NAUSIVIOS CHORA

A Journal in Naval Sciences and Technology

PART C:

Natural Sciences and Mathematics

ΝΑΥΣΙΒΙΟΣ ΧΩΡΑ

Περιοδική Έκδοση Ναυτικών Επιστημών

ΜΕΡΟΣ Γ:

Φυσικές Επιστήμες και Μαθηματικά

Volume 7/2018 - Τεύχος 7/2018

TABLE OF CONTENTS

On solving Maxwell's equations for a charging capacitor

C. J. Papachristou[1] and A. N. Magoulas[2]

[1]*Department of Physical Sciences, Hellenic Naval Academy, Piraeus 18539, Greece*
E-mail: papachristou@snd.edu.gr

[2]*Department of Electrical Engineering, Hellenic Naval Academy, Piraeus 18539, Greece*
E-mail: aris@snd.edu.gr

Abstract. The charging capacitor is used as a standard paradigm for illustrating the concept of the Maxwell "displacement current". A certain aspect of the problem, however, is often overlooked. It concerns the conditions for satisfaction of the Faraday-Henry law both in the interior and the exterior of the capacitor. In this article the situation is analyzed and a mathematical process is described for obtaining expressions for the electromagnetic field that satisfy the full set of Maxwell's equations both inside and outside the capacitor.

Keywords: Maxwell equations, Faraday-Henry law, displacement current, capacitor

PACS: 03.50.De, 41.20.-q, 01.55.+b, 01.40.Fk

1. Introduction

The charging capacitor is the standard paradigm used in intermediate-level Physics courses, textbooks and articles to demonstrate the significance of the Maxwell "displacement current" (see, e.g., [1-7]). The point is correctly made that, without this "current" term, the static Ampère's law would be incomplete with regard to explaining the conservation of charge as well as the existence of electromagnetic radiation. Also, the line integral of the magnetic field around a closed curve would be an ill-defined concept.

There is, however, a certain subtlety of the situation which is often passed by. It concerns the satisfaction of the Faraday-Henry law both inside and outside the capacitor. Indeed, although care is taken to ensure that the expressions used for the electromagnetic (e/m) field satisfy the Ampère-Maxwell law, no such care is exercised with regard to the Faraday-Henry law. As it turns out, the usual formulas for the e/m field satisfy this latter law only in the special case where the capacitor is being charged at a constant rate. But, if the current responsible for charging the capacitor is time-dependent, this will also be the case with the magnetic field outside the capacitor. This, in turn, implies the existence of an "induced" electric field in that region, contrary to the usual assertion that the electric field outside the capacitor is zero. Moreover, the time dependence of the magnetic field inside the capacitor is not compatible with the assumption that the electric field in that region is uniform, as the case would be in a static situation.

The purpose of this article is to exhibit the theoretical inconsistencies inherent in the "classical" treatment of the charging capacitor problem and to describe a mathematical process

ISSN:1791-4469

for deriving expressions for the e/m field that satisfy the full set of the Maxwell equations (including, of course, the Faraday-Henry law) both inside and outside the capacitor.

After a preliminary discussion of the concept of the electric current through a loop (Section 2), the standard "textbook" approach to the charging-capacitor example in connection with the concept of the displacement current is presented in Section 3. New and more general solutions of the Maxwell system of equations in the interior and the exterior of the capacitor are then derived in Sections 4 and 5, respectively.

2. The current through a loop

Before we proceed to write the Ampère-Maxwell law in its integral form, we must carefully define the concept of the *total current through a loop C* (where by "loop" we mean a closed curve in space).

Proposition. Consider a region R of space within which the distribution of charge, expressed by the volume charge density ρ, is time-independent ($\partial\rho/\partial t=0$). Let C be an oriented loop in R, and let S be any open surface in R bordered by C and oriented accordingly. We define the total current through C as the surface integral of the current density $\vec{J}$ over S:

$$I_{in} = \int_S \vec{J} \cdot \overrightarrow{da} \qquad (1)$$

Then, the quantity I_{in} has a well-defined value independent of the particular choice of S (that is, I_{in} is the same for all open surfaces S bounded by C).

