A Hamiltonian Formulation of General Relativity

Francis Tong

March 30, 2006

Introduction

The Hamiltonian formulation is a very general idea that has been applied to many different areas of physics. In this report, the Hamiltonian approach will first briefly be discussed in the context of classical theory. The Hamiltonian approach will then be applied to general relativity to obtain what is called the ADM (Arnowitt-Deser-Misner) formulation of general relativity. Finally, some applications of the ADM formulation, such as numerical relativity will be discussed.

Hamiltonian Formulation of a Classical Theory

For classical systems, the standard Hamiltonian approach is as follows:

1. A Lagrangian density $\mathscr{L}$ is first determined by examining the system. In many physical systems, this is just the kinetic energy minus the potential energy of the system. The $\mathscr{L}$ is expressed in terms of generalized coordinates q_i and velocities $\dot{q}_i$.

2. Once $\mathscr{L}$ is determined, the canonical momenta can be defined by

$$p_i = \frac{\partial \mathscr{L}}{\partial \dot{q}_i} \tag{1}$$

The resulting system of equations can usually be solved to obtain $\dot{q}_i$ as functions of p_i.

3. The Hamiltonian density $\mathscr{H}$ is then defined by $\mathscr{L}$ as:

$$\mathscr{H} = \sum_i p_i \dot{q}_i - \mathscr{L}(q_i, \dot{q}_i) \tag{2}$$

where $\dot{q}_i$ are expressed as functions of p_i. $\mathscr{H}$ is thus expressed in terms of generalized coordinates q_i and canonical momenta p_i. The Hamiltonian density completely contains the dynamics of the system.

4. To extract the equations of motion, a pair of equations called Hamilton's equations are used. Hamilton's equations are:

$$\begin{aligned} \dot{q}_i &= \frac{\partial \mathscr{H}}{\partial p_i} \\ \dot{p}_i &= -\frac{\partial \mathscr{H}}{\partial q_i} \end{aligned} \tag{3}$$

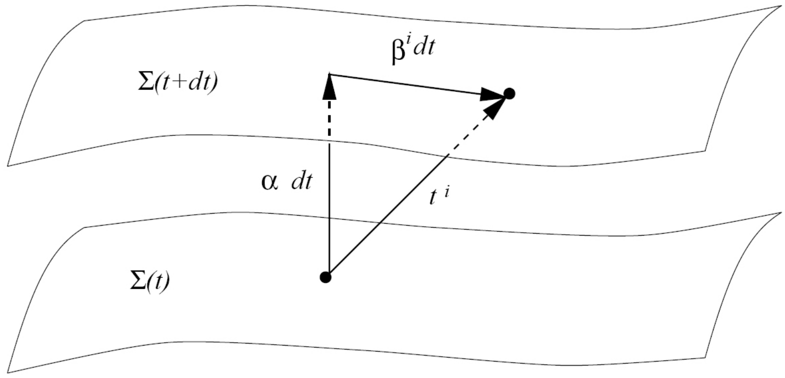

Figure 1: ADM 3+1 foliation of spacetime [3]

For time-dependent Hamiltonians, we get an additional equation:

$$\frac{\partial \mathscr{H}}{\partial t} = -\frac{\partial \mathscr{L}}{\partial t} \tag{4}$$

For systems where the Hamiltonian is independent of time, the Hamiltonian is the total energy of the system.

Hamiltonian Formulation of General Relativity

The formulation discussed here is called the ADM (Arnowitt-Deser-Misner) formulation and was first proposed in 1962. In some literature it is also referred to as the Cauchy or 3+1 formulation, the reasons which will soon become obvious. It has found much success in the area of numerical relativity and serves as a standard approach to solving problems there.

