{t^3}(g_0^{-1}\alpha_0^{-1}(x-x_0), \alpha_0^{-1}(x-x_0))$$
$$+\frac{6}{t^2}(g_0^{-1}\alpha_0^{-1}(x-x_0), y+y_0+2\alpha_0^{-1}a_0)$$
$$+\frac{2}{t}\left[(y,g_0^{-1}y)+(y,g_0^{-1}y_0)+(y_0,g_0^{-1}y_0)\right]$$

$$+\frac{6}{t}\left[(y+y_0, g_0^{-1}\alpha_0^{-1}a_0) + (g_0^{-1}\alpha_0^{-1}a_0, \alpha_0^{-1}a_0)\right].$$

Equivalently, this can be expressed in a manifestly positive form

$$S_0(t,x,y,x_0,y^0) = \frac{1}{2t}(y-y_0, g_0^{-1}(y-y_0)) + \frac{6}{t^3}(g_0^{-1}z, z) \qquad (3.50)$$

with

$$z = \alpha_0^{-1}(x-x_0) + \frac{t}{2}(y+y_0+2\alpha_0^{-1}a_0).$$

Using these expressions one proves the following property of extremals quite similarly to the proof of Proposition 2.4.

Proposition 3.4. *There exists $r' \le r$ such that for $(x,y) \in B_{r'/2}(x_0) \times B_{r'/2t}(y_0)$, the solution of the boundary value problem $(x,y)(0) = (x_0,y^0)$ and $(x,y)(t) = (x,y)$ for system (3.5), which exists and is unique under the additional assumption $(p^0,q_0) \in B_{c^3/t^3} \times B_{c^2/t^2}$, furnishes the absolute minimum for the functional (1.12) corresponding to the Hamiltonian (3.4).*

Proposition 3.5. *Let the absolute minimum of the functional (1.12) for Hamiltonian (3.4) be given by a piecewise-smooth curve. Then this curve is in fact a smooth characteristic and it contains no conjugate points.*

Proof. The first part is proved as in Proposition 2.5. Let us prove the Jacobi condition for the degenerate Hamiltonian (3.4), i.e. that a minimising characteristic does not contain conjugate points. First of all we claim that if a solution $(x(\tau),p(\tau))$ of the Hamiltonian system (1.1) corresponding to an arbitrary Hamiltonian of form (2.1) with non-negative-definite (but possibly singular) matrix G furnishes a minimum for the functional (1.12), then the curve $\eta(\tau)=0$ gives a minimum for the functional (1.12) defined on curves with fixed endpoints $\eta(0)=\eta(t)=0$ and corresponding to the quadratic time dependent Hamiltonian $\tilde{H}_t$ of form (1.17). In fact, let the extremal $(x(\tau),p(\tau))$ furnish a minimum for (1.12) corresponding to the Hamiltonian (2.1). This implies that for arbitrary $\epsilon>0$ and smooth $\eta(\tau)$ such that $\eta(0)=\eta(t)=0$:

$$\int_0^t \max_w [(p(\tau)+\epsilon w)(\dot{x}(\tau)+\epsilon\dot{\eta}(\tau))$$

$$-H(x+\epsilon\eta, p(\tau)+\epsilon w) - p(\tau)\dot{x}(\tau) + H(x(\tau),p(\tau))]\,d\tau \ge 0.$$

Since H is a polynomial in p of second degree, one has

$$H(x+\epsilon\eta, p+\epsilon w) = H(x,p) + \epsilon\frac{\partial H}{\partial x}(x,p)\eta + \epsilon\frac{\partial H}{\partial p}(x,p)w + O(\epsilon^3)$$

$$+\frac{1}{2}\epsilon^2\Big[\left(\frac{\partial^2 H}{\partial x^2}(x,p)\eta,\eta\right) + 2\left(\frac{\partial^2 H}{\partial x\partial p}(x,p)(\eta+O(\epsilon)),w\right)$$

$$+\left(\left(\frac{\partial^2 H}{\partial p^2}(x,p)+O(\epsilon)\right)w,w\right)\Big]$$

with $O(\epsilon)$ independent of w. Substituting this expression in the previous formula, integrating by parts the term $p(\tau)\dot{\eta}$, then using equations (1.1) and dividing by ϵ^2 yields

$$\int_0^t \max_w [w\dot{\eta} - \frac{1}{2}\left(\frac{\partial^2 H}{\partial x^2}(x,p)\eta, \eta\right) - 2\left(\frac{\partial H}{\partial x \partial p}(x,p)(\eta + O(\epsilon)), w\right)$$

$$-\frac{1}{2}\left(\left(\frac{\partial^2 H}{\partial p^2}(x,p) + O(\epsilon)\right) w, w\right)]\, d\tau + O(\epsilon) \geq 0.$$

Taking the limit as $\epsilon \to 0$ one gets

$$\int_0^t \max_w (w\dot{\eta} - \tilde{H}_t(\eta, w))\, d\tau \geq 0,$$

as claimed.

Now let H have the form (3.4). Then $\tilde{H}_t$ has this form as well, only it is time dependent. One sees easily that with this additional generalisation the analogue of Theorem 3.1 is still valid and therefore, the first part of Proposition 3.5 as well. Suppose now that on a characteristic $x(\tau)$ of (3.5) that furnishes minimum for the corresponding functional (1.12), the points $x(s_1), x(s_2)$ are conjugate, where $0 \leq s_1 < s_2 < t$. Then there exists a solution v, w of the Jacobi equation (1.16) on the interval $[s_1, s_2]$ such that $v(s_1) = v(s_2) = 0$ and v is not identically zero. Then the curve $\tilde{v}(s)$ on $[0, t]$, which is equal to v on $[s_1, s_2]$ and vanishes outside this interval, gives a minimum for the functional (1.12) corresponding to Hamiltonian $\tilde{H}_t$. But this curve is not smooth at $\tau = s_2$, which contradicts to the first statement of Proposition 3.5.

We can now prove the analogue of Tonelli's theorem for the Hamiltonian (3.4), namely the global existence of the boundary value problem.

Proposition 3.6. *For arbitrary $t > 0$ and arbitrary x_0, y^0, x, y, there exists a solution of the Hamiltonian system (3.5) with boundary conditions $(x, y)(0) = (x_0, y^0)$, $(x, y)(t) = (x, y)$, which furnish global minimum for the corresponding functional (1.12).*

Proof. The only difference from the proof of Proposition 3.5 is that the radius of balls can depend on y^0, but this is not of importance, because the proof is given by means of exhausting $\mathcal{R}^m$ by compact sets.

As a consequence, we have

Proposition 3.7. *Propositions 2.7 and 2.8 hold also for Hamiltonians of the form (3.4).*

To conclude we give some estimates on the derivatives of the two-point function

Proposition 3.8. *For arbitrary j and $l \leq j$*

$$\frac{\partial^j S}{\partial x^l \partial y^{j-l}}(t, \tilde{x}, \tilde{y}, x_0, y^0) = t^{-(l+1)} R(t, y_0),$$

where $R(t, y_0)$ has a regular asymptotic expansion in powers of t and y^0.

Proof. This is proved by induction on j using (1.19).

This Proposition, together with (1.19), suggests that the function $tS(t, \tilde{x} + x, \tilde{y} + y; x_0, y^0)$ can be expressed as a regular asymptotic expansion in the variables x/t and y. This important consequence will be used in the next chapter for effective calculations of the two-point function. We shall also need there estimates for the higher derivatives of the solutions of the Cauchy problem for (3.5) with respect to initial momenta, which one easily gets from Theorem 3.1 together with the Taylor expansion of the solutions up to any order.

Proposition 3.9. *Let x^0, x^1, p^0, p^1 denote respectively x, y, p, q. The following estimates hold*

$$\frac{\partial^2 X^I}{\partial p_0^J \partial p_0^K} = O\left(t^{6-I-J-K}\right).$$

More generally, if H has sufficiently many bounded derivatives, then

$$\frac{\partial^K X^I}{\partial p_0^{I_1} ... \partial p_0^{I_K}} = O\left(t^{3K-I-I_1-...-I_K}\right).$$

4. General regular Hamiltonians depending quadratically on momenta

We now consider here general regular Hamiltonians (RH). These are the Hamiltonians for which, roughly speaking, the boundary-value problem enjoys the same properties as for their quadratic (or Gaussian) approximation. As we shall see in the next chapter, the main term of the small time asymptotics for the corresponding diffusion is then also the same as for the Gaussian diffusion approximation. In fact, the motivation for the following definition will be better seen when we consider formal power series solutions of the corresponding Hamilton-Jacobi equation in the next chapter, but rigorous proofs seem to be simpler to carry out for boundary value-problem for Hamiltonian systems.

Since the Gaussian diffusions were classified in the previous chapter by means of the Young schemes, it is clear that RH should also be classified by these schemes.

Definition. *Let $\mathcal{M} = \{m_{M+1} \geq m_M \geq ... \geq m_0 > 0\}$ be a non-degenerate sequence of positive integers (Young scheme). Let X^I denote Eucleadian space $\mathcal{R}^{m_I}$ of dimension m_I with coordinates x^I, $I = 0, ..., M$, and $Y = X^{M+1} = \mathcal{R}^{M+1}$. Let $p_I, I = 0, ..., M$, and $q = p^{M+1}$ be the momenta corresponding to x^I and y respectively. The $\mathcal{M}$-degree, $\deg_{\mathcal{M}} P$, of a polynomial P in the variables $x^1, ..., x^M, y = x^{M+1}$ is by definition the degree, which one gets prescribing the degree I to the variable x^I, $I = 0, ..., M+1$. A RH corresponding to a given Young scheme is by definition a function of the form*

$$H(x, y, p, q) = \frac{1}{2}(g(x^0)q, q) - R_1(x, y)p_0 - ...$$

$$-R_{M+1}(x,y)p_M - R_{M+2}(x,y)q - R_{2(M+2)}(x,y), \tag{4.1}$$

where the $R_I(x,y)$ are (vector-valued) polynomials in the variables $x^1,...x^M, y = x^{M+1}$ of the $\mathcal{M}$-degree $\deg_{\mathcal{M}} R_I = I$ with smooth coefficients depending on x^0, and $g(x^0)$ depends only on the variable x^0 and is nondegenerate everywhere. Moreover, the matrices $\frac{\partial R_I}{\partial x^I}$ (which, due to the condition on $\deg_{\mathcal{M}} R_I$, depend only on x^0) have everywhere maximal rank, equal to m_{I-1}, and the polynomial $R_{2(M+1)}$ is bounded from below. When the coefficients of the polynomials R_I are uniformly bounded in x_0, we shall say that the RH has bounded coefficients.

