# From probabilities to Hamilton's principle



2009 J. Phys. A: Math. Theor. 42 315202

(http://iopscience.iop.org/1751-8121/42/31/315202)

J. Phys. A: Math. Theor. **42** (2009) 315202 (6pp) doi:10.1088/1751-8113/42/31/315202

# From probabilities to Hamilton's principle

**V Kapsa and L Skála**[1]

Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Prague 2, Czech Republic

E-mail: skala@karlov.mff.cuni.cz

Received 17 March 2009, in final form 11 June 2009
Published 13 July 2009
Online at stacks.iop.org/JPhysA/42/315202

**Abstract**
It is shown that Hamilton's principle of classical mechanics can be derived from the relativistic invariance of the generalized spacetime Fisher information corresponding to the statistical description of results of measurement of the coordinates and the motion in space by means of the probability density and probability density current.

PACS numbers: 45.20.Jj, 01.55.+b, 02.50.Cw

## 1. Introduction

Hamilton's principle of classical mechanics [1–5] (sometimes called the principle of stationary or least action) is a very powerful integral principle making it possible to derive equations of motion of classical mechanics. Till now, this principle has been understood as one of the basic postulates of classical mechanics [2–5].

The aim of this paper is different. We want to show that Hamilton's principle can be derived from two general requirements. The first one—statistical description of the measurement of the coordinates and the motion in space by means of the probability density and probability density current—is related to the statistical character of results of measurements. The second one is the relativistic invariance of the corresponding generalized spacetime Fisher information, important characteristics of the probability distributions known from mathematical statistics.

The importance of the Fisher information [6] as the starting point to obtaining the most important equations of motion of physics was realized in [7–9]. However, Hamilton's principle and its derivation from the spacetime Fisher information, the role of the probability density and probability density current and their representation in a way convenient for transition to classical mechanics were not discussed in [7–9]. For this reason, the main ideas making it possible to derive Hamilton's principle from the generalized spacetime Fisher information are discussed in this paper.

[1] Also at Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2 L 3G1, Canada.

## 2. Probability density

We begin our discussion with measurement of the coordinate $x$. For the sake of simplicity, we consider the one-dimensional case only.

Results of repeated measurements of the coordinate $x$ can be characterized by the mean values

$$\langle x\rangle = \int x\rho(x,t)\,\mathrm{d}x, \tag{1}$$

$$\langle x^2\rangle = \int x^2\rho(x,t)\,\mathrm{d}x, \tag{2}$$

where the integration is carried out over the whole space, $\rho(x,t) \geqslant 0$ is a normalized probability density

$$\int \rho\,\mathrm{d}x = 1 \tag{3}$$

and

$$\lim_{x\to\pm\infty} x^n\rho = 0, \qquad n = 0, 1, 2. \tag{4}$$

Normalization condition (3) and equation (4) are assumed to be valid at all times $t$.

## 3. Uncertainty relation and Fisher information

Now, we perform integration by parts with respect to the variable $x$ in equation (3) and get [6, 10–15]

$$x\rho|_{x=-\infty}^{\infty} - \int x\frac{\partial\rho}{\partial x}\,\mathrm{d}x = 1. \tag{5}$$

Using equation (4) the first term in this equation equals zero and we get

$$\int x\frac{\partial\rho}{\partial x}\,\mathrm{d}x = -1. \tag{6}$$

This simple result has interesting consequences. Putting

$$u = x\sqrt{\rho} \tag{7}$$

and

$$v = \frac{1}{\sqrt{\rho}}\frac{\partial\rho}{\partial x}, \tag{8}$$

the Schwarz inequality

$$(u,u)(v,v) \geqslant |(u,v)|^2 \tag{9}$$

for the inner product $(u,v) = \int u^* v\,\mathrm{d}x$ yields the 'uncertainty' relation [6–15]

$$\langle x^2\rangle I \geqslant 1. \tag{10}$$

Here,

$$I = \int \frac{1}{\rho}\left(\frac{\partial\rho}{\partial x}\right)^2 \mathrm{d}x \tag{11}$$

is the so-called Fisher information, important characteristics of the probability distributions known from mathematical statistics and information theory [6–9, 16].

## 4. Probability density current

For physical systems, we must give not only the probability density $\rho(x, t)$ but also some quantity describing the motion of a given system in space. It is also needed for describing the transition between two mutually moving inertial systems.