Foliation of Spacetime

Foliation of spacetime is the breaking of the spacetime manifold into a one-parameter family of three dimensional spacelike hypersurfaces parameterized by a time function t. (see figures 1 and 2) The hypersurfaces have timelike normal vectors and spacelike tangent vectors. We can characterize this foliated spacetime as follows: let n^a be a unit normal vector field to the hypersurface Σ_t and let t^a be a vector field on the spacetime manifold. We

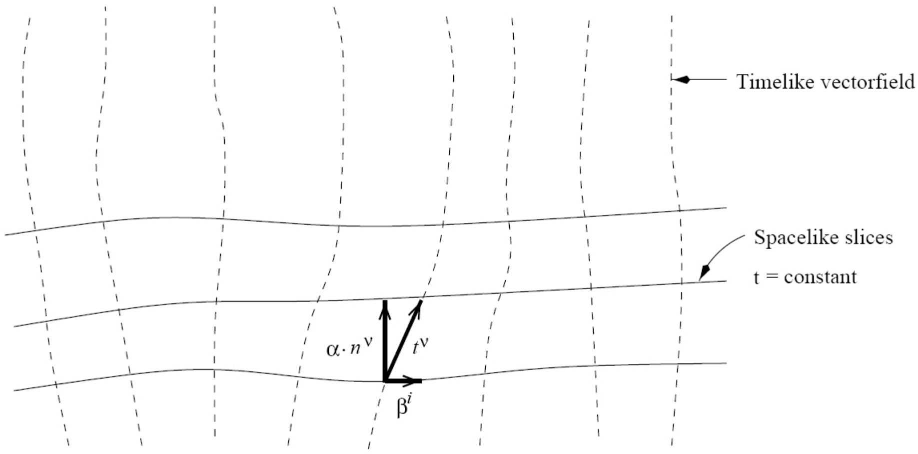

Figure 2: ADM 3+1 foliation of spacetime, a different perspective [4]

then interpret t^a as being the "flow of time" through spacetime and h_{ab} as an induced spatial metric on every surface Σ_t. The spatial metric is related to the spacetime metric by $h_{ab} = g_{ab} + n_a n_b$.

It will turn out to be convenient to decompose t^a into its normal and tangential parts with respect to the surfaces Σ_t. We can do this if we let t^a be defined so that $t^a \nabla_a = 1$. We define a lapse function α and a shift vector β as follows:

$$\alpha = -g_{ab} t^a n^b \tag{5}$$

$$\beta^a = h^a{}_b t^b \tag{6}$$

The lapse function measures the rate of flow of proper time τ with respect to coordinate time t as one moves normally to Σ_t along n^a. The shift vector measures how much the local spatial coordinate system shifts tangential to Σ_1 when moving from Σ_1 to Σ_2 along n^a.

The lapse vector is αn^a, so we have

$$\alpha n^a + \beta^a = t^a$$

$$n^a = \frac{t^a - \beta^a}{\alpha} \tag{7}$$

The lapse function and shift vector are not dynamical because they describe how coordinates move in time from one hypersurface to the next. From the definition of the spatial metric we can write g_{ab} (and thus g^{ab}) in terms of three terms h_{ab}, α, and β^a.

$$g^{ab} = h^{ab} - n^a n^b = h^{ab} - \alpha^{-2}(t^a - \beta^a)(t^b - \beta^b) \tag{8}$$

This shows that choosing h_{ab}, α, and β^a as our field variables is equivalent to using g^{ab}.

Causal Structure of Spacetime

In the ADM formulation, the spacelike hypersurfaces that result from foliation are actually Cauchy surfaces. To define the Cauchy surface, we need to first develop the following formal properties of the causal structure of spacetime.

A vector field in spacetime is causal if it is timelike or null. A curve is causal if its tangent vector is everywhere causal. The *chronological future* $I^+(p)$ of an event p on spacetime M is defined as the set of events that can be reached by a future directed timelike curve from p. For a set of events, the *chronological future* is defined as $I^+(S) = \underset{p \in S}{\cup} I^+(p)$. In other words, $D(S)$ is the set of all future events that can be reached from every event contained in the set S). Analogous definitions exist for the *chronological pasts* $I^-(p)$ and $I^-(S)$.

An *achronal set* $S \subset M$ is defined as having the property that there does not exist events $p, q \in S$ such that $q \in I^+(p)$, i.e. $I^+(S) \cap S = \emptyset$. The *domain of dependence* $D(S)$ of S is defined as $D(S) = D^+(S) \cup D^-(S)$. That is, the domain of dependence is the complete set of events whose information is known because information is known of the S.

