Hamiltonian Mechanics

1. Introduction:

Named after Hamilton (who also invented quaternions, as a means of representing spatial rotations in "mechanics"), it embodies the ultimate stage of the formalisation of classical mechanics, pursued, very roughly speaking of course, through Kepler, Newton, Maupertuis and Lagrange, among others. Hamiltonian mechanics has two main sources, namely Lagrangian mechanics on the one hand, with its derivation from the principle of least action and the Euler-Lagrange variational equations, and geometrical optics on the other hand, with Fermat's and Huyghens' principles. Nowadays, these have been somehow subsumed under a common mathematical theory, namely (differential) symplectic geometry which takes place on arbitrary manifolds, endowed with a symplectic structure. Here we shall simply introduce the basic objects in a "flat" vector space, emphasizing the immediate connection with mechanics and Newton's fundational law ("$f = ma$"). We will then (in section 3) try to just list with brief comments some of the more specialized topics, applications, etc. Happily enough, there are now – at least – two easily available references ([A] and [G]) which provide different and complementary introductions to the theory. The reader will find that [G] is the more "elementary", containing almost no symplectic or even differential geometry, but many examples treated in detail, historical remarks etc.; by contrast, [A] is more "advanced" and, although starting from the basics, proceeds to current themes of research.

The main character in Hamiltonian mechanics is Hamilton's function $H = H(p, q)$, which, in the physical cases, is nothing but an expression of the energy of the system through the phase variables (p, q). These – multidimensional – quantities describe the positions (q) and the momenta (p) of the parts of the system at hand (see section 2). An important feature of Hamiltonian mechanics (w.r.t. Lagrangian mechanics) is the emphasis placed on the momenta as primary, independent variables, rather than the "velocities" $\dot{q}$ (we use the customary dot to denote time derivation $\frac{d}{dt}$), momenta having often to do with "amplitudes" in physical systems, e.g. the amplitude of a physical pendulum or the semimajor axis of the ellipse describing the trajectory of a planet. Henceforth we shall also restrict ourselves to the discrete (finite dimensional) theory, describing systems of point-like particles. The continuous theory emerged later, giving rise to classical (relativistic as well as non relativistic) field theories and then helping to formalize quantum field theories (see [G] Chapter 12). We also note that the apparatus of Hamiltonian mechanics played a fundamental role in the discovery of quantum mechanics (or "wave mechanics" as it was then called) through the "analogy" between classical mechanics and geometrical optics, which is built into the formal framework. This was already noted by Hamilton himself (as one of the two sources pointed out above) and served as a convincing guide to de Broglie in the discovery of wave-particle dualism. We refer the reader to [G] for some historical information; the optics-mechanics analogy is outlined in

[G] (§10-8) and [A] (§46). Nowadays, as may be apparent from the above and section 3 below, Hamiltonian mechanics and its influence through symplectic geometry have become pervasive both in physics and mathematics, often under guises which are not easily traced back to “mechanics”.

2. Basic concepts:

Let us first write down the fundamental or “canonical” equations (about the use of this adjective, see the curious footnote in [G], p.342). Let $(p, q, t) = (p_1, \ldots, p_n, q_1, \ldots, q_n, t) \in \mathbb{R}^{2n+1}$ denote a generic point of the extended phase space, where “extended” refers to the inclusion of the time variable and we shall henceforth use standard mutidimensional notation; let $H(p, q, t)$ be a function on this space; we shall ignore questions of smoothness, domains etc. although they of course have to be made precise in a detailed treatment. Then the associated equations read:

$$\dot{p} = -\frac{\partial H}{\partial q}; \qquad \dot{q} = \frac{\partial H}{\partial p}. \tag{1}$$

To see the connection with Newton's fundamental law, set $n = 1$ and $H(p, q) = \frac{p^2}{2m} + V(q)$, V being some function of the one dimensional space variable q. Then, the system (1) reduces to the equation $m\ddot{q} = -\frac{\partial V}{\partial q}$, together with the defining relation $p = m\dot{q}$. The latter relation identifies p with the physical momentum (product of the mass with the velocity) whereas the former relation is precisely Newton's law, for a force f derived from the potential V, i.e. with $f = -\frac{\partial V}{\partial q}$, a relation which defines the force f. ¿From the original viewpoint, Hamilton's mechanics is a generalization of Lagrangian mechanics, and given a mechanical system, the Hamiltonian $H(p, q, t)$ is computed in practice from its Lagrangian $\mathcal{L}(\dot{q}, q, t)$, using a Legendre transformation; in many cases of physical interest, namely when the kinetic energy $T(\dot{q}, q)$ is quadratic in the velocities $\dot{q}$ and the forces derive from a potential $V(q, t)$, it turns out that one has simply $\mathcal{L} = T - V$ and $H = T + V$ where p and $\dot{q}$ are linked by the defining relation $p = \partial\mathcal{L}/\partial\dot{q}$ (see [A] §15 and [G] §8.1).