All results of the previous section hold for this more general class of Hamiltonians with clear modifications. The proofs are similar, but with notationally heavier. We omit the details and give only the main estimates for the derivatives of the solution of the corresponding Hamilton system with respect to the initial momenta. These estimates play a central role in all proofs. To obtain these estimates, one should choose the convenient coordinates in a neighbourhood of initial point, which were described in the previous chapter, in Theorem 1.2.1. Let us note also that the assumption of the boundedness of the coefficients of the polynomials in (4.1) insures the uniformity of all estimates with respect to the initial value x_0^0, and is similar to the assumptions of boundedness of the functions A, V, G defining the non-degenerate Hamiltonians of Section 2.

Theorem 4.1. *There exist positive constants K, t_0, c_0 such that for all $c \in (0, c_0]$, $t \in (0, t_0]$ the solution of the Hamiltonian system (1.1) corresponding to the regular Hamiltonian (4.1) exists on the interval $[0,t]$ whenever the initial values of the variables satisfy the estimates*

$$|x_0^1| \leq \frac{c}{t}, ..., |x_0^{M+1}| \leq \left(\frac{c}{t}\right)^{M+1},$$

$$|p_0^{M+1}| \leq \left(\frac{c}{t}\right)^{M+2}, ..., |p_0^1| \leq \left(\frac{c}{t}\right)^{2M+2}, \quad |p_0^0| \leq \left(\frac{c}{t}\right)^{2M+3}.$$

On the interval $0 < t < t_0$ the growth of the solution is governed by the estimates

$$|X(t)^I| \leq K\left(1 + \left(\frac{c}{t}\right)^I\right), \ |P(t)^I| \leq K\left(1 + \left(\frac{c}{t}\right)^{2M+3-I}\right), \ I = 0, ..., M+1, \tag{4.2}$$

the derivatives with respect to initial momenta have the form

$$\left[\frac{\partial(X^0, ..., X^{M+1})}{\partial(P_0^0, ..., P_0^{M+1})}\right]_{IJ} = t^{2M+3-I-J}\beta_{IJ}(1 + O(t)), \tag{4.3}$$

$$\left[\frac{\partial(X^0, ..., X^{M+1})}{\partial(P_0^0, ..., P_0^{M+1})}\right]^{-1}_{IJ} = t^{-(2M+3-I-J)}\gamma_{IJ}(1 + O(t)), \tag{4.4}$$

where β_{IJ}, γ_{IJ} are matrices of the maximal rank $\min(m_I, m_J)$, and for higher derivatives one has the estimates

$$\frac{\partial^K X^I}{\partial p_0^{I_1} ... \partial p_0^{I_K}} = O\left(t^{(3+2M)K-I-I_1-...-I_K}\right). \tag{4.5}$$

Clearly, the Lagrangians corresponding to degenerate Hamiltonians are singular. However, it turns out that the natural optimisation problems corresponding to degenerate regular Hamiltonians are problems of the calculus of variations for functionals depending on higher derivatives. To see this, consider first a Hamiltonian (3.1) such that $n = k$ and the map $y \mapsto a(x,y)$ is a diffeomorphism for each x. Then the change of variables $(x,y) \mapsto (x,z)$: $z = -a(x,y)$ implies the change of the momenta $p_{old} = p_{new} - (\partial a/\partial x)^t q_{new}$, $q_{old} = -(\partial a/\partial y)^t p_{new}$ and the Hamiltonian (3.1) takes the form

$$\frac{1}{2}\left(\frac{\partial a}{\partial y} g(x,y(x,z))(\frac{\partial a}{\partial y})^t q, q\right) + (z,p)$$

$$+\left(\frac{\partial a}{\partial y} b(x,y(x,z)) - \frac{\partial a}{\partial x} z, q\right) - V(x,y(x,z)).$$

In the case of Hamiltonian (3.4), the new Hamiltonian takes the form

$$\frac{1}{2}\left(\alpha(x)g(x)\alpha^t(x)q, q\right) + (z,p)$$

$$+\left(\alpha(x)(b(x) + \beta(x)y + \frac{1}{2}(\gamma(x)y,y)) - (\frac{\partial a}{\partial x} + \frac{\partial \alpha}{\partial x} y, z), q\right) - V(x,y)$$

with $y = -\alpha(x)^{-1}(z + a(x))$. This Hamiltonian is still regular of form (3.4), but at the same time it has the form (1.36) of a Hamiltonian corresponding to the problem of the calculus of variations with Lagrangian depending on first and second derivatives. Therefore, all results of the previous section correspond to the solution of the problems of that kind. In general, not all regular Hamiltonians (4.1) can be transformed to the form (1.36) but only a subclass of them. General results on regular Hamiltonians give the existence and the estimates for the solutions of problems with Lagrangian (1.35). For example, one has the following result.

Theorem 4.2. *Let $x_0, ..., x_n \in \mathcal{R}^n$ and let a smooth function L be given by the formula*

$$L(x_0, ..., x_m, z) = \frac{1}{2}\left(g(x_0)(z + \alpha(x_0, ..., x_m), z + \alpha(x_0, ..., x_m)\right) + V(x_0, ..., x_m), \tag{4.6}$$

with $g(x_0)$ being strictly positive-definite matrix, α and V being polynomials of m-degree $m+1$ and $2(m+1)$ respectively, where $\deg_m x_j = j$, and V being positive. Then there exists a solution of equation (1.30) with boundary conditions (1.25), which provides the absolute minimum for functional (1.24).

One also can specify the additional conditions under which this solution is unique for small times.

We have noticed that the Hamiltonians with the Young scheme whose entries are equal may correspond to the Lagrangians depending on $(M+2)$ derivatives. The general RH with the Young scheme $(m_{M+1}, m_M, ..., m_0)$ may correspond to the variational problems with Lagrangians depending on $M+2$ derivatives of

m_0 variables, $M+1$ derivatives of $(m_1 - m_0)$ variables and so on. Theorem 4.2 can be generalised to cover these cases as well.

To conclude this section let us give a simple example of non-regular Hamiltonian, whose quadratic approximation is regular of the first rabk (at least in a neighbourhood of almost every point) but for which the solution of the boundary value problem does not exist even locally. In this simplest case x and y are one-dimensional and H even does not depend on x. Let

$$H(x,y,p,q) = -f(y)p + \frac{1}{2}q^2 \tag{4.7}$$

with an everywhere positive function f, which therefore can not be linear (consequently H is not regular). The corresponding Hamiltonian system has the form

$$\dot{x} = -f(y), \quad \dot{y} = q, \quad \dot{p} = 0, \quad \dot{q} = f'(y)p.$$

Therefore $\dot{x}$ is always negative and there is no solution of the Hamiltonian system joining (x_0, y_0) and (x, y) whenever $x > x_0$, even for small positive t. On the other hand, if $f(y)$ is a nonlinear diffeomorphism, say $f(y) = y^3 + y$, then it is not difficult to prove the global existence of the solutions to the boundary value problem for the corresponding Hamiltonian (4.7), though H is still non-regular. In fact, regularity ensures not only the existence of the solutions but also some "nice" asymptotics for them.

5. Hamiltonians of exponential growth in momenta

In this section we generalise partially the results of Section 2 to some non-degenerate Hamiltonians, which are not quadratic in momentum. First we present a theorem of existence and local uniqueness for a rather general class of Hamiltonians and then give some asymptotic formulas for the case mainly of interest, when the Hamiltonians increase exponentially in momenta.

Definition 1. *We say that a smooth function $H(x,p)$ on $\mathcal{R}^{2m}$ is a Hamiltonian of uniform growth, if there exist continuous positive functions $C(x)$, $\kappa(x)$ on $\mathcal{R}^n$ such that*

(i) $\frac{\partial^2 H}{\partial p^2}(x,p) \geq C^{-1}(x)$ for all x, p;

(ii) if $|p| \geq \kappa(x)$, the norms of all derivatives of H up to and including the third order do not exceed $C(x)H(x,p)$ and moreover,

$$C^{-1}(x)H(x,p) \leq \|\frac{\partial H}{\partial p}(x,p)\| \leq C(x)H(x,p), \tag{5.1}$$

$$\|\frac{\partial g}{\partial p}(x,p)\| \leq C(x)H(x,p)\frac{\partial^2 H}{\partial p^2}(x,p), \tag{5.2}$$

where

$$g(x,p) = \frac{\partial^2 H}{\partial p \partial x}\frac{\partial H}{\partial p} - \frac{\partial^2 H}{\partial p^2}\frac{\partial H}{\partial x}; \tag{5.3}$$

(iii) for some positive continuous function $\delta(x)$ one has

$$\left| \left(\frac{\partial^2 H}{\partial p^2}(x,p) \right)^{-1} \right| \leq C(x) \left| \left(\frac{\partial^2 H}{\partial p^2}(x+y,p+q) \right)^{-1} \right| \tag{5.4}$$

whenever $|y| \leq \delta$, $|q| \leq \delta$, $|p| \geq \kappa(x)$.

The main properties of the boundary-value problem for such Hamiltonians are given in Theorems 5.1, 5.2 below. The function $H(x,p) = \alpha(x)\cosh p$ with $\alpha(x) > 0$ is a simple example. In fact in this book we are interested in the finite-dimensional generalisations of this example, which are described in Theorem 5.3 below.

Let a Hamiltonian of uniform growth be given. Following the same plan of investigation as for quadratic Hamiltonians we study first the Cauchy problem for the corresponding Hamiltonian system.

Proposition 5.1. *For an arbitrary neighbourhood $U(x_0)$ of x_0 there exist positive K, c_0, t_0 such that if $H(x_0,p_0) \leq c/t$ with $c \leq c_0$, $t \leq t_0$, then the solution $X(s) = X(s,x_0,p_0)$, $P(s) = P(s,x_0,p_0)$ of the Hamiltonian system exists on $[0,t]$; moreover, on this interval $X(s) \in U(x_0)$, $\|P(s)-p_0\| \leq K(t+c)$ and*

$$\frac{\partial X(s)}{\partial p_0} = s\frac{\partial^2 H}{\partial p^2}(x_0,p_0)(1+O(c)), \quad \frac{\partial P(s)}{\partial p_0} = 1 + O(c). \tag{5.5}$$

If in addition, the norms of the derivatives of H of order up to and including k do not exceed $C(x)H(x,p)$ for large p, then $\frac{\partial^l X(s)}{\partial p_0^l} = O(c)$ for $l \leq k-2$. In particular,

$$\frac{\partial^2 X(s)}{\partial p_0^2}(s) = s\frac{\partial^3 H}{\partial p^3}(x_0,p_0) + O(c^2).$$