Proof. By the equation of continuity for the electric charge (see, e.g., [8], Chap. 6) and by the fact that the charge density ρ inside the region R is static ($\partial\rho/\partial t=0$), we have that $\vec{\nabla}\cdot\vec{J}=0$. Therefore, within this region of space the current density has the properties of a solenoidal field. In particular, the value of the surface integral of $\vec{J}$ will be the same for all open surfaces S sharing a common border C.

As an example, let us consider a circuit carrying a time-dependent current $I(t)$. If the circuit does not contain a capacitor, no charge is piling up at any point and the charge density at any elementary segment of the circuit is constant in time. Moreover, at each instant t, the current I is constant along the circuit, its value changing only with time. Now, if C is a loop encircling some section the circuit, as shown in Fig. 1, then, at each instant t, the same current $I(t)$ will pass through any open surface S bordered by C. Thus, the integral in (1) is well defined for all t, assuming the same value $I_{in}=I(t)$ for all S.

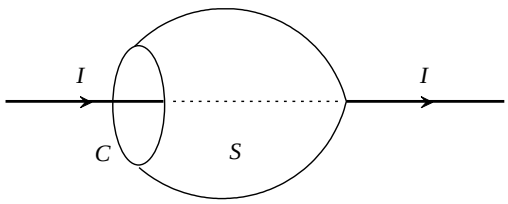


Figure 1

Things change if the circuit contains a capacitor which is charging or discharging. It is then no longer true that the current $I(t)$ is constant along the circuit; indeed, $I(t)$ is zero inside the capacitor and nonzero outside. Thus, the value of the integral in (1) depends on whether the surface S does or does not contain points belonging to the interior of the capacitor.

3. Maxwell displacement current in a charging capacitor

Figure 2 shows a simple circuit containing a capacitor that is being charged by a time-dependent current $I(t)$. At time t, the plates of the capacitor, each of area A, carry charges $\pm Q(t)$.

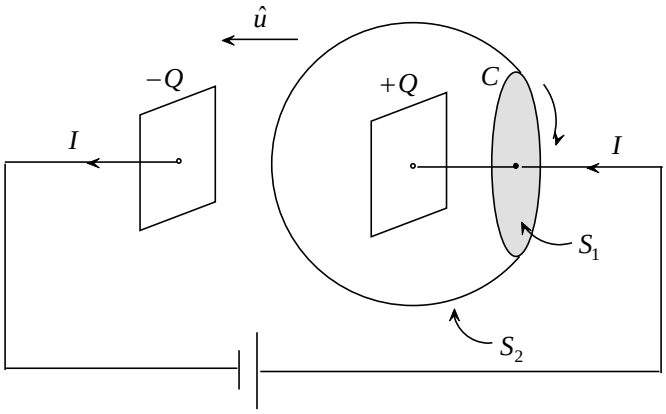


Figure 2

Assume that we encircle the current I by an imaginary plane loop C parallel to the positive plate and oriented in accordance with the "right-hand rule", consistently with the direction of I (this direction is indicated by the unit vector $\hat{u}$). The "current through C" is here an ill-defined notion since the value of the integral in Eq. (1) is $I_{in}=I$ for the flat surface S_1 and $I_{in}=0$ for the curved surface S_2 (Fig. 2). This, in turn, implies that Ampère's law of magnetostatics [1-4,8] cannot be valid in this case, given that, according to this law, the integral of the magnetic field $\vec{B}$ along the loop C, equal to $\mu_0 I_{in}$, would not be uniquely defined but would depend on the surface S bounded by C.

Maxwell restored the single-valuedness of the closed line integral of $\vec{B}$ by introducing the so-called *displacement current*, which is essentially the rate of change of a time-dependent electric field:

$$\vec{J}_d = \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \quad \Leftrightarrow \quad I_d = \int_S \vec{J}_d \cdot \overrightarrow{da} = \varepsilon_0 \int_S \frac{\partial \vec{E}}{\partial t} \cdot \overrightarrow{da} \qquad (2)$$

The *Ampère-Maxwell law* reads:

ISSN:1791-4469

$$\vec{\nabla}\times\vec{B}=\mu_0\vec{J}+\varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}\quad\Leftrightarrow$$
$$\oint_C\vec{B}\cdot\overrightarrow{dl}=\mu_0 I_{in}+\varepsilon_0\mu_0\int_S\frac{\partial\vec{E}}{\partial t}\cdot\overrightarrow{da}\equiv\mu_0\,(I+I_d)_{in} \tag{3}$$

where I_{in} is given by Eq. (1).

