

## Principle of maximum caliber and quantum physics

Ignacio J. General

*School of Science and Technology, Universidad Nacional de San Martín, and CONICET,  
25 de Mayo y Francia, San Martín, 1650 Buenos Aires, Argentina*



(Received 15 May 2018; published 11 July 2018)

MaxCal is a variational principle that can be used to infer distributions of paths in the phase space of dynamical systems. It has been successfully applied to different areas of classical physics, in particular statistical mechanics in and out of equilibrium. In this work, guided by the analogy of the formalism of MaxCal with that of the path integral formulation of quantum mechanics, we explore the extension of its applications to the realm of quantum physics, and show how the Lagrangians of both relativistic and nonrelativistic quantum fields can be built from MaxCal, with a suitable set of constraints. Related, the details of the constraints allow us to reinterpret the concept of inertia.

DOI: [10.1103/PhysRevE.98.012110](https://doi.org/10.1103/PhysRevE.98.012110)

### I. INTRODUCTION

The *Principle of Maximum Caliber* (MaxCal), first proposed by Jaynes [1] as a generalization of the principle of *Maximum Entropy* (MaxEnt) [2,3], has become more popular only in the last two decades [4] and is starting to prove itself as a powerful tool in physics. It has been shown that MaxCal contains the more usual descriptions of the dynamics of physical systems: the principle of least action, the Euler–Lagrange equations, and Newton’s second law.

In the next section we present the derivation of the least action principle from MaxCal, and show that the result of MaxCal is more general than just that principle; in fact, MaxCal allows more trajectories than just the one that extremizes the action. This resembles the case of quantum physics, where many trajectories are allowed. But what are those extra trajectories permitted by MaxCal? Making a connection with the path integral formulation of quantum mechanics, we show that the probability of those trajectories is extremely suppressed for classical values of the action, resulting in an agreement of MaxCal and classical physics.

In the following section, and based on the relation found between MaxCal and quantum mechanics, we explore the relevance of the principle to quantum physics. In particular, we show that both relativistic (Klein–Gordon and Dirac) and nonrelativistic (Schrödinger) versions of quantum mechanics can be derived from MaxCal with a suitable choice of constraints.

### II. PRINCIPLE OF LEAST ACTION AS A CONSEQUENCE OF MAXCAL, AND ITS RELATION TO QUANTUM MECHANICS

Maximization of the caliber is a variational principle that allows the inference of probability distributions compatible with a set of given constraints [1,4]. This principle is a straight generalization of the MaxEnt principle [2,3], where microstates—microscopic realizations of the conformation of a system—are replaced by *microtrajectories* between two

points in the phase space—microscopic realizations of the passage of a system from one point to another.

It has already been shown to be a successful tool to derive several relations related not only to equilibrium but also to systems out of equilibrium. Fick’s first and second laws of diffusion, Fourier’s law of heat transfer, Newton’s second law, Brownian motion, Onsager’s reciprocal relationships, Prigogine’s principle of minimum entropy production, are some examples of this [5–8].

Here, following Wang [6] (with a slightly different point of view), we want to show that the principle of least action can be obtained starting from MaxCal by choosing suitable constraints. Let us take a system that moves from points  $a$  to  $b$  in its phase space, subjected to the following constraints:

(1) Each individual path  $i$  between  $a$  and  $b$  is characterized by a well-defined physical property  $A_{ab}(i)$ , and its average over all possible trajectories is  $A_{ab}$ .

(2) The sum of the probabilities for all paths is 1.

Mathematically, we have to maximize the caliber

$$S(a,b) = - \sum_{i=1}^N p_i(ab) \ln p_i(ab), \quad (1)$$

constrained by

$$A_{ab} = \sum_{i=1}^N p_i(ab) A_{ab}(i) \quad \text{and} \quad \sum_{i=1}^N p_i(ab) = 1. \quad (2)$$

Here,  $p_i(ab)$  represents the probability that the system will follow path  $i$  when going from points  $a$  to  $b$  in the phase space, and  $N$  is the number of possible paths connecting those two points. Using the Lagrange multipliers method, we define the auxiliary function

$$S' = - \sum_{i=1}^N p_i(ab) \ln p_i(ab) - \lambda \left( \sum_{i=1}^N p_i(ab) - 1 \right) - \eta \left( \sum_{i=1}^N p_i(ab) A_{ab}(i) - A_{ab} \right) \quad (3)$$

and calculate stationary points with respect to  $p_i$  and the two multipliers  $\lambda$  and  $\eta$ :

$$\frac{\partial S'}{\partial p_i} = 0 \Rightarrow -\ln p_i - 1 - \lambda - \eta A_{ab}(i) = 0, \quad (4)$$

$$\frac{\partial S'}{\partial \lambda} = 0 \Rightarrow \sum_{i=1}^N p_i = 1, \quad (5)$$

$$\frac{\partial S'}{\partial \eta} = 0 \Rightarrow \sum_{i=1}^N p_i A_{ab}(i) = A_{ab}. \quad (6)$$