Proof. We can suppose that $\|p_0\| > 2\max_{x\in U(x_0)} \kappa(x)$, because the case of p_0 from any fixed compact is considered trivially). Let

$$T(t) = \min(t, \sup\{s > 0 : X(s) \in U(x_0), \quad \|P(s)\| > \kappa\}).$$

Using Definition 5.1 and the conservation of H along the trajectories of the Hamiltonian flow one obtains

$$\|\dot{P}(s)\| \leq \left| \frac{\partial H}{\partial x} \right| \leq C(X(s))H(X(s),P(s)) = C(X(s))H(x_0,p_0) \leq C(X(s))\frac{c}{t};$$

hence $|P(s) - p_0| = O(c)$ and similarly $|X(s) - x_0| = O(c)$ for $s \leq T(t)$. If one chooses small c_0 in such a way that the last inequalities would imply $X(s) \in U(x_0)$ and $|P(s)| > \kappa(X(s))$, the assumption $T(t) < t$ would lead to a contradiction. Consequently, $T(t) = t$ for such c_0. It remains to prove (5.5). For p_0 (or, equivalently, c/t) from any bounded neighbourhood of the origin the result is trivial. Let us suppose therefore again that $|p_0|$ is large enough. Following the lines of the proof of Lemma 2.2 let us differentiate the integral form of the Hamiltonian equations

$$\begin{cases} X(s) = x_0 + s\frac{\partial H}{\partial p}(x_0,p_0) + \int_0^s (s-\tau)g(X,P)(\tau)\,d\tau \\ P(s) = p_0 - \int_0^s \frac{\partial H}{\partial x}(X,P)(\tau)\,d\tau \end{cases}, \tag{5.6}$$

to obtain

$$\begin{cases} \frac{\partial X(s)}{\partial p_0} = s\frac{\partial^2 H}{\partial p^2}(x_0,p_0) + \int_0^s (s-\tau)\left(\frac{\partial g}{\partial x}\frac{\partial X}{\partial p_0} + \frac{\partial g}{\partial p}\frac{\partial P}{\partial p_0}\right)(\tau)\,d\tau \\ \frac{\partial P(s)}{\partial p_0} = 1 - \int_0^s \left(\frac{\partial^2 H}{\partial x^2}\frac{\partial X}{\partial p_0} + \frac{\partial^2 H}{\partial x \partial p}\frac{\partial P}{\partial p_0}\right) d\tau \end{cases}.$$

Considering now the matrices $v(s) = \frac{1}{s}\frac{\partial X(s)}{\partial p_0}$ and $u(s) = \frac{\partial P(s)}{\partial p_0}$ as vectors of the Banach space of continuous $m \times m$-matrix-valued functions $M(s)$ on $[0,t]$ with the norm $\sup\{\|M(s)\| : s \in [0,t]\}$ one deduces from the previous equations that

$$\begin{cases} v = \frac{\partial^2 H}{\partial p^2}(x_0,p_0) + O(t^2)H^2(x_0,p_0)v + O(t)\max_s |\frac{\partial g}{\partial p}(X(s),P(s))|u \\ u = 1 + O(t^2)H(x_0,p_0)v + O(t)H(x_0,p_0)u \end{cases}.$$

Due to (5.2),(5.4),

$$\left|\frac{\partial g}{\partial p}(X(s),P(s))\right| \left|\left(\frac{\partial^2 H}{\partial p^2}(x_0,p_0)\right)^{-1}\right|$$

$$\leq C(x_0)\left|\frac{\partial g}{\partial p}(X(s),P(s))\right| \left|\left(\frac{\partial^2 H}{\partial p^2}(x(s),P(s))\right)^{-1}\right|$$

$$\leq H(X(s),P(s))C(X(s))C(x_0) = C(X(s))C(x_0)H(x_0,p_0),$$

and thus the previous system of equations can be written in the form

$$\begin{cases} v = \frac{\partial^2 H}{\partial p^2}(x_0,p_0) + O(c^2)v + O(c)\frac{\partial^2 H}{\partial p^2}(x_0,p_0)u \\ u = 1 + O(tc)v + O(c)u \end{cases}.$$

From the second equation one gets

$$u = (1+O(c))(1+(tc)v), \tag{5.7}$$

and inserting this in the first one yields

$$v = \frac{\partial^2 H}{\partial p^2}(x_0,p_0)(1+O(c)) + O(c^2)v,$$

and consequently

$$v = (1+O(c))\frac{\partial^2 H}{\partial p^2}(x_0,p_0),$$

which yields the first equation in (5.5). Inserting it in (5.7) yields the second equation in (5.5). Higher derivatives with respect to p_0 can be estimated similarly by differentiating (5.6) sufficient number of times and using induction. Proposition is proved.

Theorem 5.1. *For any $c \leq c_0$ with small enough c_0 there exists t_0 such that for all $t \leq t_0$ the map $p_0 \mapsto X(t,x_0,p_0)$ defined on the domain $D_c =$*

$\{p_0 : H(x_0,p_0) \leq c/t\}$ is a diffeomorphism on its image, which contains the ball $B_{cr}(x_0)$ and belongs to the ball $B_{cr^{-1}}(x_0)$ for some r (that can be chosen arbitrary close to $C^{-1}(x_0)c$ whenever c is small enough).

Proof. From Proposition 5.1 one concludes that $\frac{1}{t}\frac{\partial X(t)}{\partial p_0}$ is bounded from below in D_c by some positive constant. This implies that the map under consideration is a diffeomorphism on its image, which one shows by the same arguments as in the proof of Theorem 2.1. To estimate this image, let us estimate $X(t,x_0,p_0)-x_0$ on the boundary of the domain D_c, namely when $H(x_0,p_0)=c/t$. From (5.6) it follows that

$$\|X(t,x_0,p_0)-x_0\| = \|t\frac{\partial H}{\partial p}(x_0,p_0) + \int_0^t (t-s)g(x,p)(s)\,ds\|$$

$$\geq tC^{-1}(x_0)H(x_0,p_0) - t^2C(x)^2H(x_0,p_0)^2 = tc^{-1}(x_0)H(x_0,p_0)(1+O(c)).$$

Since the image of the boundary of D_c is homeomorphic to S^{m-1} and therefore divides the space into two open connected components, it follows from the last estimate that the ball with the centre at x_0 and the radius rt belongs to the image of D_c, where r can be chosen arbitrary close to $C^{-1}(x_0)c$, if c is sufficiently small. Similarly one proves that

$$\|X(t,x_0,p_0)-x_0\| \leq tC(x)H(x_0,p_0)(1+O(c)),$$

which implies the required upper bound for the image of D_c.

Proposition 5.2. *There exist t_0, r, c such that if $|x-x_0| \leq rc$, the solution to the boundary value problem with the condition $x(0)=x_0$, $x(t)=x$, for the Hamiltonian system with the Hamiltonian H exists for all $t \leq t_0$ and is unique under additional condition that $H(x_0,p_0) \leq c/t$. If $|x-x_0|/t$ be outside a fixed neighbourhood of the origin, then the initial momentum p_0 and the initial velocity $\dot{x}_0$ on this solution satisfy the estimates*

$$r\frac{|x-x_0|}{t} \leq H(x_0,p_0) \leq \frac{|x-x_0|}{rt}, \quad \dot{x}_0 = \frac{x-x_0}{t}(1+O(|x-x_0|)) \tag{5.8}$$

and the two-point function $S(t,x,x_0)$ on this solution has the form

$$S(t,x,x_0) = tL(x_0, \frac{x-x_0}{t}(1+O(|x-x_0|))) + O(|x-x_0|^2). \tag{5.9}$$

If $C(x), C^{-1}(x)$, $\kappa(x)$, $\kappa^{-1}(x)$, $\delta(x)$ from Definition 5.1 are bounded (for all x uniformly), the constants t_0, r, c can be chosen independently of x_0.

Proof. Everything, except for the formula for S, follows directly from the previous theorem and the estimates used in its proof. To prove (5.9), we write

$$S(t,x,x_0) = \int_0^t L(x_0+O(|x-x_0|), \dot{x}_0 + O(|x-x_0|)H^2(x_0,p_0))\,d\tau$$

$$= tL(x_0, \frac{x-x_0}{t}(1+O(|x-x_0|)))+O(t|x-x_0|)\frac{\partial L}{\partial x}(x_0+O(c), \frac{x-x_0}{t}(1+O(c))),$$

which implies (5.9), since

$$\frac{\partial L}{\partial x}(x,v) = -\frac{\partial H}{\partial x}\left(x, \frac{\partial L}{\partial v}(x,v)\right). \tag{5.10}$$

Proposition 5.3. *For $\dot{x}_0$ (or equivalently $(x-x_0)/t$) from a bounded neighbourhood of $\frac{\partial H}{\partial p}(x,0)$*

$$\frac{1}{t}S(t,x;x_0) = L(x_0, \frac{x-x_0}{t}) + O(t); \tag{5.11}$$

moreover, if $H(x,0) = 0$ for all x, then

$$L(x_0,\dot{x}_0) = \frac{1}{2}\left(\left[\frac{\partial^2 H}{\partial p^2}(x_0,0)\right]^{-1}\left(\dot{x}_0 - \frac{\partial H}{\partial p}(x_0,0)\right), \dot{x}_0 - \frac{\partial H}{\partial p}(x_0,0)\right) + ..., \tag{5.12}$$

$$\frac{1}{t}S(t,x,x_0) = \frac{1}{2}(1+O(t))\left(\left[\frac{\partial^2 H}{\partial p^2}(x_0,0)\right]^{-1}\frac{x-\tilde{x}(t,x_0)}{t}, \frac{x-\tilde{x}(t,x_0)}{t}\right) + ..., \tag{5.13}$$

where ... in (5.12),(5.13) denote the higher terms of the expansion with respect to $\dot{x}_0 - \frac{\partial H}{\partial p}(x_0,0)$ and

$$\frac{x-\tilde{x}(t,x_0)}{t} = \frac{x-x_0}{t} - \frac{\partial H}{\partial p}(x_0,0) + O(t)$$

respectively, and each coefficient in series (5.12) differs from the corresponding coefficient of (5.13) by the value of the order $O(t)$.

Remark. The number of available terms in asymptotic series (5.12) or (5.13) depends of course on the number of existing derivatives of H.

Proof. Formula (5.11) is proved similarly to (5.9). Next, since $H(x,0) = 0$, the Lagrangian $L(x_0,v)$ (resp. the function $S(t,x;x_0)$) has its minimum equal to zero at the point $v = \frac{\partial H}{\partial p}(x_0,0)$ (resp. at $x = \tilde{x}(t,x_0)$, where $\tilde{x}$ is as usual the solution of the Hamiltonian system with initial data $x(0) = x_0, p(0) = 0$). At last, one compares the coefficients in series (5.12),(5.13) using (1.19) and the obvious relations

$$\frac{\partial^k X}{\partial p_0^k} = t\frac{\partial^{k+1} H}{\partial p_0^{k+1}}(x_0,p_0) + O(c^2), \quad \frac{\partial^k P}{\partial p_0^k} = O(t), \quad k > 1, \tag{5.14}$$

which hold for p_0 from any bounded domain.

Now we can prove the smooth equivalence of the two-point functions of the Hamiltonian $H(x,p)$ and the corresponding Hamiltonian $H(x_0,p)$ with a fixed x_0. This result plays a key role in the construction of the semi-classical asymptotics for the Feller processes given in Chapter 6.