Now in theory, if not in practice, one can forget about Lagrangian mechanics and consider the Hamiltonian $H(p, q, t)$ as the primary object of study, i.e. consider the actions p as independent variables, and not as the Legendre transforms of the velocities $\dot{q}$. Trading the n second order Lagrange equations for the $2n$-dimensional first order system (1) turned out to be an important conceptual advance. Note that in the theory of partial differential equations, the exact parallel step took much longer to be taken, with the comparatively recent introduction and development of microlocal analysis.

Equations (1) can be derived from a variational principle; namely it is easily seen that their solutions extremize the integral $\int pdq - Hdt$ of the Poincaré-Cartan differential one form $\omega^1 = pdq - Hdt$ ($pdq = \sum p_i dq_i$) over paths in the extended phase space, with boundaries which have prescribed values for the variables (q, t). The occurence of the form ω^1 may look artificial at this point, but is in fact connected with the action functional; this is where optics and mechanics meet (and merge into symplectic geometry) and also the source of the Hamilton-Jacobi method. We refer again to [A] and [G] on these subjects. Actually this variational principle is a consequence of the fact that equations (1) define the characteristic lines of ω^1; a further consequence is that if γ_1 and γ_2 are two closed paths which enclose the same tube of trajectories of (1) (i.e. flow any point of γ_1 forward and backward and you will meet γ_2 exactly once, and the same with the role of 1 and 2 inverted) then the integrals $\int \omega^1$ along γ_1 and γ_2 coincide (cf. [A] §44). In particular, take γ_i $(i = 1, 2)$ in planes $t_i = Cst$, so that $dt = 0$ along γ_i. Then we get that:

$$\int_{\gamma_1} pdq = \int_{\gamma_2} pdq. \tag{2}$$

This is an integral version of the fact that the flow of (1) defines symplectic transformations. More precisely, let g^t denote the time t transformation generated by (1), and let $\omega^2 = dp \wedge dq$ $(= \sum dp_i \wedge dq_i)$ be the standard two form. Then (2) implies that $g^{t*}\omega^2 = \omega^2$, i.e. ω^2 is preserved under the flow of (1). Now this implies that the exterior powers of ω^2 are preserved just as well, in particular the n-th power which is nothing but the usual volume element $\omega^{2n} = dp_1 \wedge \ldots \wedge dp_n \wedge dq_1 \ldots dq_n$. Note that this can be proved in one line because it amounts to the observation that the Hamiltonian vector field $(-\partial H/\partial q, \partial H/\partial p)$ is divergence free, hence preserves the volume in phase space. This property alone has innumerable consequences and reflects the fact that when equations (1) describe a physical system, it contains no friction.

Another important property in this direction is energy conservation: taking the total derivative $\frac{d}{dt}H(p, q, t)$ of the Hamiltonian along the flow of H, one immediately finds that $\frac{d}{dt}H(p, q, t) = \frac{\partial H}{\partial t}$. In particular, for autonomous systems, i.e. when $H(p, q, t) = H(p, q)$ is time independent, one finds that the value of the function H is constant along the trajectories. As we have seen above, for a large class of physical systems, H represents the total energy (kinetic energy + potential energy) of the system, and this is another indication that we are dealing with frictionless, reversible systems.

Actually, all the above properties can now be viewed from a purely geometric viewpoint, forgetting about mechanics altogether and retaining only the fact that equations (1) generate symplectic transformations. This is the point of view adopted in Chapter 8 of [A], to which we refer, whereas Chapter 9 of the same book rederives most properties from a more dynamical viewpoint. We only mention the starting point, which consists in noting that Hamilton's equations can be rewritten as $\dot{x} = JdH(x)$, where $x = (p, q)$, $dH = (\partial H/\partial p, \partial H/\partial q)$ is the gradient of H w.r.t. x and

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

(0 and 1 stand for the n dimensional null and identity matrices respectively) defines the standard symplectic structure in the $2n$ dimensional space (p, q). That is the group $Sp_n(\mathbb{R})$ of linear symplectic transformations is defined as those $A \in GL_{2n}(\mathbb{R})$ satisfying ${}^tAJA = J$ where ${}^tA$ denotes the transpose of A.