For a S that is closed and achronal, we can define the *edge* of S to be the set of events $p \in S$ such that every open neighborhood of p contains a point $q \in I^+(p)$, a point $r \in I^-(p)$, and a timelike curve from r to q that does not intersect S. Finally, a closed achronal set is known as a *slice* .

A *Cauchy surface* is thus defined as a closed achronal set Σ for which $D(\Sigma) = M$. It follows that Cauchy surfaces are edgeless (edge(Σ)=$\emptyset$) and are thus slices. It is helpful to think of Σ_t as being the entire universe at an "instant in time" t. Due to the achronal nature of Cauchy surfaces, it is possible for inextendible causal curves to intersect a given slice only once. By inextendible we mean a curve that has no endpoints. It follows that closed timelike curves are impossible and so time travel is not possible in this formulation of spacetime. Finally, a spacetime that contains a Cauchy surface as a submanifold is termed *globally hyperbolic*.

An Initial Value Formulation

An initial value formulation of general relativity theory is useful because the behaviour of the universe for all time can be determined when given some initial data about the system. This gives rise to a predictive (and retrodictive) theory of general relativity. Having developed the tools to describe the causal structure of spacetime, we now see the reason for using Cauchy surfaces in our 3+1 foliation of spacetime. The use of a Cauchy surfaces as our spatial slices gives rise to an initial value formulation of general relativity.

As we move from one hypersurface to the next along the time flow of t^a, the components of h_{ab} change on each successive hypersurface in accordance with Einstein's field equations. The initial data we need is analogous to the initial data in classical mechanics of initial position and velocity. In our ADM formulation, we require the spatial metric h_{ab} and its time derivative $\dot{h}_{ab}$ as initial data.

The Lagrangian Density

In a way analogous to classical formulations, we obtain the Hamiltonian density $\mathscr{H}$ of our formulation from the Lagrangian density $\mathscr{L}$. The Lagrangian density for vacuum space is

$$\mathscr{L} = \sqrt{-g}R \tag{9}$$

where R is the scalar curvature. However, we require $\mathscr{L}$ in terms of variables that describe our hypersurface Σ_t. To do this, we need to make use of four facts:

Remark 1 We see that the Lagrangian, as it stands, depends on the determinant g of the matrix g_{ab}. We would like to replace it with an expression of terms that describe the hypersurface Σ_t. This can be done by using the following relation:

$$\sqrt{-g} = \alpha\sqrt{h} \tag{10}$$

Remark 2 From the definition of the Einstein tensor G_{ab}, we have R as:

$$
\begin{aligned}
G_{ab} &= R_{ab} - \frac{1}{2}Rg_{ab} \\
-Rg_{ab}n^a n^b &= 2(G_{ab}n^a n^b - R_{ab}n^a n^b) \\
-Rn_b n^b &= 2(G_{ab}n^a n^b - R_{ab}n^a n^b)
\end{aligned}
$$

$$
R = 2(G_{ab}n^a n^b - R_{ab}n^a n^b) \tag{11}
$$

Remark 3 From the Gauss-Codacci equation, which relates the spatial curvature ${}^{(3)}R$ to the spacetime curvature R, we have the following constraint relationship:

$$
G_{ab}n^a n^b = \frac{1}{2}\left[{}^{(3)}R - K_{ab}K^{ab} + K^2\right] \tag{12}
$$

where K_{ab} is the extrinsic curvature of Σ_t and K it's trace. We've introduced K_{ab} as an intermediate variable because $\mathscr{L}$, as we will see, takes a simpler form if expressed in terms of K_{ab}. It represents $\dot{h}_{ab}$ in the Lagrangian density because R depends on $\dot{h}_{ab}$ only through K_{ab}.