Now, the standard "textbook" approach to the charging capacitor problem goes as follows: Outside the capacitor the electric field vanishes everywhere, while inside the capacitor the electric field is uniform – albeit time-dependent – and has the static-field-like form

$$\vec{E}=\frac{\sigma(t)}{\varepsilon_0}\,\hat{u}=\frac{Q(t)}{\varepsilon_0 A}\,\hat{u} \tag{4}$$

where $\sigma(t)=Q(t)/A$ is the surface charge density on the positive plate at time t. This density is related to the current I, which charges the capacitor, by

$$\sigma'(t)=\frac{Q'(t)}{A}=\frac{I(t)}{A} \tag{5}$$

(the prime indicates differentiation with respect to t). Thus, inside the capacitor,

$$\frac{\partial\vec{E}}{\partial t}=\frac{\sigma'(t)}{\varepsilon_0}\,\hat{u}=\frac{I(t)}{\varepsilon_0 A}\,\hat{u} \tag{6}$$

Outside the capacitor the time derivative of the electric field vanishes everywhere and, therefore, so does the displacement current.

Now, on the flat surface S_1 the total current through C is $(I+I_d)_{in}=I+0=I(t)$. The Ampère-Maxwell law (3) then yields:

$$\int_C\vec{B}\cdot\overrightarrow{dl}=\mu_0 I(t) \tag{7}$$

On the curved surface S_2, the total current through C is $(I+I_d)_{in}=0+I_{d,in}=I_{d,in}$, where the quantity on the right assumes a nonzero value only for the portion S_2' of S_2 which lies inside the capacitor. This quantity is equal to

$$I_{d,in}=\varepsilon_0\int_{S_2'}\frac{\partial\vec{E}}{\partial t}\cdot\overrightarrow{da}=\frac{I(t)}{A}\int_{S_2'}\hat{u}\cdot\overrightarrow{da} \tag{8}$$

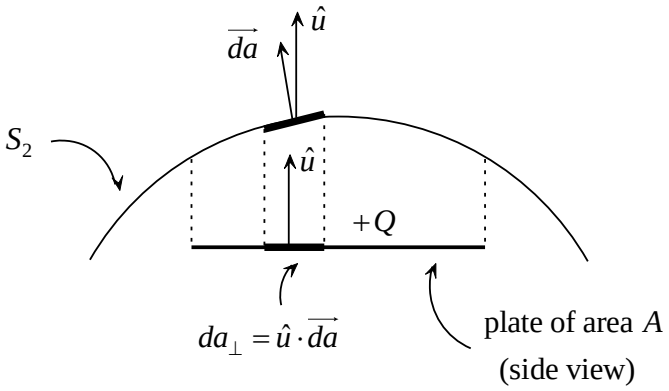


Figure 3

The dot product in the integral on the right of (8) represents the projection of the surface element $\overrightarrow{da}$ onto the axis defined by the unit vector $\hat{u}$ (see Fig. 3). This is equal to the projection $da_\perp$ of an elementary area da of S_2' onto the flat surface of the plate of the capacitor. Eventually, the integral on the right of (8) equals the total area A of the plate. Hence, $I_{d,in}=I(t)$ and, given that $I_{in}=0$ on S_2, the Ampère-Maxwell law (3) again yields the result (7).

So, everything works fine with regard to the Ampère-Maxwell law, but there is one law we have forgotten so far; namely, the *Faraday-Henry law*! According to that law, a time-changing magnetic field is always accompanied by an electric field (or, as is often said, "induces" an electric field). So, the electric field outside the capacitor cannot be zero, as claimed previously, given that the time-dependent current $I(t)$ is expected to generate a time-dependent magnetic field. For a similar reason, the electric field inside the capacitor cannot have the static-field-like form (4) (there must also be a contribution from the rate of change of the magnetic field between the plates).

An exception occurs if the current I which charges the capacitor is constant in time, since in this case the magnetic field will be static everywhere. This is actually the assumption silently or explicitly made in many textbooks (see, e.g., [2], Chap. 21). Physically this means that the capacitor is being charged at a constant rate. But, in the general case where $I(t)\neq$constant, the preceding discussion regarding the charging capacitor problem needs to be significantly revised in order to take into account the entire set of the Maxwell equations; in particular, the Ampère-Maxwell law as well as the Faraday-Henry law.