By analogy with continuum mechanics, it is possible to introduce the probability density current $j$ related to the 'velocity' $v$ [12–14]

$$j = \rho v. \tag{12}$$

The quantities $\rho$ and $v$ can be in general expressed in terms of two real functions $s_1 = s_1(x, t)$ and $s_2 = s_2(x, t)$ as follows:

$$\rho = \mathrm{e}^{-2s_2/\eta}, \tag{13}$$

$$v = \frac{1}{m}\frac{\partial s_1}{\partial x}. \tag{14}$$

Here, $m$ is the mass of the system and $\eta > 0$ is a constant given by the normalization condition (3). The last equation has the form analogous to that between the velocity $v$ and momentum $p$ in classical mechanics $v = p/m$, where $p = (\partial S/\partial x)$ and $S$ is the classical Hamilton action. It is seen that instead of $\rho$ and $j$, the state of the system can be described by the functions $s_1$ and $s_2$ (see also [10–15]).

We note that a similar approach to writing $\rho$ and $j$ in terms of functions $s_1$ and $s_2$ was successfully used in quantum mechanics where the wavefunction $\psi$ can be written as [10–15] (see also [17–19])

$$\psi = \mathrm{e}^{(\mathrm{i}S_1 - S_2)/\hbar}, \tag{15}$$

where functions $S_1$ and $S_2$ are real functions having similar meaning as $s_1$ and $s_2$. The corresponding quantum-mechanical probability density and probability density current can then be expressed in the form

$$\rho = \mathrm{e}^{-2S_2/\hbar}, \tag{16}$$

$$j = \frac{\hbar}{2m\mathrm{i}}[\psi^* \nabla\psi - (\nabla\psi)^*\psi] = \frac{\rho}{m}\frac{\partial S_1}{\partial x} \tag{17}$$

analogous to equations (12)–(14). In the limit of classical mechanics, the probability density (16) can be approximated by the Dirac function $\delta(x - x_{\mathrm{cl}})$, where $x_{\mathrm{cl}} = x_{\mathrm{cl}}(t)$ is the classical trajectory and $S_1$ becomes the classical action $S$ [12, 15]. Analogous transition will be performed in section 6.

## 5. Generalized uncertainty relation

To derive a generalized uncertainty relation for $s_1$ and $s_2$ we first substitute equation (13) into equation (11). The resulting formula for $I$ has the form [12, 15]

$$I = \frac{4}{\eta^2}\int \left(\frac{\partial s_2}{\partial x}\right)^2 \mathrm{e}^{-2s_2/\eta}\,\mathrm{d}x. \tag{18}$$

Now we note that

$$I \leqslant I', \tag{19}$$

where

$$I' = \frac{4}{\eta^2} \int \left[ \left( \frac{\partial s_1}{\partial x} \right)^2 + \left( \frac{\partial s_2}{\partial x} \right)^2 \right] \mathrm{e}^{-2s_2/\eta} \, \mathrm{d}x \tag{20}$$

is a generalized Fisher information [12]. Physical importance of $I'$ is given by the fact that it takes into account not only the form of the probability distribution given by $\rho$ (or $s_2$) but also an analogous distribution given by $j$ (or $s_1$ and $s_2$). It can be shown that the generalized Fisher information $I'$ is related to the kinetic energy in quantum mechanics [12]. It is evident uncertainty relation for $I'$ following from equations (10) and (19)

$$\langle x^2 \rangle I' \geqslant 1 \tag{21}$$

is closely related to the Heisenberg uncertainty relations [12, 13, 15].

## 6. Hamilton's principle

To derive Hamilton's principle we require the relativistic invariance of the generalized spacetime Fisher information. A similar approach has been successfully used also for the derivation of the Klein–Gordon and Dirac equations [7–10, 12].

The relativistically invariant spacetime generalization of the Fisher information $I'$ has the form

$$J' = J_1' + J_2', \tag{22}$$

where the first term

$$J_1' = \frac{4}{\eta^2} \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \int \left[ \frac{1}{c^2} \left( \frac{\partial s_1}{\partial t} \right)^2 - \left( \frac{\partial s_1}{\partial x} \right)^2 \right] \mathrm{e}^{-2s_2/\eta} \, \mathrm{d}x \, \mathrm{d}t \tag{23}$$

depends on the motion in space described by $s_1$ and the form of the probability distribution $\rho = \exp(-2s_2/\eta)$ and the second term

$$J_2' = \frac{4}{\eta^2} \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \int \left[ \frac{1}{c^2} \left( \frac{\partial s_2}{\partial t} \right)^2 - \left( \frac{\partial s_2}{\partial x} \right)^2 \right] \mathrm{e}^{-2s_2/\eta} \, \mathrm{d}x \, \mathrm{d}t \tag{24}$$

depends on $s_2$ only. Here, $1/(t_2 - t_1)$ is a normalization factor with respect to time and $c$ is the speed of light. Time integration is performed from $t_1$ when the initial conditions were given (first measurement or preparation of the system in the state given by $\rho(x, t_1)$ and $j(x, t_1)$). At times $t \in (t_1, t_2)$ no measurement is performed. At $t_2$, the system interacts with the measuring apparatus again (second measurement).