Starting from there, one defines the group of –not necessarily linear– symplectic transformations as those which preserve the form ω^2 and builds the theory along these lines. It is then easy to extend the framework to symplectic *manifolds*, which are $2n$ dimensional manifolds endowed with a symplectic form, i.e. a nondegenerate 2-form. An important case is provided by the cotangent bundle T^*M of any n dimensional manifold M; if one thinks of the variables q as describing the "configuration space" M and the p's as describing the fibers of the bundle, there is a natural one form $\omega^1 = pdq$ and its derivative $\omega^2 = dp \wedge dq$ provides the symplectic structure. In this way, one theoretically includes Lagrangian mechanics into the framework, as the symplectic manifolds which appear as cotangent bundle of a "configuration space".

At this point, one could forcefully argue that we have not "solved" any "real" problem, nor made it clear what the effective power of this apparatus might be, with respect to "just writing the equations of motion" i.e. $f = ma$, or a fortiori Lagrange's equations and trying to make the most out of them .... Apart from its aesthetic appeal, which, as should be apparent already, has had a tremendous unifying and suggestive power, the usefulness of the Hamiltonian framework may perhaps be found in two very different directions. The first is the development of a powerful perturbation theory, and this a mainly "analytic" side (see section 3 for a few more words on this). The second is the development of symplectic geometry, which can be considered just as fundamental and pervasive as euclidean geometry. The facts are that it took much longer to realize its ubiquity and that the history is still very young (symplectic *topology* is now barely ten years old – it was born around 1985), yet already very rich. We refer to the dynamical systems and geometry volumes of the Russian Math. Encyclopedia for first-hand accounts.

We close this section however with a few words which point to the continuation of the story and provide some keywords to look for. The Hamiltonian framework makes it easy to devise –and effectively write down– transformations which preserve the form of the equations of motions (i.e. (1)). These are *canonical transformations* which preserve the symplectic structure (see [G] Chapter 9 and more briefly [A] §47). One of the goals of these transformations is to let integrals of the motion appear, i.e. functions $F(p, q)$ which are constant along the trajectories. Assuming that the system is autonomous, one computes from (1) that

$$\frac{d}{dt}F(p, q) = \frac{\partial F}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial H}{\partial q}.$$

The quantity on the r.h.s. is the *Poisson bracket* of F with H, denoted $\{F, H\}$ and is obviously skewsymmetric in its arguments. It makes the space of Hamiltonians into an –infinite dimensional– Lie algebra (checking the validity of the Jacobi identity is easy). We can restate this by saying that F is an integral of the motion if its bracket with H vanishes; note that the skewsymmetry makes it obvious that H itself satisfies this condition. An important, albeit trivial observation, is then that if a variable, say q_i, does not appear in the Hamiltonian $H(p, q)$, the conjugate variable (i.e. p_i) is an integral of the motion. Moreover, if *two* such variables q_i and q_j occur, they are "in involution", i.e. their Poisson bracket $\{q_i, q_j\}$ –obviously– vanishes. One goes on to prove (see [A] §49) that for a large class of systems, integrating the system via quadratures is the same as finding a set of n integrals in involution, i.e. such that all mutual Poisson brackets vanish. One can then express the original Hamiltonian as a fuction of these integrals alone. In other words there exists a canonical transformation $(p, q) \rightarrow (p', q')$ such that in the new variables (called action-angle variables), the Hamiltonian is a function of the p' variables alone, i.e. $H(p, q) = H'(p')$. The solution of the problem in the action-angle variables is then obvious (solve (1) with H independent of the (q, t)). Systems which are amenable to such a procedure are called "integrable". In particular, all one dimensional autonomous systems are integrable, since the Hamiltonian itself provides the only necessary integral of the motion.

3. Subfields, developments, ramifications:

Certainly the most favourable case in mathematics occurs when the problem at hand can be solved in "closed form" (what this means depends on the field); and as a rule, this rarely happens! Here it means of course "integrating" the equations or, say, reducing the solution to quadratures, i.e. having to deal with an integrable system, as outlined above. These systems are few and precious; they are also, so to speak, "far between" among all systems. More precisely, the first such system to be discovered, and still one of the most important, was the two-body problem: two isolated point-like massive bodies subjected to mutual gravitational attraction only. This was "solved" by Newton, using in fact more geometry than the "differential equations" he had just discovered, and represented a landmark in the history of science (for the modern treatment, see [A] §8 as well as the more detailed Chapter 3 of [G]). Apart from one dimensional autonomous systems, which are all integrable as noted above, because of energy conservation, very few "nontrivial" integrable sytems were known until recently, when the appearance of soliton theory and related techniques made it possible to exhibit many more. The search for and study of integrable systems has now become a whole subfield of classical mechanics, often using mostly algebraic techniques, especially drawn from the theory of Lie algebras. The systems thus exhibited have sometimes very little to do with mechanics and are used for instance in the representation theory of Lie algebras and Lie groups.