Combining Eqs. (4) and (5), it is possible to eliminate  $\lambda$  and arrive at an expression for the probability of path  $i$ :

$$p_i(ab) = \frac{1}{Z} e^{-\eta A_{ab}(i)}, \quad (7)$$

where we define the partition function as

$$Z = \sum_{i=1}^N e^{-\eta A_{ab}(i)}. \quad (8)$$

Notice that

$$-\frac{\partial(\ln Z)}{\partial \eta} = A_{ab}. \quad (9)$$

Equation (7) shows the relation between the probability of path  $i$  and its corresponding property  $A_{ab}(i)$ : path  $i$  between  $a$  and  $b$  has maximum probability  $p_i(ab)$  when  $A_{ab}(i)$  is minimum. Notice that if, in the last sentence, we replace “ $A_{ab}(i)$ ” by the word “action,” we are then stating the principle of least action with a caveat: MaxCal [plus constraints in Eqs. (2)] contains the least action principle as its most probable outcome, but it is not equivalent to it since MaxCal allows other trajectories with lower probability. But, since classically there is only one allowed trajectory, we can infer that  $\eta$  must be high, thus suppressing the extra trajectories, and leaving only the most probable one.

To be more specific, consider a simple example, where a system can go from  $a$  to  $b$  via three possible paths, with values of the action of 1, 2, and 3 (in arbitrary units), respectively. The probability of these cases, calculated via Eqs. (7) and (8), are shown in Table I. There we can see that the probability of the paths with higher action decrease extremely fast if  $\eta$  is large. Hence, a classical system could be perfectly described by MaxCal with a large enough value of the multiplier  $\eta$ .

As noted by Davis and Gonzalez [9], there is a strong resemblance between the MaxCal formalism and the path integral formulation of quantum mechanics [10]. More specifically, from the path integral perspective a particle can go from  $a$  to  $b$  via any physical path connecting them, and each path  $i$

contributes an exponential factor of the action to the probability amplitude  $P^{\text{amp}}$  of the process:

$$P_i^{\text{amp}}(ab) \propto e^{\frac{i}{\hbar} A_{ab}(i)}. \quad (10)$$

A comparison with Eq. (7) suggests that  $\eta$  is inversely proportional to Planck’s constant, making its value very large. Hence, when the action is classical (large values of  $A$ ), this would make the probability of nonminimal action trajectories vanish, so that the results of MaxCal agree with those of classical physics.

We have to mention that, although we have compared Eqs. (7) and (10), their meanings are not the same. Equation (7) is the probability of the particle taking path  $i$ , and, thus, the path with minimum action will be the most likely. Also, as we already said, if  $\eta$  is large, other paths will be highly suppressed.

On the other hand, Eq. (10) is not the probability but the probability amplitude of path  $i$ . That is, in quantum mechanics we have to add the  $P^{\text{amp}}$  of every path, then take the absolute value and square it in order to get the probability of the particle going from  $a$  to  $b$  (notice the identity of the path is lost here, because there is no such thing as a precise path in quantum mechanics):

$$p(ab) \propto \left| \sum_{i=1}^N P_i^{\text{amp}}(ab) \right|^2. \quad (11)$$

In this case, the largest contribution to  $p(ab)$  comes from the path which makes  $A_{ab}(i)$  stationary: in the limit of large  $A_{ab}(i)/\hbar$  the imaginary exponential oscillates so fast that most paths cancel each other out in the final amplitude; the only path not canceled is the one which makes the action stationary, i.e., the classical path (see Ref. [10] for more details).

### III. QUANTUM MECHANICAL EQUATIONS FROM MAXCAL

Gonzalez *et al.* [7] have shown that Newton’s second law can be derived from MaxCal imposing two constraints, one related to the magnitude of the squared displacement, and the other related to the probability distribution of the position. Inspired by this, and noticing the above-mentioned relation between MaxCal and the path integral formulation of quantum mechanics, we ask ourselves what kind of constraints are needed in order for the different quantum mechanical Lagrangians to appear in the probability of path  $i$  as given by MaxCal?