Remark 4 From the definition of the Ricci tensor R_{ab}, we have:

$$
\begin{aligned}
R_{ab} &= R_{acb}{}^{c} \\
R_{ab}n^a n^b &= R_{acb}{}^{c}n^b n^a \\
&= -(\nabla_a\nabla_c - \nabla_c\nabla_a)n^c n^a \\
&= -n^a(\nabla_a\nabla_c - \nabla_c\nabla_a)n^c \\
&= (\nabla_a n^a)(\nabla_c n^c) - \nabla_a(n^a\nabla_c n^c) - (\nabla_c n^a)(\nabla_a n^c) + \nabla_c(n^a\nabla_a n^c) \\
&= K^2 - K_{ac}K^{ac} - \nabla_a(n^a\nabla_c n^c) + \nabla_c(n^a\nabla_a n^c)
\end{aligned}
$$

The last two terms of the last equation are divergences and can be neglected. We therefore get the equation

$$
R_{ab}n^a n^b = K^2 - K_{ac}K^{ac} \tag{13}
$$

Using equations (10) to (13), we can now express the Lagrangian density in terms of the variables of the hypersurface:

$$\begin{aligned}
\mathscr{L} &= \sqrt{-g}R \\
&= \alpha\sqrt{h}R \\
&= 2\alpha\sqrt{h}(G_{ab}n^a n^b - R_{ab}n^a n^a) \\
&= 2\alpha\sqrt{h}(\frac{1}{2}\left[{}^{(3)}R - K_{ab}K^{ab} + K^2\right] - K^2 - K_{ab}K^{ab})
\end{aligned} \tag{14}$$

$$\mathscr{L} = \alpha\sqrt{h}({}^{(3)}R + K_{ab}K^{ab} - K^2)$$

The Hamiltonian Density

Recall in classical theory we obtained the Hamiltonian density from the Lagrangian density by the Legendre transformation in equation (2):

$$\mathscr{H} = \sum_i p_i \dot{q}_i - \mathscr{L}(q_i, \dot{q}_i)$$

For general relativity, we choose the spatial metric h_{ab} to be our set of "generalized coordinate" q_i. The canonical momenta p_i is replaced with p^{ab} to conform with our new notation. We therefore have

$$\mathscr{H} = p^{ab}\dot{h}_{ab} - \mathscr{L}(q_i, \dot{q}_i) \tag{15}$$

Our goal is to perform a variation of $\mathscr{H}$ with respect to α and β_a. We will therefore require $\mathscr{H}$ and hence p^{ab} and $\dot{h}_{ab}$ in terms of α, β_a, and h_{ab}. By definition of canonical momentum,

$$\begin{aligned}
p^{ab} &= \frac{\partial \mathscr{L}}{\partial \dot{h}_{ab}} \\
&= \sqrt{h}\alpha\left[\frac{\partial^{(3)}R}{\partial \dot{h}_{ab}} + \frac{\partial(K_{ab}K^{ab})}{\partial \dot{h}_{ab}} - \frac{\partial K^2}{\partial \dot{h}_{ab}}\right] \\
&= \sqrt{h}(K^{ab} - h^{ab}K)
\end{aligned} \tag{16}$$

where we had

$$\begin{aligned}
\frac{\partial K_{ab}}{\partial \dot{h}_{ab}} &= \frac{1}{2\alpha} \\
\frac{\partial\, {}^{(3)}R}{\partial \dot{h}_{ab}} &= 0 \\
\frac{\partial K^2}{\partial \dot{h}_{ab}} &= \frac{h^{ab}K}{\alpha}
\end{aligned}$$

We now need K_{ab} in terms of α, β_a, and h_{ab}. The extrinsic curvature of a surface Σ is defined as

$$K_{ab} = \nabla_a n_b \tag{17}$$

where n_b is a field orthogonal to Σ and tangent to timelike geodesics that do not intersect. To relate K_{ab} to the metric, we make use of the following property of Lie derivatives:

$$\begin{aligned}
\mathcal{L}_n g_{ab} &= n^c \nabla_c g_{ab} + g_{cb} \nabla_a v^c + g_{ac} \nabla_b v^c \\
&= \nabla_a n_b + \nabla_b v_a \\
&= 2\nabla_a n_b
\end{aligned} \tag{18}$$

where the second line holds when ∇_a is the natural derivative operator corresponding to the metric g_{ab} and the third line holds because K_{ab} is symmetric. Substituting this into our definition of K_{ab},