Theorem 5.2. *Let the assumptions of Theorem 5.1 hold and $H(x,0)=0$ for all x. Then there exists a smooth map $z(t,v,x_0)$ defined for v from the ball of the radius rc/t such that*

(i) for fixed t, x_0 the map $v \mapsto z(t,v,x_0)$ is a diffeomorphism on its image,

(ii) for v from a bounded domain

$$\|z(t,x,x_0)-v\| = O(t)\|v\| + O(t), \tag{5.15}$$

(ii) if v is outside a neighbourhood of the origin, then

$$z(t,v,x_0) = (1+\omega(t,v))D_t v + O(t), \tag{5.16}$$

where $\omega(t,v)=O(|x-x_0|)$ is a scalar function and D_t is a linear diffeomorphism of $\mathcal{R}^d$ of the form $1+O(t)$ with a uniformly bounded derivative in t;

(iii) z takes $S(t,x,x_0)$ into $tL(x,(x-x_0)/t)$, i.e.

$$L(x_0, z(t, \frac{x-x_0}{t}, x_0) = \frac{1}{t} S(t,x,x_0). \tag{5.17}$$

Proof. It follows from Propositions 5.2, 5.3 and E2. More precisely, one repeats the proof of Propositions E1, E2 of Appendix E to obtain a diffeomorphism that takes the function $L(x_0,v)$ in the function $S(t,x,x_0)/t$ considered both as the functions of $v=(x-x_0)/t$ and depending on t, x_0 as parameters. Due to Proposition 5.3, the linear and the local parts D_3, D_2 of the required diffeomorphism have the form $1+O(t)$. Due to (5.9), the dilatation coefficient ω from (E1) has the order $O(|x-x_0|)$. To get (5.16) one needs then only take in account the necessary shift on the difference between minimum points of L and S which is of the order $O(t)$ due to Proposition 5.3.

We shall concentrate now on a more concrete class of Hamiltonians and shall obtain for these Hamiltonians more exact estimates of the objects introduced above. For any vector p we shall denote by $\bar{p}$ a unit vector in the direction p, i.e. $\bar{p}=p/\|p\|$.

Definition 5.2. *We say that a smooth function $H(x,p)$ on $\mathcal{R}^{2m}$ is a Hamiltonian of exponential growth, if there exist two positive continuous functions $a(x,\bar{p}) \geq b(x,\bar{p})$, on $\mathcal{R}^m \times S^{m-1}$ and a positive continuous function $C(x)$ on $\mathcal{R}^m$ such that*

(i) for p outside a neighbourhood of the origin, the norms of all derivatives of H up to and including the third order do not exceed $C(x)H(x,p)$;

(ii) $H(x,p) \leq C(x)\exp\{a(x,\bar{p})|p|\}$ for all x,p;

(iii) $\frac{\partial^2 H}{\partial p^2}(x,p) \geq C^{-1}(x)\exp\{b(x,\bar{p})|p|\}$ for all x,p.

Notice that the condition on the growth of the Hamiltonian implies that the matrix of the second derivatives of the corresponding Lagrangian $L(x,\dot{x})$ tends to zero, as $\dot{x}\to\infty$ (see Remark after Theorem 5.3), which means that the corresponding problem of the calculus of variations has certain degeneracy at infinity. Nevertheless, similarly to the case of Hamiltonians from Definition 5.1, one can show the existence of the solution to the boundary-value problem (with the condition $x(0) = x_0, x(t) = x$) for the Hamiltonian systems with Hamiltonians of exponential growth and the uniqueness of such solution for $|x - x_0| < t^\Delta$, where $\Delta \in (0, 1)$ depends on the difference $a(x, \bar{p}) - b(x, \bar{p})$ and can be chosen arbitrary small whenever this difference can be chosen arbitrary small. We are not going into detail, because actually we are interested in a more restrictive class of Hamiltonians that satisfy both Definitions 5.1 and 5.2. This class of Hamiltonians is provided by the Lévy-Khintchine formula with the Lévy measure having finite support, namely, the Hamiltonians of this class are given by the formula

$$H(x,p) = \frac{1}{2}(G(x)p,p) - (A(x),p) + \int_{\mathcal{R}^m \setminus \{0\}} \left(e^{-(p,\xi)} - 1 + \frac{(p,\xi)}{1+\xi^2} \right) d\nu_x(\xi), \tag{5.18}$$

where $G(x)$ is a nonnegative matrix, ν_x is a so called Lévy measure on $\mathcal{R}^m \setminus \{0\}$, which means that

$$\int_{\mathcal{R}^m \setminus \{0\}} \min(\xi^2, 1)\, d\nu_x(\xi) < \infty$$

for all x, and the support of the Lévy measure ν is supposed to be a bounded set in $\mathcal{R}^m$. The last assumption insures that function (5.18) is well defined for all complex p and is an entire function with respect to p. We suppose that all $G(x), A(x), \nu_x$ are continuously differentiable (at least thrice). Notice that

$$\left(\frac{\partial^2 H}{\partial p^2}(x,p)v, v \right) = (G(x)v, v) + \int_{\mathcal{R}^m \setminus \{0\}} (\xi, v)^2 e^{-(p,\xi)}\, d\nu_x(\xi),$$

and therefore $\frac{\partial^2 H}{\partial x^2}$ is always nonnegative. Moreover, one sees directly that for Hamiltonian (5.18), properties (i),(ii) from Definition 5.2 hold with the function a being the support function of the set $-supp\,\nu_x$, i.e.

$$a(x, \bar{p}) = \max\{(\bar{p}, -\xi) : -\xi \in supp\,\nu_x\}. \tag{5.19}$$

The following results give simple sufficient conditions on ν_x that ensure that corresponding function (5.18) is a Hamiltonian of exponential growth. We omit rather simple proofs.

Proposition 5.4. *(i) Let the function β on $\mathcal{R}^m \times S^{m-1}$ be defined by the formula*

$$\beta(x, \bar{v}) = \sup\{r : r\bar{v} \in supp\,\nu_x\}.$$

If β is continuous and everywhere positive, then Hamiltonian (5.18) is of exponential growth with the function a defined in (5.19) and any continuous $b(x,p) < a(x,p)$.

(ii) If there exists $\epsilon > 0$ such that for any $\bar{v} \in S^{m-1}$ the convex hull of $supp\,\nu_x \cap \{\xi : (\xi, \bar{v}) \geq \epsilon\}$ depends continuously on x and has always nonempty

interior, then function (5.18) is of exponential growth with the function a defined as above and

$$b(x,\bar v)=\min_{w\in S^{m-1}}\max\{|(w,\xi)|:\xi\in supp\,\nu_x\cap\{\xi:(-\xi,\bar v)\geq\epsilon\}\}. \qquad (5.20)$$

Examples. Let $m=2$, G and A vanish in (5.18), and let $\nu_x=\nu$ does not depend on x. If the support of ν consists of only three points, then H of form (5.18) is not of exponential growth. Actually in this case $\frac{\partial^2 H}{\partial p^2}$ tends to zero, as p tends to infinity along some directions, and one can show that the boundary-value problem have no solution for some pairs of points x, x_0. On the other hand, if the support of ν consists of four vertices of a square with the centre at the origin, then H of form (5.18) again is not of exponential growth. However, it satisfies the condition of Proposition 5.4 (ii) with $\epsilon=0$, and one can prove that for this Hamiltonian the boundary-value problem is always solvable.

In order that function (5.18) would satisfy all conditions of Definition 5.1, it seems necessarily to make some assumptions on the behaviour of ν near the boundary of its support. We are not going to describe the most general assumptions of that kind. In the next statement we give only the simplest sufficient conditions.

Theorem 5.3. *Let ν_x have a convex support, containing the origin as an inner point, with a smooth boundary, $\partial\, supp\,\nu_x$, depending smoothly on x and having nowhere vanishing curvature, and moreover, let $\nu_x(d\xi)=f(x,\xi)\,d\xi$ in a neighbourhood of $\partial\, supp\,\nu_x$ with a continuous f not vanishing on $\partial\, supp\,\nu_x$. Then H of form (5.18) satisfies the requirements of both Definitions 5.1 and 5.2 with $a(x,\bar p)$ given by (5.19), and moreover for large p and some continuous $C(x)$*

$$C^{-1}(x)|p|^{-(m+1)/2}\exp\{a(x,\bar p)|p|\}\leq H(x,p)\leq C(x)|p|^{-(m+1)/2}\exp\{a(x,\bar p)|p|\}, \qquad (5.21)$$

$$\frac{\partial^2 H}{\partial p^2}(x,p)\geq C^{-1}(x)\frac{H(x,p)}{|p|},$$

$$C^{-1}(x)\frac{H^d(x,p)}{|p|^{d-1}}\leq\det\frac{\partial^2 H}{\partial p^2}(x,p)\leq C(x)\frac{H^d(x,p)}{|p|^{d-1}}, \qquad (5.22)$$

and

$$\max(|g(x,p)|,|\frac{\partial g}{\partial p}(x,p)|)\leq C(x)\frac{H(x,p)}{|p|},$$

Remark. If the above $f(x,\xi)$ vanishes at $\partial\, supp\,\nu_x$, but has a non-vanishing normal derivative there, then the same holds but with $m+2$ instead of $m+1$ in (5.21).

Proof. Clear that for p from any compact set H is bounded together with all its derivatives and the matrix of the second derivatives is bounded from below by a positive constant. It is also obvious that (5.22) implies (5.2). In order to obtain the precise asymptotics of H as $p\to\infty$ notice first that for large p

the behaviour of H and its derivatives is the same (asymptotically) as by the function

$$\tilde{H}(x,p) = \int_{U(x,\nu)} \exp\{|p|(\bar{p},\xi)\} f(x,-\xi)\, d\xi,$$

where $U(x,\nu)$ is an arbitrary small neighbourhood of $\partial\, supp\, \nu_x$. To estimate this integral we consider it as the Laplace integral with the large parameter $|p|$ (depending on the additional bounded parameter $\bar{p}$). The phase of this integral $S(\xi,\bar{p}) = (\xi,\bar{p})$ takes its maximum at the unique point $\xi_0 = \xi_0(\bar{p},x)$ on the boundary $\partial\, supp\, \nu_x$ of the domain of integration, the unit vector $\bar{p}$ provides an outer normal vector to this boundary at ξ_0, and the value of this maximum is given by the support function (5.19). Moreover, this maximum is not degenerate (due to the condition of not vanishing curvature) in the sense that the normal derivative of S at ξ_0 (the derivative with respect to ξ in the direction $\bar{p}$) does not vanish (because it equals $\bar{p}$) and the $(m-1)\times(m-1)$-matrix $A(x,\xi_0)$ of the second derivatives of S restricted to $\partial\, supp\, \nu_x$ at ξ_0 is not degenerate. Thus by the Laplace method (see e.g. Proposition B5) one finds for large $|p|$

$$H(x,p) = |p|^{-(d+1)/2} \exp\{a(x,\bar{p})|p|\} f(x,-\xi_0)$$
$$\times (2\pi)^{(d-1)/2} \left(\det A(x,\xi_0)\right)^{-1/2} (1 + O(|p|^{-1})). \tag{5.23}$$