Whatever the wealth of integrable systems that have been produced, they remain a very thin set (of infinite codimension, meager in the categorical sense etc.) in the "space of Hamiltonians", for any sensible topology. This point of view was particularly emphasized by Poincaré, whose investigations led to the development of modern topology. Perturbation theory, which began as far back as Lagrange and Laplace, is interested in exploring the neighbourhoods of these integrable systems, as represented for instance by simple forms of the *three* body problem, e.g. when one of the bodies is "very small" w.r.t. the other two. Before however studying a given system close to an integrable one, it is sensible to make sure that the systems in the vicinity are not themselves integrable; this is far from being clear a priori and amounts to an impossibility statement akin to the "generic" insolvability statement of Galois theory. In practical terms, it has to do with the generic *divergence* of the perturbation series which were long ago devised by astronomers. Poincaré was the first to recognize the importance of the problem and to prove (roughly sixty years after Galois) the generic insolvability (nonintegrability) of the three-body problem under some mild conditions. This prompted the relatively recent development of yet another branch of the subject, which studies formal criteria for *non*integrability.

Returning to perturbation theory, the problem was certainly on the mind of Newton, at least in connection with the three-body problem, but the subject mainly grew out of astronomy during the late 18-th and the 19-th century, culminating in Poincaré's masterpiece *Les méthodes nouvelles de la mécanique céleste* (available in English, Dover Publ.). Formally, starting from an integrable system, one assumes that it has been "integrated", i.e. that one has found a set of variables (p, q), called action-angle variables, such that he Hamiltonian h depends on the actions only ($h = h(p)$); in practice, the integration may of course turn out to be quite hairy and cumbersome. One then looks at a nearby system, i.e. one studies systems with Hamiltonians of the form $H(p, q) = h(p) + \epsilon f(p, q)$ where ϵ is a small parameter (introductions to perturbation theory may be found in [A], Chapter 10 and [G], Chapter 11). This was termed by Poincaré "the fundamental problem of dynamics"; why? Well, perhaps because the theory "in the large", just does not exist; that is, given a Hamiltonian $H(p, q)$ picked at random (generic in a suitable topology), one cannot say very much about the flow of the associated equations of motion. There are famous case studies, such as again the general three-body problem (which is far from being completely "understood"), and there are some favourable classes of systems which have been singled out. We mention in particular hyperbolic systems, which have been the subject of a long and remarkable string of investigations (Hadamard, Hedlund, Morse, Anosov, Sinaï, Bowen, Ruelle, and others). The prototype of these is geodesic flow on a manifold with negative sectional curvatures. The global situation there is fairly well understood, but one should note that the techniques are not really Hamiltonian in nature (actually hyperbolic systems exist and are studied in a wider setting) and it does not include systems of "physical" origin, i.e. with a Hamiltonian of type "kinetic energy + potential energy", which are just never –or almost never– hyperbolic. Again nature has been subtle in picking systems of mixed type, which are not to-date amenable to any general theory. We also mention the recent renewal of variational calculus (or methods) which has led to some general assertions on the existence of trajectories of certain types (periodic and homoclinic in particular) for fairly large classes of systems. A nice condition which is well-suited for the use of variational methods is convexity, for instance the strict convexity of the energy surface $H(p, q) = E$ one is interested in. Note that hyperbolicity and convexity are of course mutually exclusive and Poincaré had already perceived that the geodesic flow on a convex surface (positive curvature) is actually *more* intricate than on a hyperbolic surface (negative curvature).

The next step –or say another direction– is to step back from dynamical systems altogether and view the study of Hamiltonian systems as a chapter of symplectic geometry. The reader will find an introduction to this more abstract field in [A] (Chapter 8) and it has undergone very important developments recently, together with the newborn symplectic topology. At this point, many statements about general (in particular non perturbative) Hamiltonian systems become purely geometric and lose their dynamical flavor, although dynamical methods (differential equations), of course still play an important role, and will probably continue to do so.

4. References:

[A] V.I. Arnold, *Mathematical Methods in Classical Mechanics*

[G] H. Goldstein, *Classical Mechanics*