4. The Maxwell equations inside the capacitor

We consider a parallel-plate capacitor with circular plates of radius a, thus of area $A=\pi a^2$. The space in between the plates is assumed to be empty of matter. The capacitor is being charged by a time-dependent current $I(t)$ flowing in the $+z$ direction. The z-axis is perpendicular to the plates (the latter are therefore parallel to the xy-plane) and passes through their centers, as seen in Fig. 4 (by $\hat{u}_z$ we denote the unit vector in the $+z$ direction).

ISSN:1791-4469

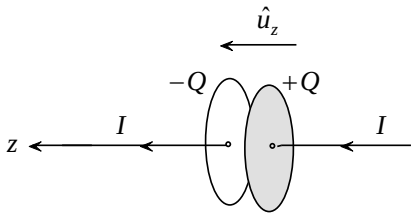


Figure 4

The capacitor is being charged at a rate $dQ/dt=I(t)$, where $+Q(t)$ is the charge on the right plate (as seen in the figure) at time t. If $\sigma(t)=Q(t)/\pi a^2=Q(t)/A$ is the surface charge density on the right plate, then the time derivative of σ is given by Eq. (5).

We assume that the plate separation is very small compared to the radius a, so that the electromagnetic (e/m) field inside the capacitor is practically independent of z, although it *does* depend on the normal distance ρ from the z-axis. (We will not be concerned with edge effects, thus we will restrict out attention to points that are not too close to the edges of the plates.) In cylindrical coordinates (ρ,φ,z) the magnitude of the e/m field at any time t will thus only depend on ρ (it will not depend on the angle φ, as follows by the symmetry of the problem).

We assume that the positive and the negative plate of the capacitor of Fig. 4 are centered at $z=0$ and $z=d$, respectively, on the z-axis, where, as mentioned above, the plate separation d is much smaller than the radius a of the plates. The interior of the capacitor is then the region of space with $0\le\rho<a$ and $0<z<d$.

The magnetic field inside the capacitor is azimuthal, of the form $\vec{B}=B(\rho,t)\hat{u}_\varphi$. As noted in Sec. 3, a standard practice is to assume that, at all t, the electric field in this region is uniform, of the form

$$\vec{E}=\frac{\sigma(t)}{\varepsilon_0}\,\hat{u}_z \tag{9}$$

while everywhere outside the capacitor the electric field vanishes. With this assumption the magnetic field inside the capacitor is found to be [2,3,6]

$$\vec{B}=\frac{\mu_0 I(t)\rho}{2\pi a^2}\,\hat{u}_\varphi=\frac{\mu_0 I(t)\rho}{2A}\,\hat{u}_\varphi \tag{10}$$

Expressions (9) and (10) must, of course, satisfy the Maxwell system of equations in empty space, which system we write in the form [1,8]

$$\begin{aligned} &(a)\quad \vec{\nabla}\cdot\vec{E}=0 \qquad && (c)\quad \vec{\nabla}\times\vec{E}=-\frac{\partial\vec{B}}{\partial t}\\ &(b)\quad \vec{\nabla}\cdot\vec{B}=0 \qquad && (d)\quad \vec{\nabla}\times\vec{B}=\varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t} \end{aligned} \tag{11}$$

By using cylindrical coordinates (see Appendix) and by taking into account that $\sigma'(t)=I(t)/A$ [Eq. (5)], it is not hard to show that (9) and (10) satisfy three of Eqs. (11), namely, (*a*), (*b*) and (*d*). This is not the case with the Faraday-Henry law (11*c*), however, since by (9) and (10) we find that $\vec{\nabla}\times\vec{E}=0$, while

$$\frac{\partial\vec{B}}{\partial t}=\frac{\mu_0 I'(t)\rho}{2A}\,\hat{u}_\varphi\ .$$

An exception occurs if the current I is constant in time, i.e., if the capacitor is being charged at a constant rate, so that $I'(t)=0$. But, for a current $I(t)$ with arbitrary time dependence, the pair of fields (9) and (10) does not satisfy the third Maxwell equation.