Similarly one finds that $\frac{\partial H}{\partial p}(x,p)$ and $\left(\frac{\partial^2 H}{\partial p^2}(x,p)v,v\right)$ equal respectively to

$$|p|^{-(d+1)/2} \exp\{a(x,\bar{p})|p|\} f(x,-\xi_0)\xi_0$$
$$\times (2\pi)^{(d-1)/2} \left(\det A(x,\xi_0)\right)^{-1/2} (1 + O(|p|^{-1})). \tag{5.24}$$

and

$$|p|^{-(d+1)/2} \exp\{a(x,\bar{p})|p|\} f(x,-\xi_0)(\xi_0,v)^2$$
$$\times (2\pi)^{(d-1)/2} \left(\det A(x,\xi_0)\right)^{-1/2} (1 + O(|p|^{-1})). \tag{5.25}$$

Similarly one finds the asymptotic representations for other derivatives of H, which implies (5.1),(5.21) and the required upper bounds for all derivatives of H. To get a lower bound for the eigenvalues of the matrix of the second derivatives of H notice that due to the above formulas $\left(\frac{\partial^2 H}{\partial p^2}(x,p)v,v\right)$ is of the same order as $H(x,p)$ whenever v is not orthogonal to ξ_0. If $v = v_0$ is such that $(v_0,\xi_0) = 0$, then the major term of the corresponding asymptotic expansion vanishes, which means the drop in at least one power of $|p|$. To get (5.22) one must show that the second term in this expansion does not vanish for any such v_0. This follows from the general explicit formula for this term (see e.g. in Proposition B5) and the fact that the amplitude in the corresponding Laplace integral has zero of exactly second order at ξ_0. To complete the proof of the Proposition, it remains to note that writing down the major terms of the expansions of $\frac{\partial^2 H}{\partial p^2}, \frac{\partial H}{\partial x}, \frac{\partial^2 H}{\partial p \partial x}, \frac{\partial H}{\partial p}$ one sees that the terms proportional to $|p|^{-(d+1)}$ cancel in the expansions for g or its derivative in p, which implies the required estimates for g.

Remark. Notice that from the formulas

$$v = \frac{\partial H}{\partial p}\left(x, \frac{\partial L}{\partial v}(x,v)\right), \quad \frac{\partial^2 L}{\partial v^2}(x,v) = \left(\frac{\partial^2 H}{\partial p^2}\right)^{-1}\left(x, \frac{\partial L}{\partial v}(x,v)\right), \tag{5.26}$$

connecting the derivatives of H and its Legendre transform L, it follows that if H is a Hamiltonian from Theorem 5.3, the first (resp. the second) derivative of the corresponding Lagrangian $\frac{\partial L}{\partial v}(x,v)$ (resp. $\frac{\partial^2 L}{\partial v^2}(x,v)$ increases like $\log|v|$ (resp. decreases like $|v|^{-1}$) as $|v| \to \infty$.

Proposition 5.5. *For a Hamiltonian from Theorem 3.1, if $|x - x_0| \leq rc$, $t \leq t_0$ and $(x-x_0)/t$ does not approach the origin, one has the following estimates for the initial momentum $p_0 = p_0(t, x, x_0)$ and the two-point function $S(t, x, x_0)$ of the solution to the boundary-value problem with conditions $x(0) = x_0$, $x(t) = x$ (recall that the existence and uniqueness of this solution is proved in Theorem 5.1):*

$$|p_0|\left(1 + \frac{O(\log(1+|p_0|))}{|p_0|}\right) = \frac{1}{a(x_0, \bar{p}_0)}\log(1 + \frac{|x-x_0|}{t}) + O(1), \tag{5.27}$$

$$-\sigma t + C|x - x_0| \leq S(t, x; x_0) \leq \sigma t + C|x - x_0|. \tag{5.28}$$

with some constants σ, C.

Proof. Estimate (5.27) follow directly from (5.8) and (5.24). Next, from (5.22) one has

$$|\log t|^{-1}\frac{|x-x_0|}{Ct} \leq \frac{\partial^2 H}{\partial p^2}(x_0, p_0) \leq C\frac{|x-x_0|}{t} \tag{5.29}$$

for small t and some constant C, which implies, due to (5.5), that for θ not approaching zero

$$C^{-1}|x - x_0|^{-1} \leq \left(\frac{\partial X}{\partial p_0}\right)^{-1}(t, x, x_0 + \theta(x - x_0)) \leq |\log t|C|x - x_0|^{-1}. \tag{5.30}$$

Hence, from (1.22) and (5.13) one obtains

$$-\sigma t + C|x - x_0| \leq S(t, x; x_0) \leq \sigma t + C|x - x_0||\log t|.$$

In order to get rig of $\log t$ on the r.h.s. (which is not very important for our purposes) one needs to argue similarly to the proof of (5.31) below. We omit the details.

Now one can use the same arguments as in Section 2 to get the following

Proposition 5.6. *The statements (and the proofs) of Proposition 2.4-2.8 are valid for Hamiltonians from Theorem 3.1. In particular, for $t \leq t_0$, $|x-x_0| \leq r$ with small enough r, t_0 there exists a unique solution to the boundary value problem for the Hamiltonian system with Hamiltonian H that provides the global minimum for the corresponding problem of the calculus of variations.*

Further we shall need the estimates for the function z from Theorem 5.2 and its derivatives.

Proposition 5.7. *Let H belong to the class of Hamiltonians described in Theorem 3.1 and z be the corresponding mapping from Theorem 5.2. If $v = (x-x_0)/t$ does not approach the origin and $|x-x_0| \leq rc$, then*

$$p_0 - \frac{\partial L}{\partial v}(x_0, v) = O(|x-x_0|), \quad \frac{\partial L}{\partial v}(x_0, v) - \frac{\partial L}{\partial v}(x_0, z(t, v, x_0)) = O(|x-x_0|), \tag{5.31}$$

and

$$\frac{\partial \omega}{\partial v} = O(t), \quad \frac{\partial \omega}{\partial t} = O(|v|), \quad \frac{\partial^2 \omega}{\partial v^2} = O(t)|v|^{-1}|\log t|. \tag{5.32}$$

Proof. Plainly for any w

$$\frac{\partial L}{\partial v}(x_0, w) - \frac{\partial L}{\partial v}(x_0, v) = \int_0^1 \frac{\partial^2 L}{\partial v^2}(x_0, v + s(w-v))\, ds(w-v),$$

$$\leq \max_{s \in (0,1)} \left(\frac{\partial^2 H}{\partial p^2}(x_0, \frac{\partial L}{\partial v}(x_0, v + s(w-v))) \right)^{-1} (w-v).$$

First let $w = \dot{x}_0 = \frac{\partial H}{\partial p}(x_0, p_0)$. From the first equation in (5.6) and the estimates for g in Theorem 5.3 it follows that

$$|\dot{x}_0 - v| = O(t)|g(x,p)| = O(t)|p|^{-1}H^2(x,p)$$

$$= O(t)|\log t|^{-1}H^2(x,p) = O(t^{-1})|x-x_0|^2|\log t|^{-1}.$$

Therefore, due to (5.22), one has

$$\left| \frac{\partial L}{\partial v}(x_0, \dot{x}_0) - \frac{\partial L}{\partial v}(x_0, v) \right| = O(|\log t|) \frac{t}{|x-x_0|} \frac{|x-x_0|^2}{t} |\log t|^{-1} = O(|x-x_0|),$$

i.e. the first inequality in (5.31). Now let $w = z(t, v, x_0)$. In that case it follows from (5.16) that $|z(t,v,x_0) - v| = O(t^{-1})|x-x_0|^2$ only, i.e. without an additional multiplier of the order $|\log t|$ as in the previous situation. Therefore, the previous arguments would lead here to an additional multiplier of the order $|\log t|$ on the r.h.s. of the second inequality (5.31). Hence, one needs here a more careful consideration. Namely, as it was noted in the proof of Theorem 5.3, the matrix $\frac{\partial^2 H}{\partial p^2}(x_0, p)$ has the maximal eigenvalue of the order $H(x_0, p)$ and the corresponding eigenvector is asymptotically (for large p) proportional to $\xi_0 = \xi_0(x_0, \bar{p})$. Other eigenvalues are already of the order $|p|^{-1}H(x,p)$. Therefore, in order to obtain the required estimate it is enough to show that the ratio of the projection of the vector $z(t,v,x_0) - v$ on the direction of $\xi_0(x_0, \bar{p})$ (for p around p_0) and the projection of $z(t,v,x_0) - v$ on the perpendicular hyperplane is not less than of the order $|\log t|$. But the vector $z(t,v,x_0) - v$ is proportional to $v = (x-x_0)/t$, which in its turn is close to $\frac{\partial H}{\partial p}(x_0, p_0)$. Hence, one must prove that the vector $\frac{\partial H}{\partial p}(x_0,p_0)$ lies essentially in the direction of $\xi_0(x_0,\bar{p}_0)$. But this is surely true, because due to (5.24) the principle term of the asymptotics of $\frac{\partial H}{\partial p}(x_0,p_0)$ is proportional to ξ_0 and next terms differs exactly by the multiplier of the order $|\log t|^{-1}$.

Let us turn now to the proof of (5.32). Differentiating (5.17) with respect to x yields

$$\frac{\partial S}{\partial x} = (1+\omega)\frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0))D_t + \left(\frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0)), D_t v\right)\frac{\partial \omega}{\partial v}.$$

Hence

$$\frac{\partial \omega}{\partial v} = \frac{\frac{\partial S}{\partial x} - (1+\omega)\frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0))D_t}{\left(\frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0)), D_t v\right)}. \tag{5.33}$$

To estimate the denominator in this formula notice that

$$\left(\frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0)), D_t v\right) = \left(\frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0)), z\right)(1+O(|x-x_0|)$$

$$\geq H(x_0, \frac{\partial L}{\partial v}(x_0, z(t,\frac{x-x_0}{t},x_0)))(1+O(|x-x_0|)),$$

which is of the order $|x-x_0|/t$. Turning to the estimate of the nominator we present it as the sum of two terms, first being the difference between the final momentum $\frac{\partial S}{\partial x}$ and the initial momentum p_0 on a trajectory, which is of the order $O(|x-x_0|)$ due to Lemma 5.1, and the second being the difference between p_0 and $\frac{\partial L}{\partial v}(x_0,z)$, which is of the order $O(|x-x_0|)$ due to (5.31). Consequently, taking into account the estimates for the nominator and denominator gives the first estimate in (5.32).