The Lagrangians we propose to find are the nonrelativistic Schrödinger, and the relativistic Klein–Gordon and Dirac. We start with the latter two because they are more straightforward.

#### A. Klein–Gordon Lagrangian

Figure 1 depicts the paths connecting phase-space points  $a$  and  $b$ , where each path  $i$  is discretized in time steps  $\Delta t$ , resulting in  $N$  total paths, with  $n_i$  number of segments each. Let us propose the following constraint, where we use Einstein’s notation (repeated indexes in the same term, one as a subscript, the other as a superscript, mean summation with respect to

TABLE I. Probability as a function of the action of three different paths (with values of their action of 1, 2, and 3, in arbitrary units), for different  $\eta$  values

$A_{ab}(i)$	1	2	3
$p(i)$ (with $\eta = 1$ )	0.66	0.24	0.09
$p(i)$ (with $\eta = 10$ )	0.99	$10^{-5}$	$10^{-9}$
$p(i)$ (with $\eta = 100$ )	1	$10^{-44}$	$10^{-87}$

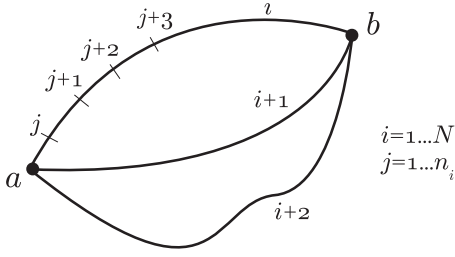


FIG. 1. Points  $a$  and  $b$  can be connected by  $N$  different paths, each designated by an index  $i$ . And each path's length is discretized in  $n_i$  intervals, designated by index  $j$ .

them):

$$\left\langle \sum_{j=1}^{n_i} \frac{\Delta \phi_j}{\Delta x_\mu} \frac{\Delta \phi_j}{\Delta x^\mu} \Delta^4 x \right\rangle_i = a, \quad (12)$$

$$\left\langle \sum_{j=1}^{n_i} |\phi_j|^2 \Delta^4 x \right\rangle_i = b. \quad (13)$$

Now, using the Lagrange multipliers technique, we extremize the caliber, subjected to the above constraints, and impose also that the sum of probabilities must equal 1. That is, we define the function

$$S' = \sum_{i=1}^N p_i \ln(p_i) - \lambda \sum_{i=1}^N p_i - \beta \sum_{i=1}^N \left( \sum_{j=1}^{n_i} \frac{\Delta \phi_j}{\Delta x_\mu} \frac{\Delta \phi_j}{\Delta x^\mu} \Delta^4 x \right) - \gamma \sum_{i=1}^N \left( \sum_{j=1}^{n_i} |\phi_j|^2 \Delta^4 x \right), \quad (14)$$

and requiring its derivatives with respect to  $p_i$  and to the multipliers,  $\lambda$ ,  $\beta$ , and  $\gamma$ , to be zero, we obtain the probability of path  $i$ :

$$p_i(ab) = \frac{1}{Z} e^{-A_{ab}(i)}, \quad (15)$$

where

$$A_{ab}(i) = -\beta \sum_{j=1}^{n_i} (\partial_\mu \phi_j)(\partial^\mu \phi_j) \Delta^4 x - \gamma \sum_{j=1}^{n_i} |\phi_j|^2 \Delta^4 x, \quad (16)$$

and

$$Z = \sum_{i=1}^N e^{-A_{ab}(i)}. \quad (17)$$

Notice in Eq. (16) we have replaced the rates of change of the field  $\Delta \phi_j / \Delta x^\mu$  by the derivatives  $\partial_\mu \phi_j$ .  $Z$  is a normalization factor.

Finally, recalling Eq. (7) and comparing with Eq. (15), we see that  $A$  is the action. And since the action is related to the Lagrangian through  $A = \int L d^4 x$ , we can conclude—after transforming from the discrete to the continuous case—that the corresponding Lagrangian is

$$L = \frac{1}{2} (\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 |\phi|^2. \quad (18)$$