To remedy the situation and restore the validity of the full set of Maxwell's equations in the interior of the capacitor, we must somehow correct the expressions (9) and (10) for the e/m field. To this end, we employ the following *Ansatz*:

$$\vec{E}=\left(\frac{\sigma(t)}{\varepsilon_0}+f(\rho,t)\right)\hat{u}_z\ ,\quad \vec{B}=\left(\frac{\mu_0 I(t)\rho}{2A}+g(\rho,t)\right)\hat{u}_\varphi\ ;\qquad (12)$$
$$\sigma'(t)=I(t)/A$$

where $f(\rho,t)$ and $g(\rho,t)$ are functions to be determined consistently with the given current function $I(t)$ and the given initial conditions. It is easy to check that the solutions (12) automatically satisfy the first two Maxwell equations (11*a*) and (11*b*). By the Faraday-Henry law (11*c*) and the Ampère-Maxwell law (11*d*) we get the following system of partial differential equations:

$$\frac{\partial f}{\partial\rho}=\frac{\partial g}{\partial t}+\frac{\mu_0 I'(t)\rho}{2A}\qquad (a)$$
$$\frac{\partial(\rho g)}{\partial\rho}=\varepsilon_0\mu_0\frac{\partial(\rho f)}{\partial t}\qquad (b)\qquad (13)$$

Note in particular that the "classical" solution with $f(\rho,t)\equiv 0$ and $g(\rho,t)\equiv 0$ is possible only if $I'(t)=0 \Leftrightarrow I=$constant in time (i.e., if the capacitor is being charged at a constant rate), as mentioned earlier.

As a special case, let us assume that the functions f and g are time-independent, i.e., $\partial f/\partial t = \partial g/\partial t=0 \Leftrightarrow f=f(\rho)$, $g=g(\rho)$. From (13*a*) we get (ignoring an arbitrary constant):

$$f(\rho)=\frac{\mu_0 I'(t)\rho^2}{4A}\ .$$

This can only be valid if $I'(t)=$constant $\Leftrightarrow I''(t)=0$. On the other hand, (13*b*) yields: $\rho g=$constant $\equiv\lambda \Leftrightarrow g(\rho)=\lambda/\rho$. In order for $g(\rho)$ to be finite for $\rho=0$, we must set $\lambda=0$, so that $g(\rho)\equiv 0$. The solution (12) for the e/m field inside the capacitor is then written:

ISSN:1791-4469

$$\vec{E} = \left(\frac{\sigma(t)}{\varepsilon_0} + \frac{\mu_0 I'(t)\rho^2}{4A} \right) \hat{u}_z \ , \quad \vec{B} = \frac{\mu_0 I(t)\rho}{2A} \hat{u}_\varphi \ ; \qquad (14)$$
$$I''(t) = 0 \ , \quad \sigma'(t) = I(t)/A$$

We notice that, since $I''(t)=0$, Eq. (6) is still valid and the displacement current inside the capacitor is again given by $I_d = I(t)$. What *is* different here is the correction to the electric field in order for the Faraday-Henry law to be satisfied.

5. The Maxwell equations outside the capacitor

We recall that the positive and the negative plate of the capacitor of Fig. 4 are centered at $z=0$ and $z=d$, respectively, on the z-axis, where the plate separation d is much smaller than the radius a of the plates. The space exterior to the capacitor consists of points with $\rho > 0$ and $z \notin (0,d)$, as well as points with $\rho > a$ and $0 < z < d$. (In the former case we exclude points on the z-axis, with $\rho=0$, to ensure the finiteness of our solutions in that region.)

The e/m field outside the capacitor is usually described mathematically by the equations [2,3,6]

$$\vec{E} = 0 \ , \quad \vec{B} = \frac{\mu_0 I(t)}{2\pi\rho} \hat{u}_\varphi \qquad (15)$$

As the case is with the standard solutions in the interior of the capacitor, the solutions (15) fail to satisfy the Faraday-Henry law (11*c*) (although they do satisfy the remaining three Maxwell equations), since $\vec{\nabla} \times \vec{E} = 0$ while

$$\frac{\partial \vec{B}}{\partial t} = \frac{\mu_0 I'(t)}{2\pi\rho} \hat{u}_\varphi \ .$$

As before, an exception occurs if the current I is constant in time, i.e., if the capacitor is being charged at a constant rate, so that $I'(t)=0$.