The condition that $J'$ has the same value in all physically equivalent inertial systems can be written in the variational form

$$\delta J' = 0. \tag{25}$$

Physically relevant functions $s_1 = s_1(x, t)$ and $s_2 = s_2(x, t)$ are the only such functions $s_1$ and $s_2$ that obey this condition.

To get the non-relativistic approximation to $J'$ we write $s_1$ in the form

$$s_1 = m_0 c^2 t + \bar{s_1}, \tag{26}$$

where $m_0$ is the rest mass of the system and $\bar{s_1} = \bar{s_1}(x, t)$ is a real function. It leads to

$$\left( \frac{\partial s_1}{\partial t} \right)^2 = \left( m_0 c^2 + \frac{\partial \bar{s_1}}{\partial t} \right)^2 = m_0^2 c^4 + 2 m_0 c^2 \frac{\partial \bar{s_1}}{\partial t} + \left( \frac{\partial \bar{s_1}}{\partial t} \right)^2. \tag{27}$$

Now, we assume that the term $\partial\bar{s}_1/\partial t$ is small with respect to the rest energy $m_0c^2$ and the last term in equation (27) can be neglected. Then we obtain

$$J_1' = \frac{4m_0^2c^2}{\eta^2} + \frac{4}{\eta^2}\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\!\!\int\left[2m_0\frac{\partial\bar{s}_1}{\partial t} - \left(\frac{\partial\bar{s}_1}{\partial x}\right)^2\right]\mathrm{e}^{-2s_2/\eta}\,\mathrm{d}x\,\mathrm{d}t. \tag{28}$$

Further, we take into consideration that $s_1$ does not appear in equation (6) for $\rho$. It shows that modifying $s_1$ it is possible to introduce additional functions or 'potentials' into the theory that can describe external fields in which the system moves. Therefore, in agreement with physical experience we can introduce an external potential $V(x,t)$ by the transition

$$\frac{\partial\bar{s}_1}{\partial t} \to \frac{\partial\bar{s}_1}{\partial t} + V \tag{29}$$

analogous to similar prescriptions known from classical and quantum mechanics [7–15]. It leads to

$$J_1' = \frac{4m_0^2c^2}{\eta^2} + \frac{8m_0}{\eta^2(t_2-t_1)}\int_{t_1}^{t_2}\!\!\int\left[\frac{\partial\bar{s}_1}{\partial t} + V - \frac{1}{2m_0}\left(\frac{\partial\bar{s}_1}{\partial x}\right)^2\right]\mathrm{e}^{-2s_2/\eta}\,\mathrm{d}x\,\mathrm{d}t. \tag{30}$$

Now, we assume that the probability density $\rho(x,t)$ is close to zero everywhere except for a very narrow region along a classical trajectory $x_{\mathrm{cl}} = x_{\mathrm{cl}}(t)$. In such a case, the probability density $\rho$ can be approximated by the $\delta$-function at the point $x_{\mathrm{cl}}$

$$\rho(x,t) = \mathrm{e}^{-2s_2/\eta} \approx \delta(x - x_{\mathrm{cl}}). \tag{31}$$

This classical limit can also be obtained for $\eta \to 0+$ corresponding to the limit $\hbar \to 0+$ known from transition from quantum to classical mechanics (see e.g. [20, 21]). Then we get

$$J_1' = \frac{4m_0^2c^2}{\eta^2} + \frac{8m_0}{\eta^2(t_2-t_1)}\left\{S(x_{\mathrm{cl}},t)\Big|_{t_1}^{t_2} + \int_{t_1}^{t_2}\left[V(x_{\mathrm{cl}},t) - \frac{1}{2m_0}\left(\frac{\partial S(x_{\mathrm{cl}},t)}{\partial x_{\mathrm{cl}}}\right)^2\right]\mathrm{d}t\right\}, \tag{32}$$

where $S$ denotes the function $\bar{s}_1$ in the classical limit. Analogously, we can calculate

$$J_2' = \frac{4}{\eta^2}\frac{1}{t_2-t_1}\int_{t_1}^{t_2}\left[\frac{1}{c^2}\left(\frac{\partial s_2(x_{\mathrm{cl}},t)}{\partial t}\right)^2 - \left(\frac{\partial s_2(x_{\mathrm{cl}},t)}{\partial x_{\mathrm{cl}}}\right)^2\right]\mathrm{d}t. \tag{33}$$

Taking into consideration that the function $\delta(x - x_{\mathrm{cl}})$ is even with respect to $x_{\mathrm{cl}}$ we see that its spatial derivative at this point must equal zero. Similarly, the form of the function $\delta(x - x_{\mathrm{cl}})$ does not change in time and its time derivative equals zero. Therefore, taking into account the form of derivatives of $\rho$ following from equation (31), we can assume that $J_2'$ does not contribute to $\delta J'$ in the limit of classical mechanics.