In the last step we assigned the following values to the multipliers to recover the well-known Klein–Gordon Lagrangian:

$$\beta \rightarrow -1/2, \quad (19)$$

$$\gamma \rightarrow m^2/2. \quad (20)$$

We see that the mass is associated with the constraint in Eq. (13), via the  $\gamma$  multiplier (it has to be this way, since in relativistic field theories, the Lagrangian term quadratic in the field is the one that gives mass to the field). To interpret this, let us recall from the theory of Lagrange multipliers [11] that the multiplier has a specific meaning: it is the derivative of the function being extremized (the caliber  $S$ ) with respect to the value of the constraint [ $b$  in Eq. (13)]:  $m^2 \sim \partial S / \partial b$ . Thus, the mass of the field is associated with the rate of change of the caliber with respect to  $b$ . That is, a large  $m$  signifies that changing  $b$  would largely change  $S$ , getting it out of the stationary point (in which it *wants* to stay, according to MaxCal). On the contrary, a field with a small  $m$  would barely change the value of  $S$  when varying  $b$ . In conclusion, the mass has to do with the slope of the curve  $S$  vs  $b$ : zero slope associated with zero mass, and increasing slopes associated with increasing mass. Interpreting this in a more intuitive physical way, a field with zero mass could be thought of as having no constraints on the values that the field can take in different points of spacetime. A massive field, on the other hand, cannot take any value. This suggests an interpretation of the concept of inertia: The value of the field of a large mass cannot change easily, since the change in  $S$  would be large. Thus, the field (or particle) can only change its state of motion slowly, with a large inertia. But a light particle can move fast and change its path easily, since it has no cost to change the values of the field (thus, small inertia).

## B. Dirac Lagrangian

Analogously, we now propose constraints which will lead to the Dirac Lagrangian ( $L = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi$ ). These are

$$\left\langle \sum_{j=1}^{n_i} \bar{\psi}_j \gamma^\mu \frac{\Delta \psi_j}{\Delta x_\mu} \Delta^4 x \right\rangle_i = a, \quad (21)$$

$$\left\langle \sum_{j=1}^{n_i} \bar{\psi}_j \psi_j \Delta^4 x \right\rangle_i = b, \quad (22)$$

along with the constraint of total probability equal to 1. Applying Lagrange multipliers to extremize the caliber, we arrive at an expression for the probability of path  $i$  as that in Eq. (15), but with the action now given by

$$A_{ab}(i) = -\beta \sum_{j=1}^{n_i} \bar{\psi}_j \gamma^\mu (\partial_\mu \psi_j) \Delta^4 x - \gamma \sum_{j=1}^{n_i} \bar{\psi}_j \psi_j \Delta^4 x. \quad (23)$$

Or, going to the continuous limit,

$$L = i \bar{\psi} \gamma^\mu (\partial_\mu \psi) - m \bar{\psi} \psi, \quad (24)$$

where we have assigned the following values to the multipliers:

$$\beta \rightarrow -i, \quad (25)$$

$$\gamma \rightarrow m. \quad (26)$$

We see that, in this case, it is  $m$  that is associated with the cost of having defined values of the square of the field, and not  $m^2$  as in the previous case. This difference seems interesting, but should not be surprising, since the fields themselves are intrinsically different, with Klein–Gordon fields being scalar and Dirac fields vectorial.

### C. Schrödinger Lagrangian

We could repeat the above process again, this time to show that the Schrödinger Lagrangian can also be obtained by proposing suitable constraints. These are

$$\left\langle \sum_{j=1}^{n_i} \frac{\Delta\psi}{\Delta x_j} \frac{\Delta\psi^*}{\Delta x_j} \Delta t \right\rangle_i = a, \quad (27)$$

$$\left\langle \sum_{j=1}^{n_i} \psi^* \frac{\Delta\psi}{\Delta t} \Delta t \right\rangle_i = b, \quad (28)$$

$$\left\langle \sum_{j=1}^{n_i} \psi^* \psi \Delta t \right\rangle_i = c. \quad (29)$$

And again, using Lagrange multipliers and taking the continuous limit, we get

$$L = \beta(\partial_j\psi)(\partial_j\psi^*) + \gamma\psi^*\dot{\psi} + \delta\psi^*\psi. \quad (30)$$

The assignment of multipliers this time is

$$\beta \rightarrow \frac{\hbar^2}{8\pi^2 m}, \quad (31)$$

$$\gamma \rightarrow \frac{\hbar}{4\pi i}, \quad (32)$$

$$\delta \rightarrow V. \quad (33)$$

Not surprisingly, mass is not associated in this case with the quadratic term in the field, but with the one with the spatial derivatives. In general, relativistic theories like Klein–Gordon and Dirac have mass provided by the Lagrangian term that is quadratic in the field; nonrelativistic theories, like Schrödinger, have mass in the terms with spatial derivatives. This is a consequence of relativistic and nonrelativistic theories obeying different dispersion relations:  $E^2 = p^2 + m^2$  and  $E = p^2/2m$ , respectively. Recalling that, in quantum mechanics, the  $\hat{p}$  operator is associated with spatial derivatives, we see the reason why the Schrödinger Lagrangian has  $m$  dividing them.