$$\begin{aligned}
K_{ab} &= \frac{1}{2} \mathcal{L}_n g_{ab} \\
&= \frac{1}{2} \mathcal{L}_n (h_{ab} - n_a n_b) \\
&= \frac{1}{2} \mathcal{L}_n h_{ab} \\
&= \frac{1}{2} \left[n^c \nabla_c h_{ab} + h_{cb} \nabla_a v^c + h_{ac} \nabla_b v^c \right] \\
&= \frac{1}{2\alpha} \left[\alpha n^c \nabla_c h_{ab} + h_{cb} \nabla_a \alpha v^c + h_{ac} \nabla_b \alpha v^c \right] \\
&= \frac{1}{2\alpha} h_a{}^c h_b{}^d \left[\mathcal{L}_t h_{cd} - \mathcal{L}_\beta h_{cd} \right] \\
&= \frac{1}{2\alpha} h_a{}^c h_b{}^d \left[\dot{h}_{ab} - D_a \beta_b - D_b \beta_a \right]
\end{aligned} \tag{19}$$

where we have used the appropriate versions of equation 18 when needed and the fact that $\dot{h}_{ab}$ is simply the Lie derivative of h^{ab} with respect to t: $\dot{h}_{ab} = \mathcal{L}_t h^{ab}$

We can now use (15) to write an expression for our surface Σ_t:

$$
\begin{aligned}
\mathscr{H} &= p^{ab}\dot{h}_{ab} - \mathscr{L}(q_i, \dot{q}_i) \\
&= -\sqrt{h}\alpha^{(3)}R + \frac{\alpha}{\sqrt{h}}\left[p^{ab}p_{ab} - \frac{1}{2}p^2\right] + 2p^{ab}D_a\beta_b \\
&= \sqrt{h}\left[\alpha\left[-^{(3)}R + h^{-1}p^{ab}p_{ab} - \frac{1}{2}h^{-1}p^2\right] - 2\beta_b\left[D_a(h^{-1/2}p^{ab})\right] + 2D_a(h^{-1/2}\beta_b p^{ab})\right] \\
&= \sqrt{h}\left[\alpha\left[-^{(3)}R + h^{-1}p^{ab}p_{ab} - \frac{1}{2}h^{-1}p^2\right] - 2\beta_b\left[D_a(h^{-1/2}p^{ab})\right]\right]
\end{aligned}
\tag{20}
$$

where p is the trace of p^{ab} and we neglected the boundary term in the last line because we assume a sufficiently large spatial surface so that the boundary effects are negligible.

Constraint and Evolution Equations

To determine the Hamiltonian H, which we need in order get the constraint and evolution equations, we integrate $\mathscr{H}$ over the hypersurface Σ_t using the fixed spatial volume element ${}^{(3)}\mathbf{e}$. (Note: ${}^{(3)}\mathbf{e}\sqrt{h} = \epsilon$ where ϵ is the natural volume element associated with the metric h_{ab}.)

$$H = \int_{\Sigma} \mathscr{H}^{(3)}\mathbf{e} \tag{21}$$

To obtain the constraint equations of the sytem, we perform a variation of H with respect to α and β. From this we get the two constraint equations (22, 23):

$$-^{(3)}R + \frac{p^{ab}p_{ab}}{h} - \frac{p^2}{2h} = 0 \tag{22}$$

This equation constrains the Hamiltonian. It forces the first term of the Hamiltonian density (20) to vanish.

$$D_a(\frac{p^{ab}}{\sqrt{h}}) = 0 \tag{23}$$

This equation constrains the momentum. p^{ab} cannot change with respect to a. This constraint can be removed if one were to choose a superspace as the configuration space.[6] With this configuration space, all possible momenta already satisfy equation (23).

The last step of the formulation is to obtain the evolution equations from which all of spacetime can be determined when given any initial data. To do this, we expand Hamilton's equations in an analogous way to the classical case:

$$\begin{aligned}\mathcal{L}_t h^{ab} = \dot{h}_{ab} = & \frac{\delta H}{\delta p^{ab}} \\ = & \frac{2\alpha}{\sqrt{h}}(p_{ab} - \frac{h_{ab}p}{2}) + D_a\beta_b + D_b\beta_a\end{aligned} \tag{24}$$