To find more general solutions that satisfy the entire set of the Maxwell equations, we work as in the previous section. Thus, we assume the following general form of the e/m field everywhere outside the capacitor:

$$\vec{E} = f(\rho,t)\,\hat{u}_z \ , \quad \vec{B} = \left(\frac{\mu_0 I(t)}{2\pi\rho} + g(\rho,t) \right) \hat{u}_\varphi \qquad (16)$$

where f and g are functions to be determined consistently with the given current function $I(t)$. The solutions (16) automatically satisfy the first two Maxwell equations (11*a*) and (11*b*). By Eqs. (11*c*) and (11*d*) we get the following system of partial differential equations:

$$\frac{\partial f}{\partial \rho} = \frac{\partial g}{\partial t} + \frac{\mu_0 I'(t)}{2\pi\rho} \qquad (a)$$
$$\frac{\partial(\rho g)}{\partial \rho} = \varepsilon_0 \mu_0 \frac{\partial(\rho f)}{\partial t} \qquad (b) \qquad (17)$$

Again, the usual solution with $f(\rho,t)\equiv 0$ and $g(\rho,t)\equiv 0$ is possible only if $I'(t)=0$, i.e., if the capacitor is being charged at a constant rate.

As a special case, let us assume that the functions f and g are time-independent, i.e., $f=f(\rho)$, $g=g(\rho)$. From (17a) we get:

$$f(\rho) = \frac{\mu_0 I'(t)}{2\pi} \ln(\kappa\rho)$$

where κ is a positive constant quantity having dimensions of inverse length. This can only be valid if $I'(t)=\text{constant} \Leftrightarrow I''(t)=0$. On the other hand, (17$b$) yields: $\rho g=\text{constant} \equiv \lambda \Leftrightarrow g(\rho)=\lambda/\rho$. Since $\rho>0$, by assumption, we could now let $\lambda\neq 0$. For reasons of continuity, however (see below), we set $\lambda=0$, so that $g=0$. The solution (16) for the e/m field outside the capacitor is then written:

$$\vec{E} = \frac{\mu_0 I'(t)}{2\pi} \ln(\kappa\rho)\, \hat{u}_z \,, \quad \vec{B} = \frac{\mu_0 I(t)}{2\pi\rho}\, \hat{u}_\varphi \,; \qquad (18)$$
$$I''(t) = 0$$

Note, in particular, that the magnetic field in the strip $0 < z < d$ is continuous for $\rho=a$, since the expression for $\vec{B}$ in (18) matches the corresponding expression in (14) upon substituting $\rho=a$ (remember that $A=\pi a^2$). No analogous continuity exists, however, for the electric field. Physically, this may be attributed to fringing effects at the edges of the plates.

6. Summary

The purpose of this article is to point out the need to revisit the problem of the charging capacitor, as this is discussed in connection with the Maxwell displacement current, and to carefully examine the expressions for the e/m field both in the interior and the exterior of this system. As was noted, the standard formulas assumed for this field, tailor-made to satisfy the Ampère-Maxwell law, fail to satisfy the Faraday-Henry law except in the special case where the capacitor is being charged at a constant rate. We have derived general expressions for the e/m field that satisfy the full set of Maxwell's equations for arbitrary charging rate of the system. These results may reduce to the familiar set of equations in the case of a constant charging rate.

Note

This article is an extensively revised and expanded version of an article published previously in letter form [9]. In particular, the results contained in Sec. 5 of this article are new.

ISSN:1791-4469

Appendix: Vector operators in cylindrical coordinates

Let $\vec{A}$ be a vector field, expressed in cylindrical coordinates (ρ, φ, z) as

$$\vec{A} = A_\rho(\rho,\varphi,z)\hat{u}_\rho + A_\varphi(\rho,\varphi,z)\hat{u}_\varphi + A_z(\rho,\varphi,z)\hat{u}_z \ .$$

The *div* and the *rot* of this field, in this system of coordinates, are written respectively as follows:

$$\vec{\nabla}\cdot\vec{A} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho A_\rho) + \frac{1}{\rho}\frac{\partial A_\varphi}{\partial\varphi} + \frac{\partial A_z}{\partial z} \ ,$$

$$\vec{\nabla}\times\vec{A} = \left(\frac{1}{\rho}\frac{\partial A_z}{\partial\varphi} - \frac{\partial A_\varphi}{\partial z}\right)\hat{u}_\rho + \left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial\rho}\right)\hat{u}_\varphi + \frac{1}{\rho}\left(\frac{\partial}{\partial\rho}(\rho A_\varphi) - \frac{\partial A_\rho}{\partial\varphi}\right)\hat{u}_z \ .$$

In particular, if the vector field is of the form $\vec{A} = A_\varphi(\rho)\hat{u}_\varphi + A_z(\rho)\hat{u}_z$, then $\vec{\nabla}\cdot\vec{A} = 0$.