Differentiating (5.17) with respect to t yields

$$\frac{\partial S}{\partial t} = L(x_0,z) + t\frac{\partial L}{\partial v}D_t\left(\frac{\partial \omega}{\partial t}v - \frac{1}{t}(1+\omega)\frac{\partial z}{\partial v}v\right)$$

up to a nonessential smaller term. Since S satisfies the Hamilton-Jacobi equation it follows that

$$\frac{\partial \omega}{\partial t} = \frac{(1+\omega)\left(\frac{\partial L}{\partial v}, \frac{\partial z}{\partial v}v\right) - H(x,\frac{\partial S}{\partial x}) - L(x_0,z)}{t\frac{\partial L}{\partial v}(x_0,z)v}.$$

The nominator is of the order $O(|x-x_0|^2)/t$, because the main term has the form

$$\frac{\partial L}{\partial v}(x_0,z)z - L(x_0,z) - H(x,\frac{\partial S}{\partial x})$$

$$= \left(H(x_0,\frac{\partial L}{\partial v}) - H(x,\frac{\partial L}{\partial v})\right) + \left(H(x,\frac{\partial L}{\partial v}) - H(x,\frac{\partial S}{\partial x})\right),$$

which is of the order $O(|x-x_0|^2)t^{-1}$, due to the estimates of the first order derivatives of H. The denominator is of the order $|x-x_0|$, which proves the second formula in (5.32).

Differentiating (5.17) two times in x yields

$$t\frac{\partial^2 S}{\partial x^2}-(1+\omega)\frac{\partial^2 L}{\partial v^2}-\frac{\partial \omega}{\partial v}\otimes\frac{\partial^2 L}{\partial v^2}v-\left(\frac{\partial L}{\partial v}\otimes\frac{\partial \omega}{\partial v}+\frac{\partial \omega}{\partial v}\otimes\frac{\partial L}{\partial v}\right)=\left(\frac{\partial L}{\partial v},v\right)\frac{\partial^2 \omega}{\partial v^2}. \tag{5.34}$$

The coefficient at $\frac{\partial^2\omega}{\partial v^2}$ in this formula was already found to be of the order $O(|x-x_0|/t)$. Thus one needs to show that the l.h.s. of (5.34) has the order $t|\log t|$. All tensor products on the l.h.s. of (5.34) certainly have this order due to the first estimate in (5.32). Next,

$$t\frac{\partial^2 S}{\partial x^2}=\left(\frac{\partial^2 H}{\partial p^2}\right)^{-1}(x_0,p_0)(1+O(|x-x_0|)=\left(\frac{\partial^2 H}{\partial p^2}\right)^{-1}(x_0,p_0)+O(t)|\log t|.$$

Therefore, it remains to show that

$$\left(\frac{\partial^2 H}{\partial p^2}\right)^{-1}(x_0,p_0)-\left(\frac{\partial^2 H}{\partial p^2}\right)^{-1}(x_0,\frac{\partial L}{\partial v}(x_0,z(v)))=O(t)|\log t|. \tag{5.35}$$

Using (5.31) and the mean value theorem for the difference on the l.h.s. of (5.35) leads directly to the estimate $O(t)|\log t|^2$ for this difference. But slightly more careful considerations, similar to those used in the proof of the second inequality (5.31) allow to decrease the power of $|\log t|$, which gives the required estimate.

Proposition 5.8. *Let H, z and v be the same as in Proposition 5.7. Then:*

$$\frac{\partial z}{\partial v}(t,v,x_0)=1+O(|x-x_0|),\quad \frac{\partial^2 z}{\partial v^2}(t,v,x_0)=O(t)(1+\log^+\frac{|x-x_0|}{t}), \tag{5.36}$$

where we used the usual notation $\log^+ M=\max(0,\log M)$.

Proof. For v outside a neighbourhood of the origin, it follows directly from Proposition 5.7. Notice only that for brevity we used always the estimate $|\log t|$ for $|p_0|$, but in formula (5.36) we have restored a more precise estimate for $|p_0|$ from (5.27). For the bounded v formulas (5.36) follow from Proposition 5.3 and explicit formulas for z from the proof of Lemma E2.

6. Complex Hamiltonians and calculus of variations for saddle-points

Here we discuss the solutions to the boundary-value problem for complex Hamiltonians depending quadratically on momenta. As we have seen in Sect.1, in the case of real Hamiltonians, a solution to the Hamiltonian system furnishes a minimum (at least locally) for a corresponding problem of calculus of variations. It turns out that for complex Hamiltonians the solutions of Hamiltonian equations have the property of a saddle-point. Let $x=y+iz\in\mathcal{C}^m$, $p=\xi+i\eta\in\mathcal{C}^m$ and

$$H=\frac{1}{2}(G(x)p,p)-(A(x),p)-V(x), \tag{6.1}$$

where

$$G(x) = G_R(x) + iG_I(x), \quad A(x) = A_R(x) + iA_I(x), \quad V(x) = V_R(x) + iV_I(x)$$

are analytic matrix-, vector-, and complex valued functions in a neighbourhood of the real plane. Moreover, G is non-degenerate, G_R, G_I are symmetric and G_R is non-negative for all x. Under these assumptions, one readily sees that all the results and proofs of Lemma 2.1, Lemma 2.2, Theorem 2.1 and Propositions 2.1-2.3 on the existence of the solutions of the Hamiltonian system and the asymptotic expansions for the two-point function are valid for this complex situation, where x_0 is considered to be any complex number in the domain where G, A, V are defined, and initial momenta are complex as well. In particular, one gets therefore the existence of the family $\Gamma(x_0)$ of complex characteristics joining x_0 with any point x from some complex domain $D(x_0)$ in time $t \leq t_0$, and the corresponding complex momentum field $p(t,x) = (\xi + i\eta)(t, y, z)$, which in its turn defines the complex invariant Hilbert integral (1.6). Let us clarify what optimisation problem is solved by the complex characteristics of the family $\Gamma(x_0)$.

Notice first that

$$Re\, H = \frac{1}{2}(G_R\xi, \xi) - \frac{1}{2}(G_R\eta, \eta) - (G_I\xi, \eta) - (A_R, \xi) + (A_I, \eta) - V_R, \qquad (6.2)$$

$$Im\, H = \frac{1}{2}(G_I\xi, \xi) - \frac{1}{2}(G_I\eta, \eta) + (G_R\xi, \eta) - (A_I, \xi) - (A_R, \eta) - V_I, \qquad (6.3)$$

and if $(x(s), p(s))$ is a complex solution to (1.1), then the pairs (y, ξ), (y, η), (z, ξ), (z, η) are real solutions to the Hamiltonian systems with Hamiltonians ReH, ImH, ImH, $-ReH$ respectively. Next, if G_R^{-1} exists, then

$$(G_R + iG_I)^{-1} = (G_R + G_I G_R^{-1} G_I)^{-1} - iG_R^{-1} G_I (G_R + G_I G_R^{-1} G_I)^{-1},$$

and if G_I^{-1} exists, then

$$(G_R + iG_I)^{-1} = G_I^{-1} G_R (G_I + G_R G_I^{-1} G_R)^{-1} - i(G_I + G_R G_I^{-1} G_R)^{-1}.$$

Therefore, since G_R, G_I are symmetric, $(G^{-1})_R, (G^{-1})_I$ are also symmetric. Moreover, $G_R > 0$ is equivalent to $(G^{-1})_R > 0$, and $G_I > 0$ is equivalent to $(G^{-1})_I < 0$.

By definition, the Lagrangian corresponding to the Hamiltonian H is

$$L(x, \dot{x}) = (p\dot{x} - H(x, p))|_{p=p(x)} \qquad (6.4)$$

with $p(x)$ uniquely defined from the equation $\dot{x} = \frac{\partial H}{\partial p}(x, p)$. Therefore, the formula for L is the same as in the real case, namely

$$L(x, \dot{x}) = \frac{1}{2}(G^{-1}(\dot{x} + A(x)), \dot{x} + A(x)) + V(x). \qquad (6.4')$$

Consequently

$$ReL(y,z,\dot y,\dot z)=\xi(y,z,\dot y,\dot z)\dot y-\eta(y,z,\dot y,\dot z)\dot z-ReH(y,z,\xi(y,z,\dot y,\dot z),\eta(y,z,\dot y,\dot z)), \tag{6.5}$$

$$ImL(y,z,\dot y,\dot z)=\eta(y,z,\dot y,\dot z)\dot y+\xi(y,z,\dot y,\dot z)\dot z-ImH(y,z,\xi(y,z,\dot y,\dot z),\eta(y,z,\dot y,\dot z)), \tag{6.6}$$

where $(\xi,\eta)(y,z,\dot y,\dot z)$ are defined from the equations

$$\dot y=\frac{\partial ReH}{\partial \xi}=G_R\xi-G_I\eta-A_R,\quad \dot z=-\frac{\partial ReH}{\partial \eta}=G_R\eta+G_I\xi-A_I. \tag{6.7}$$

Proposition 6.1 *For all* ξ,η

$$\xi\dot y-\eta(y,z,\dot y,\dot z)\dot z-ReH(y,z,\xi,\eta(y,z,\dot y,\dot z))\le ReL(y,z,\dot y,\dot z)$$

$$\le \xi(y,z,\dot y,\dot z)\dot y-\eta\dot z-ReH(y,z,\xi(y,z,\dot y,\dot z),\eta), \tag{6.8}$$

or equivalently

$$ReL(y,z,\dot y,\dot z)=\max_\xi\min_\eta(\xi\dot y-\eta\dot z-ReH(x,p))=\min_\eta\max_\xi(\xi\dot y-\eta\dot z-ReH(x,p)). \tag{6.8$'$}$$

In other words, $ReL(y,z,\dot y,\dot z)$ is a saddle-point for the function $(\xi\dot y-\eta\dot z-ReH(x,p))$. Moreover, ReL is convex with respect to $\dot y$ and concave with respect to $\dot z$ (strictly, if $G_R>0$ strictly). Furthermore, if $G_I\ge 0$, then

$$ImL(y,z,\dot y,\dot z)=\max_\xi\min_\eta(\xi\dot z+\eta\dot y-ImH(x,p))=\min_\eta\max_\xi(\xi\dot z+\eta\dot y-ImH(x,p)), \tag{6.9}$$

i.e. $Im(y,z,\dot y,\dot z)$ is a saddle-point for the function $\xi\dot z+\eta\dot y-ImH(x,p)$. Moreover, if $G_I\ge 0$, then ImL is convex with respect to $\dot z$ and concave with respect to $\dot y$ (strictly, if $G_I>0$ strictly).

Proof. Formula (6.7) is obvious, since the function $\xi\dot y-\eta\dot z-ReH(x,p)$ is concave with respect to ξ and convex with respect to η. Furthermore, $\frac{\partial^2 L}{\partial \dot x^2}=(G^{-1})(x)$ due to (6.4'). Therefore

$$\frac{\partial^2 ReL}{\partial \dot y^2}=Re\frac{\partial^2 L}{\partial \dot x^2}=(G^{-1})_R(x),\quad \frac{\partial^2 ReL}{\partial \dot z^2}=-Re\frac{\partial^2 L}{\partial \dot x^2}=-(G^{-1})_R(x),$$

which proves the required properties of ReL, because $(G^{-1})_R\ge 0$. The statements about ImL are proved similarly.