Finally, taking into consideration that the variations of $S(x_{\mathrm{cl}},t)$ at the points $t_1$ and $t_2$ (the first and second measurement) equal zero we can write the condition (25) in terms of the integral in equation (32)

$$\delta\int_{t_1}^{t_2}\left[\frac{1}{2m_0}\left(\frac{\partial S(x_{\mathrm{cl}},t)}{\partial x_{\mathrm{cl}}}\right)^2 - V(x_{\mathrm{cl}},t)\right]\mathrm{d}t = 0. \tag{34}$$

Denoting

$$L = \frac{1}{2m_0}\left(\frac{\partial S(x_{\mathrm{cl}},t)}{\partial x_{\mathrm{cl}}}\right)^2 - V(x_{\mathrm{cl}},t) \tag{35}$$

and taking into account that $\partial S/\partial x$ equals momentum $p$ in classical mechanics [2–5, 15] we can write the condition for the trajectory $x_{\mathrm{cl}} = x_{\mathrm{cl}}(t)$ in the form of Hamilton's principle

$$\delta\int_{t_1}^{t_2} L\,\mathrm{d}t = 0, \tag{36}$$

where

$$L = \frac{p^2}{2m_0} - V. \tag{37}$$

## 7. Conclusions

We have seen that Hamilton's principle can be obtained in a few steps. In the first step, results of measurement of the coordinates are described by means of the probability density $\rho$. In the next step, the description of the motion in space is included by means of the probability density current $j$. The last two steps make it possible to derive the 'uncertainty' relation for $\langle x^2 \rangle$ and the generalized spatial Fisher information $I'$. To derive Hamilton's principle, the relativistically invariant generalized spacetime Fisher information $J'$ analogous to $I'$ is defined. Then, the condition that $J'$ has the same value in all physically equivalent inertial systems is written in the variational form. The variational condition followed by the transition to the non-relativistic classical limit leads finally to Hamilton's principle.

In summary, our discussion shows that Hamilton's principle has its origin in the statistical description of results of measurement of the coordinates and the motion in space by means of the probability density $\rho$ and probability density current $j$, respectively, and relativistic invariance of the corresponding generalized spacetime Fisher information $J'$.

## Acknowledgment

This work was supported by the MSMT grant no. 0021620835 of the Czech Republic.

## References

[1] Hamilton W R 1834 *Phil. Trans. R. Soc. Part I* 247–308
Hamilton W R 1835 *Phil. Trans. R. Soc. Part II* 95–144
[2] Goldstein H 1980 *Classical Mechanics* (Reading, MS: Addison-Wesley)
[3] Landau L D and Lifshitz E M 1976 *Mechanics* (Oxford: Pergamon)
[4] Arnold V I 1989 *Mathematical Methods of Classical Mechanics* (Berlin: Springer)
[5] Lanczos C 1970 *The Variational Principles of Mechanics* (New York: Dover)
[6] Fisher R A 1925 *Proc. Camb. Phil. Soc.* **22** 700
[7] Frieden B Roy and Soffer B H 1995 *Phys. Rev.* E **52** 2274
[8] Frieden B Roy 1998 *Physics from Fisher Information* (Cambridge: Cambridge University Press)
[9] Frieden B Roy 2004 *Science from Fisher Information: A Unification* (Cambridge: Cambridge University Press)
[10] Skála L and Kapsa V 2005 *Physica* E **29** 119
[11] Skála L and Kapsa V 2005 *Collect. Czech. Chem. Commun.* **70** 621
[12] Skála L and Kapsa V 2007 *Opt. Spectrosc.* **103** 434
[13] Skála L and Kapsa V 2008 *J. Phys. A: Math. Theor.* **41** 265302
[14] Skála L and Kapsa V 2009 *Int. J. Quantum Chem.* **109** 1626
[15] Skála L and Kapsa V 2007 *Int. Rev. Phys.* **1** 302
[16] Cover T and Thomas J 1991 *Elements of Information Theory* (New York: Wiley)
[17] Bohm D 1952 *Phys. Rev.* **85** 166
[18] Bohm D 1952 *Phys. Rev.* **85** 180
[19] Madelung E 1926 *Z. Phys.* **40** 322
[20] Davydov A S 1976 *Quantum Mechanics* (New York: Pergamon)
[21] Shankar R 1994 *Principles of Quantum Mechanics* (New York: Plenum Press)