$$\begin{aligned}\mathcal{L}_t p^{ab} = \dot{p}^{ab} = & -\frac{\delta H}{\delta h_{ab}} \\ = & -\alpha\sqrt{h}({}^{(3)}R^{ab} - \frac{{}^{(3)}Rh^{ab}}{2}) \\ & + \frac{\alpha h^{ab}}{2\sqrt{h}}(p_{cd}p^{cd} - \frac{p^2}{2}) \\ & - \frac{2\alpha}{\sqrt{h}}(p^{ac}p_c{}^b - \frac{pp^{ab}}{2}) \\ & + \sqrt{h}(D^aD^b\alpha - h^{ab}D^cD_c\alpha) \\ & + \sqrt{h}D_c(\frac{\beta^c p^{ab}}{\sqrt{h}}) - 2p^{c(a}D_c\beta^{b)}\end{aligned} \tag{25}$$

The two constraint equations (22, 23) and two evolution equations (24, 25) together constitute a constrained Hamiltonian formulation of general relativity for vacuum space. Given initial conditions that satisfy the constraint equations, the universe at any point in spacetime can, in principle, be determined by evolving the system using the evolution equations.

Applications

The ADM Hamiltonian formulation of general relativity was, for some time, the most popular model used in the field of numerical relativity. The procedure for performing computational experiments using the ADM theory is roughly as follows:

1. Make an assumption about the metric of the system.
2. Solve the constraint equations.
3. Specify the slicing conditions of spacetime. (eg values of α and K)

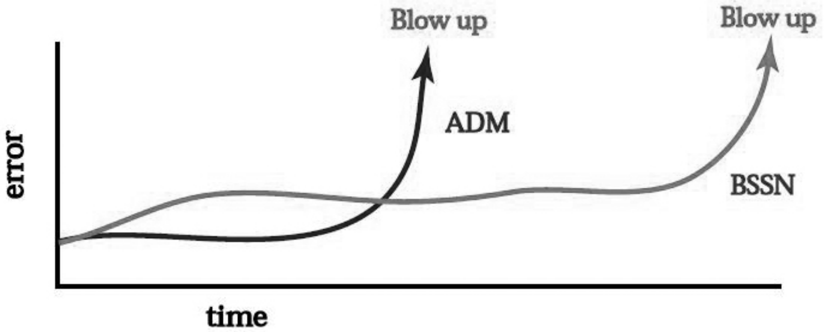

Figure 3: Stability of ADM vs BSSN formulations in numerical applications [5]

4. Evolve the variables in time using the evolution equations.
5. Extract the values of physical quantities at the end of the evolution.

Although the equations we have found are mathematically equivalent to the vacuum Einstein's equations, in practice they are not. For instance, in computer simulations, the numerical stability of two formulations can be quite different. For instance, the ADM approach, for a large number of scenarios, is capable of a shorter lived evolution than another formulation called BSSN (a conformal traceless formulation). [5] The two are compared in figure 3.

The ADM Hamiltonian model may no longer be the best one available, but it continues to be relevant in modern research. For instance, one of the new models in current use is actually a modified formulation of the ADM model. This model, called the adjusted ADM model, is basically the ADM model except it uses the extrinsic curvature K_{ab} as a dynamical variable instead of $\dot{h}_{ab}$.

The ADM Hamiltonian formulation has also been useful as an approach to the canonical quantization of general relativity. Other Hamiltonian formulations exist and continue to be studied and developed in an attempt to reach a theory of quantum gravity. The ADM approach acted as one of many springboards into this area of research and is still referred to in today's literature on the subject. The ADM formulation has stayed relevant for a long time since its creation in the 1960s. It is reasonable to expect that it will continue to be useful in the foreseeable future.

References

[1] R.M. Wald, *General Relativity*, The University of Chicago Press, Chicago, 1984.

[2] Thornton, Marion, *Classical Dynamics*, Thomson-Brooks/Cole, 2004.

[3] R.R. Correll, *Numerical Evolution of Binary Blackhole Spacetimes*, University of Texas at Austin

[4] B. Szilagyi, *Cauchy-Characteristic Matching In General Relativity*, University of Pittsburgh

[5] H. Shinkai, *Introduction to Numerical Relativity*, Institute of Physical and Chemical Research, Japan

[6] Wheeler, *Superspace and the Nature of Quantum Geometrodynamics*, Battelle Recontres, Benjamin, New York