Consider now the complex-valued functional

$$I_t(x(.))=\int_0^t L(x(\tau),\dot x(\tau))\,d\tau,$$

defined on piecewise-smooth complex curves $x(\tau)$ joining x_0 and x in time t, i.e. such that $x(0)=x_0, x(t)=x$. As in Section 1, we define $S(t,x;x_0)=I_t(X(.))$, where $X(s)$ is the (unique) characteristic of the family $\Gamma(x_0)$ joining x_0 and x in time t.

Proposition 6.2. *The characteristic $X(s) = Y(s) + iZ(s)$ of the family $\Gamma(x_0)$ joining x_0 and x in time t is a saddle-point for the functional ReI_t, i.e. for all real piecewise smooth $y(\tau), z(\tau)$ such that $y(0) = y_0$, $z(0) = z_0$, $y(t) = y$, $z(t) = z$ and $y(\tau) + iZ(\tau)$, $Y(\tau) + iz(\tau)$ lie in the domain $D(x_0)$*

$$ReI_t(Y(.) + iz(.)) \le ReI_t(Y(.) + iZ(.)) = Re\, S(t, x, x_0) \le ReI_t(y(.) + iZ(.)). \tag{6.10}$$

In particular,

$$ReI_t(Y(.) + iZ(.)) = \min_{y(.)} \max_{z(.)} ReI_t(y(.) + iz(.)) = \max_{z(.)} \min_{y(.)} ReI_t(y(.) + iz(.)). \tag{6.11}$$

If $G_I(x) \ge 0$, then similar fact holds for ImI_t, namely

$$ImI_t(Y(.) + iZ(.)) = \min_{z(.)} \max_{y(.)} ImI_t(y(.) + iz(.)) = \max_{y(.)} \min_{z(.)} ImI_t(y(.) + iz(.)). \tag{6.12}$$

Proof. Let us prove, for example, the right inequality in (6.10). Notice

$$ReI_t(y(.), Z(.)) = \int_0^t (\xi(y, Z, \dot y, \dot Z)\dot y - \eta(y, Z, \dot y, \dot Z)\dot Z$$

$$-ReH(y, Z, \xi(y, Z, \dot y, \dot Z), \eta(y, Z, \dot y, \dot z)))(\tau)\, d\tau$$

$$\ge \int_0^t (\xi(y, Z)\dot y - \eta(y, Z, \dot y, \dot Z)\dot Z - ReH(y, Z, \xi(y, Z), \eta(y, Z, \dot y, \dot Z))(\tau)\, d\tau,$$

due to the left inequality in (6.8). The last expression can be written in equivalent form as

$$\int_0^t [\xi(y, Z)\dot y - \eta(y, Z)\dot Z - ReH(y, Z, \xi(y, Z), \eta(y, Z)](\tau)\, d\tau$$

$$+ \int_0^t [(\eta(y, Z) - \eta(y, Z, \dot y, \dot Z))\dot Z$$

$$+ReH(y, Z, \xi(y, Z), \eta(y, Z)) - ReH(y, Z, \xi(y, Z), \eta(y, Z, \dot y, \dot Z))](\tau)\, d\tau.$$

Let us stress (to avoid ambiguity) that in our notation, say, $\eta(y, z)(\tau)$ means the imaginary part of the momentum field in the point $(\tau, (y + iz)(\tau)$ defined by the family $\Gamma(x_0)$, and $\eta(y, z, \dot y, \dot z)$ means the solution of equations (6.7). Now notice that in the last expression the first integral is just the real part of the invariant Hilbert integral and consequently one can rewrite the last expression in the form

$$ReS(t, x; x_0) - \int_0^t [ReH(y, Z, \xi(y, Z), \eta(y, Z, \dot y, \dot Z)) - ReH(y, Z, \xi(y, Z), \eta(y, Z))$$

$$- \left(\eta(y, Z, \dot{y}, \dot{Z}) - \eta(y, Z), \frac{\partial ReH}{\partial \eta}(y, Z, \xi(y, Z), \eta(y, Z) \right)] \, d\tau.$$

The function under the second integral is negative (it is actually the real part of the Weierstrass function), since with respect to η the function ReH is concave. It follows that

$$ReI_t(y(.), Z(.)) \geq ReS(t, x : x_0) = ReI_t(Y(.), Z(.)).$$

Further on we shall deal mostly with a particular case of Hamiltonian (6.1), namely with the case of vanishing A and a constant G.

Proposition 6.3 *If the drift A vanishes and the diffusion matrix G is constant, then formula (2.19)-(2.21) hold. More exact formulas can be written as well:*

$$\frac{\partial^2 S}{\partial x^2} = \frac{1}{t} G^{-1} (1 + \frac{1}{3} t^2 \frac{\partial^2 V}{\partial x^2}(x_0) G + O(t^2 c)),$$

$$\frac{\partial^2 S}{\partial x_0^2} = \frac{1}{t} G^{-1} (1 + \frac{1}{3} t^2 \frac{\partial^2 V}{\partial x^2}(x_0) G + O(t^2 c)), \tag{6.13}$$

$$\frac{\partial^2 S}{\partial x \partial x_0} = -\frac{1}{t} G^{-1} (1 - \frac{1}{6} t^2 \frac{\partial^2 V}{\partial x^2}(x_0) G + O(t^2 c)). \tag{6.14}$$

and,

$$\frac{\partial X}{\partial p_0} = tG(1 + \frac{1}{6} t^2 \frac{\partial^2 V}{\partial x^2}(x_0) G + O(t^2 c)), \quad \frac{\partial P}{\partial p_0} = 1 + \frac{1}{2} t^2 \frac{\partial^2 V}{\partial x^2}(x_0) G + O(t^2 c), \tag{6.15}$$

where c is from Theorem 2.1.

Proof. Under the assumptions of the Proposition

$$X(t, x_0, p_0) = x_0 + tGp_0 + \frac{1}{2} G \frac{\partial V}{\partial x}(x_0) t^2 + \frac{t^3}{6} G \frac{\partial^2 V}{\partial x^2} G p_0 + O(t^4 p_0^2),$$

$$P(t, x_0, p_0) = p_0 + \frac{\partial V}{\partial x}(x_0) t + \frac{t^2}{2} \frac{\partial^2 V}{\partial x^2} G p_0 + O(t^3 p_0^2).$$

This implies (6.15) and also the estimate

$$\frac{\partial X}{\partial x_0} = 1 + \frac{1}{2} t^2 G \frac{\partial^2 V}{\partial x^2}(x_0) + O(t^2 c).$$

These estimates imply (6.13), (6.14) due to (1.19)-(1.21).

In the theory of semiclassical approximation , it is important to know whether the real part of the action S is nonnegative.

Proposition 6.4. *(i) If G_R is strictly positive for all x, then $ReS(t, x; x_0)$ restricted to real values x, x_0 is nonnegative and convex for small enough t and $x - x_0$.*

(ii) Let G_R and the drift A vanish for all x, and let G_I be a constant positive matrix, which is proportional to the unit matrix. Then ReS restricted to real values x, x_0 is nonnegative and convex for small enough $t, x - x_0$ iff V_R is nonnegative and strictly convex with respect to $y = Re\, x$.

Proof. (i) Follows directly from representation (2.17).

(ii) It follows from (6.13), (6.14) that (under assumptions (ii)) $Re\, S(t, x, x_0)$ is convex in x and x_0 for real x and x_0 whenever V_R is convex for real V. Consequently, to prove the positivity of $Re\, S$ it is enough to prove the positivity of $S(t, \tilde{x}, x_0)$ for all x_0, because, this is a minimum of S, as a function of x_0. Using expansion (2.17) yields

$$S(t, \tilde{x}, x_0) = tV(x_0) + O(t^3)\| \frac{\partial V}{\partial x} \|^2.$$

Since $V(x)$ is nonnegative, it follows that $V(\hat{x}_0) \geq 0$ at the point $\hat{x}_0$ of its global minimum. The previous formula implies directly that $S(t, \tilde{x}, x_0)$ is positive (for small t at least) whenever $V(\hat{x}_0) > 0$. If $V(\hat{x}_0) = 0$, then S is clearly non-negative outside a neighbourhood of $\hat{x}_0$. Moreover, in the neighbourhood of $\hat{x}_0$, it can be written in the form

$$S(t, \tilde{x}, x_0) = \frac{t}{2}\left(\frac{\partial^2 V}{\partial x^2}(\hat{x}_0)(x_0 - \hat{x}_0), x_0 - \hat{x}_0 \right) + O(t|x_0 - \hat{x}_0|^3) + O(t^3|x_0 - \hat{x}_0|^2),$$

which is again non-negative for small t.

7. Stochastic Hamiltonians

The theory developed in the previous Sections can be extended to cover the stochastic generalisations of Hamiltonian systems, namely the system of the form

$$\begin{cases} dx = \frac{\partial H}{\partial p}\, dt + g(t, x) \circ dW \\ dp = -\frac{\partial H}{\partial x}\, dt - (c'(t, x) + pg'(t, x)) \circ dW, \end{cases} \tag{7.1}$$

where $x \in \mathcal{R}^n,\ t \geq 0$, $W = (W^1, ...W^m)$ is the standard m-dimensional Brownian motion ($\circ$, as usual, denotes the Stratonovich stochastic differential), $c(t, x)$ and $g(t, x) = g_{ij}(t, x)$ are given vector-valued and respectively $(m \times n)$-matrix -valued functions and the Hamiltonian $H(t, x, p)$ is convex with respect to p. Stochastic Hamiltonian system (7.1) correspond formally to the singular Hamiltonian function

$$H(t, x, p) + (c(t, x) + pg(t, x))\dot{W}(t),$$

where $\dot{W}$ is the white noise (formal derivative of the Wiener process). The corresponding stochastic Hamilton-Jacobi equation clearly has the form

$$dS + H(t, x, \frac{\partial S}{\partial x})\, dt + (c(t, x) + g(t, x)\frac{\partial S}{\partial x}) \circ dW = 0. \tag{7.2}$$

To simplify the exposition we restrict ourselves to the most important particular case, when $g = 0$ in (7.1) and the functions H and c do not depend explicitly on t. Namely, we shall consider the stochastic Hamiltonian system

$$\begin{cases} dx = \frac{\partial H}{\partial p}\,dt \\ dp = -\frac{\partial H}{\partial x}\,dt - c'(x)\,dW. \end{cases} \tag{7.3}$$

and the stochastic Hamilton-Jacobi equation

$$dS + H(x, \frac{\partial S}{\partial x})\,dt + c(x)\,dW = 0. \tag{7.4}$$

In that case the Ito and the Stratonovich differentials coincide. The generalisation of the theory to (7.1) and (7.2) is almost straightforward. As the next stage of simplification we suppose that the matrix of the second derivative of H with respect to all its arguments is uniformly bounded. An example of this situation is given by the standard quantum mechanical Hamiltonian $p^2 - V(x)$. In that important for the application case one can get rather nice results on the existence of the solution to the boundary-value problem uniform with respect to the position of the boundary values x_0, x. However, the restriction to this type of Hamiltonians is by no means necessary. More general Hamiltonians that was discussed in Sections 2-6 can be considered in this framework similarly and the result are similar to those obtained for the deterministic Hamiltonian systems of Sections 2-6.