## IV. CONCLUSION

In this work we have shown, first, that the principle of least action follows from MaxCal, with the constraint that the value of the action in each path and the average action are both known. Notice that we have not shown why the constraint has to be on the action, but we have taken the action as a fundamental quantity that describes (constrains) the different paths. But

MaxCal is more general than least action, as it allows not only the classical path which extremizes the action and is thus the most likely, but gives nonzero probability for other paths [Eq. (7)]. This fact leads to a connection between MaxCal and the path integral formulation of quantum mechanics, because probability amplitudes in that formalism are also described by exponential functions of the action [Eq. (10)]. Unless further work says otherwise, this should be taken more as an analogy than as a direct connection, since the relation we mentioned is between the *probability* of a specific path  $i$  in MaxCal, and the *probability amplitude* of a specific path  $i$  in the path integral formulation. Those concepts, although firmly related [through Eq. (11)], are not the same.

Inspired by this similarity, we probed the relevance of MaxCal to quantum mechanics; in particular, the possibility of obtaining quantum equations from that principle. And, in fact, we showed that the Klein–Gordon, Dirac, and Schrödinger Lagrangians can be obtained from MaxCal by imposing constraints on the average value of products of the field and/or its derivatives [see constraint equations (12), (13), (21), (22), and (27)–(29)]. The procedure we followed in order to get them was backwards, i.e., knowing the Lagrangians, we proposed constraints that would lead to them.

An interesting finding that arises from the constraints is that related to the mass of the fields. In the relativistic cases, we have seen that the mass of the field is generated from the constraints in Eqs. (13) and (22): from the theory of Lagrange multipliers, mass is the cost of having the average of the squared fields set to a specific value. A large mass means that the specific value of  $b$  in the constraint can hardly be changed, while a massless field would signify that the value of  $b$  is not important and can easily be changed. In turn, this appears as an explanation of the concept of inertia; that is, it is difficult to change the values of the field (and thus, the path of the particle) when it has a large mass, but not when it is light.

Related to the other constraints used in the quantum cases, one may wonder what is special about those constraints? Why not others? To answer this question we rely on the fact, well known in field theory, that Lagrangians need to be covariant and, thus, their terms need to be scalars. In this way, the terms available to a constraint (from which the Lagrangian will be generated) are only those that are scalar; for example, combinations of field and derivatives where all indices are summed over. If we also require simple combinations of field and first derivatives (higher-order derivatives are usually not needed in most theories), then the possible terms are highly restricted to those used in our constraints.

In summary, this work shows that the MaxCal principle, when complemented with the right set of constraints, is useful not only in classical mechanics but also in quantum theory because it allows the derivation of relevant Lagrangians, like Schrödinger, Klein–Gordon, and Dirac’s. Moreover, exploring the nature of the constraints and of the associated Lagrange multipliers may lead, as we have shown in relation to the mass, to insights into important physical concepts.

Finally, an intriguing question arises from the similarity of Eqs. (7) and (10). Is their relation just qualitative, as we have used it here? It would be interesting to see if there is a way to firmly connect one to the other, thus logically connecting classical to quantum mechanics.

## ACKNOWLEDGMENT

The author acknowledges support from Agencia Nacional de Promoción Científica y Tecnológica through grant PICT-2015-3832.

- 
- [1] E. T. Jaynes, *Annu. Rev. Phys. Chem.* **31**, 579 (1980).
  - [2] E. T. Jaynes, *Phys. Rev.* **106**, 620 (1957).
  - [3] E. T. Jaynes, *Phys. Rev.* **108**, 171 (1957).
  - [4] S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, *Rev. Mod. Phys.* **85**, 1115 (2013).
  - [5] Q. Wang, *Chaos, Solitons Fractals* **23**, 1253 (2005).
  - [6] Q. A. Wang, *Astrophys. Space Sci.* **305**, 273 (2006).
  - [7] D. González, S. Davis, and G. Gutiérrez, *Found. Phys.* **44**, 923 (2014).
  - [8] P. D. Dixit, J. Wagoner, C. Weistuch, S. Pressé, K. Ghosh, and K. A. Dill, *J. Chem. Phys.* **148**, 010901 (2018).
  - [9] S. Davis and D. González, *J. Phys. A: Math. Theor.* **48**, 425003 (2015).
  - [10] R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, International Series in Pure and Applied Physics (McGraw-Hill, New York, 1965).
  - [11] H. Karabulut, *Eur. J. Phys.* **27**, 709 (2006).