Theorem 7.1 [K1], [K2]. *For fixed $x_0 \in \mathcal{R}^n$ and $t > 0$ let us consider the map $P : p_0 \mapsto X(t, x_0, p_0)$, where $X(\tau, x_0, p_0)$, $P(\tau, x_0, p_0)$ is the solution to (7.3) with initial values (x_0, p_0). Let all the second derivatives of the functions H and c are uniformly bounded, the matrix $Hess_p H$ of the second derivatives of H with respect to p is uniformly positive (i.e. $Hess_p H \geq \lambda E$ for some constant λ), and for any fixed x_0 all matrices $Hess_p H(x_0, p)$ commute. Then the map P is a diffeomorphism for small $t \leq t_0$ and all x_0.*

Proof. Clear that the solution of the linear matrix equation

$$dG = B_1 G\,dt + B_2(t)\,dW, \quad G|_{t=0} = G_0, \tag{7.5}$$

where $B_j = B_j(t, [W])$ are given uniformly bounded and non-anticipating functionals on the Wiener space, can be represented by the convergent series

$$G = G_0 + G_1 + G_2 + ... \tag{7.6}$$

with

$$G_k = \int_0^t B_1(\tau) G_{k-1}(\tau)\,d\tau + \int_0^t B_2(\tau) G_{k-1}(\tau)\,dW(\tau). \tag{7.7}$$

Differentiating (7.3) with respect to the initial data (x_0, p_0) one gets that the matrix

$$G = \frac{\partial(X, P)}{\partial(x_0, p_0)} = \begin{pmatrix} \frac{\partial X}{\partial x_0} & \frac{\partial X}{\partial p_0} \\ \frac{\partial P}{\partial x_0} & \frac{\partial P}{\partial p_0} \end{pmatrix} (x(\tau, [W]), p(\tau, [W]))$$

satisfies a particular case of (7.5):

$$dG = \begin{pmatrix} \frac{\partial^2 H}{\partial p \partial x} & \frac{\partial^2 H}{\partial p^2} \\ -\frac{\partial^2 H}{\partial x^2} & -\frac{\partial^2 H}{\partial x \partial p} \end{pmatrix} (X,P)(t)\, G\, dt - \begin{pmatrix} 0 & 0 \\ c''(X(t)) & 0 \end{pmatrix} G\, dW \tag{7.8}$$

with G_0 being the unit matrix. Let us denote by $\tilde{O}(t^\alpha)$ any function that is of order $O(t^{\alpha-\epsilon})$ for any $\epsilon > 0$, as $t \to 0$. Applying the log log law for stochastic integrals [Ar] first to the solutions of system (7.3) and then calculating G_1 by (7.7) we obtain

$$G_1 = \left(t \begin{pmatrix} \frac{\partial^2 H}{\partial p \partial x} & \frac{\partial^2 H}{\partial p^2} \\ -\frac{\partial^2 H}{\partial x^2} & -\frac{\partial^2 H}{\partial x \partial p} \end{pmatrix} (x_0, p_0) + \begin{pmatrix} 0 & 0 \\ c''(x_0)\tilde{O}(t^{1/2}) & 0 \end{pmatrix} \right) \begin{pmatrix} E & 0 \\ 0 & E \end{pmatrix}$$

up to a term of order $\tilde{O}(t^{3/2})$. Application of the log log law to the next terms of series (7.5) yields for the remainder $G - G_0 - G_1$ the estimate $\tilde{O}(t^{3/2})$. Thus, we have the convergence of series (7.5) for system (7.8) and the following approximate formula for its solutions:

$$\frac{\partial X}{\partial x_0} = E + t \frac{\partial^2 H}{\partial p \partial x}(x_0, p_0) + \tilde{O}(t^{3/2}), \quad \frac{\partial X}{\partial p_0} = t \frac{\partial^2 H}{\partial p^2}(x_0, p_0) + \tilde{O}(t^{3/2}), \tag{7.9}$$

$$\frac{\partial P}{\partial x_0} = \tilde{O}(t^{1/2}), \quad \frac{\partial P}{\partial p_0} = E + t \frac{\partial^2 H}{\partial x \partial p}(x_0, p_0) + \tilde{O}(t^{3/2}). \tag{7.10}$$

These relations imply that the map $P : p_0 \mapsto X(t, x_0, p_0)$ is a local diffeomorphism and is globally injective. The last statement follows from the formula

$$X(t, x_0, p_0^1) - X(t, x_0, p_0^2) = t(1 + O(t))(p_0^1 - p_0^2),$$

which one gets by the same arguments as in the proof of Theorem 2.1. Moreover, from this formula it follows as well that $x(t, p_0) \to \infty$, as $p_0 \to \infty$ and conversely. From this one deduces that the image of the map $P : p_0 \mapsto X(t, x_0, p_0)$ is simultaneously closed and open and therefore coincides with the whole space, which completes the proof of the Theorem.

Let us define now the two-points function

$$S_W(t, x, x_0) = \inf \int_0^t (L(y, \dot{y})\, d\tau - c(y)\, dW), \tag{7.11}$$

where inf is taken over all continuous piecewise smooth curves $y(\tau)$ such that $y(0) = x_0$, $y(t) = x$, and the Lagrangian L is, as usual, the Legendre transform of the Hamiltonian H with respect to its last argument.

Theorem 7.2. *Under the assumptions of Theorem 7.1*

$$(i) \quad S_W(t, x, x_)) = \int_0^t (p\, dx - H(x,p)\, dt - c(x)\, dW), \tag{7.12}$$

where the integral is taken along the solution $X(\tau), P(\tau)$ of system (7.3) that joins the points x_0 and x in time t (and which exists and is unique due to Theorem 7.1),

$$(ii) \qquad P(t) = \frac{\partial S_W(t,x,x_0)}{\partial x}, \qquad p_0 = -\frac{\partial S_W(t,x,x_0)}{\partial x_0},$$

(iii) S satisfies equation (7.4), as a function of x,

(iv) $S(t,x,x_0)$ is convex in x and x_0.

Proof. The proof can be carried out by rather long and tedious direct differentiations with the use of the Ito formula. But fortunately, we can avoid it by using the following well known fact [SV, Su, WZ]: if we approximate the Wiener trajectories W in some (ordinary) stochastic Stratonovich equation by a sequence of smooth functions

$$W_n(t) = \int_0^t q_n(s)\, ds \tag{7.13}$$

(with some continuous functions q_n), then the solutions of the corresponding classical (deterministic) equations will tend to the solution of the given stochastic equation. For functions (7.13), equation (7.4) as well as system (7.3) become classical and results of the Theorem become well known (see, for instance, [MF1],[KM1]). In Section 1.1 we have presented these result for the case of Hamiltonians which do not depend explicitly on t, but this dependence actually would change nothing in these considerations. By the approximation theorem mentioned above the sequence of corresponding diffeomorphisms P_n of Theorem 7.1 converges to the diffeomorphism P, and moreover, due to the uniform estimates on their derivatives (see (7.9),(7.10)), the convergence of $P_n(t,x_0,p_0)$ to $P(t,x_0,p_0)$ is locally uniform as well as the convergence of the inverse diffeomorphisms $P_n^{-1}(t,x) \to P^{-1}(t,x)$. It implies the convergence of the corresponding solutions S_n to function (2.2) together with their derivatives in x. Again by the approximation arguments we conclude that the limit function satisfies equation (7.4). Let us note also that the convex property of S is due to equations (1.19),(1.20),(7.9),(7.10).

By similar arguments one gets the stochastic analogue of the classical formula to the Cauchy problem for Hamilton-Jacobi equation, namely the following result

Theorem 7.3 [TZ1],[K1]. *Let $S_0(x)$ is a smooth function and for all $t \le t_0$ and $x \in \mathcal{R}^n$ there exists a unique $\xi = \xi(t,x)$ such that $x(t,\xi) = x$ for the solution $x(\tau,\xi), p(\tau,\xi)$ of system (7.3) with initial data $x_0 = \xi, p_0 = (\partial S_0/\partial x)(\xi)$. Then*

$$S(t,x) = S_0(\xi) + \int_0^t (p\, dx - H(x,p)\, dt - c(x)\, dW) \tag{7.14}$$

(where the integral is taken along the trajectory $x(\tau,\xi), p(\tau,\xi)$) is a unique classical solution of the Cauchy problem for equation (1.4) with initial function $S_0(x)$.

Theorems 7.1, 7.2 imply simple sufficient conditions for the assumptions of Theorem 2.3 to be true. The following result is a direct corollary of Theorem 7.2.

Theorem 7.4. *Under the assumptions of Theorem 7.1 let the function $S_0(x)$ is smooth and convex. Then for $t \le t_0$ there exists a unique classical (i.e. everywhere smooth) solution to the Cauchy problem of equation (7.4) with initial function $S_0(x)$ and it is given by equation (7.14) or equivalently by the formula*

$$R_t S_0(x) = S(t,x) = \min_\xi (S_0(\xi) + S_W(t,x,\xi)). \tag{7.15}$$

One can directly apply the method of constructing generalised solution to deterministic Bellman equation from [KM1],[KM2] to the stochastic case, which gives the following result (details in [K1],[K2]):

Theorem 7.5. *For any bounded from below initial function $S_0(x)$ there exists a unique generalised solution of the Cauchy problem for equation (7.3) that is given by formula (7.15) for all $t \ge 0$.*

Let us mention for the conclusion that the results of Propositions 2.4-2.8 can be now obtained by the similar arguments for the stochastic Hamiltonians. Furthermore, most of the results of the previous Section can be obtained also for complex stochastic Hamiltonian system of form (7.3). For example, let us formulate one of the results.

Theorem 7.6. *Consider a complex stochastic Hamilton-Jacobi equation of form (7.4), with H of form (6.1) supposing that G, A, V, c are analytic in the band $\{|Im\, x| \le 2\epsilon\}$ with some $\epsilon > 0$ and such that G, A, V'', c' are uniformly bounded there together with all their derivatives. Then there exist $\delta > 0$, $t_0 > 0$ and a smooth family Γ of solutions of the corresponding Hamiltonian system (7.3) joining uniquely in time $t \le t_0$ any two points x_0, x such that $|x - x_0| \le \delta$, $|Imx| \le \epsilon$, $|Imx_0| \le \epsilon$. Moreover, all trajectories from Γ are saddle-points for the corresponding functional $\int_0^t (L(y,\dot{y})\, d\tau - c(y)\, dW)$ (in the sense of Section 2.6), and the corresponding random two-point function $S_W(t,x,x_0)$ satisfies (almost surely) equation (7.4).*