References

1. D. J. Griffiths, *Introduction to Electrodynamics*, 3rd Edition (Prentice-Hall, 1999).
2. R. K. Wangsness, *Electromagnetic Fields*, 2nd Edition (Wiley, 1986).
3. A. Shadowitz, *The Electromagnetic Field* (McGraw-Hill, 1975).
4. V. Rojansky, *Electromagnetic Fields and Waves* (Dover, 1979).
5. J. D. Jackson, *Maxwell's Displacement Current Revisited*, Eur. J. Phys. 20 (1999) 495.
6. K. T. McDonald, *Magnetic Field in a Time-Dependent Capacitor* (2017), http://www.physics.princeton.edu/~mcdonald/examples/displacement.pdf
7. K. T. Selvan, *A Revisiting of Scientific and Philosophical Perspectives on Maxwell's Displacement Current*, IEEE Antennas and Propagation Magazine, Vol. 51, No. 3 (2009) 36.
8. C. J. Papachristou, *Introduction to Electromagnetic Theory and the Physics of Conducting Solids* (META Publishing, Nov. 2017[1]); also, arXiv:1711.09969[2].
9. C. J. Papachristou, *Some Remarks on the Charging Capacitor Problem*, Advanced Electromagnetics, Vol. 7, No. 2 (2018), pp. 10-12. https://www.aemjournal.org/index.php/AEM/article/view/694

[1] http://metapublishing.org/index.php/MP/catalog/book/52

[2] https://arxiv.org/abs/1711.09969

Quantum Corrections to the Period of the Non Linear Pendulum via Path Integrals

Theodosios G. Douvropoulos

Hellenic Naval Academy, Physical Sciences Sector, Pure & Applied Physics, Hatzikyriakou Ave. Piraeus, 18539, Greece

Abstract. We treat the problem of the quantum pendulum in order to calculate the swing period of its classical counterpart. For this we apply the semiclassical path integral method to construct the system's fractional Green's function. The energy eigenvalues arise as the poles of the later, and the frequency of oscillation comes from the difference of two successive eigenvalues. The linear and non-linear potential are separately discussed and the role of the derivatives with respect to the energy of the action phase-factors, is discussed in detail.

Keywords: non linear pendulum, period of oscillation, path integral, Mathieu functions, hindered rotation, action phase factor
PACS: 03.65.Sq, 03.65.Xp, 45.50.Dd

INTRODUCTION

The simple pendulum is a famous case study in classical mechanics that still attracts attention, since it involves many problems in different branches of physics, leading to many interesting applications. Mentioning a few we can distinguish its diachronic and historical use as a clock or a metronome [1], as a seismometer [2], as a ballistic galvanometer [3], as a definite instrument for determining gravitational constant G [4] or an instrument for proving the equivalence of gravitational and inertial mass [5], part of an experimental setup for measuring viscosity [6], and many others.

It was only a couple of years after the breakthrough of quantum mechanics, that a study of the quantum pendulum appeared for the first time in science [7], by Edward Condon. Condon would soon become one of the most influential physicists in the quantum area with major contributions to atomic physics. The substantial difference in the quantum case comes from the possibility of the particle to exist in a classically inaccessible region, (barrier), where the potential energy exceeds the total energy, via the tunneling effect. Thus, quantum pendulum at most concerns with what is known as hindered or restricted rotation, which is the inhibition of rotation of molecular structures about a bond due to the presence of a sufficiently large rotational barrier, [8].

ISSN:1791-4469

In addition, the dynamics of the driven classical pendulum resembles those of the pure quantum mechanical Josephson effect [9] in superconductive structures.

In the present work we are interested to the calculation of the period of the non linear pendulum by taking in account the peculiar quantum mechanical effect of tunneling. As far as pendulum is treated inside the frame of classical mechanics governed by the Newton's laws of motion, period calculation is straightforward. Through the linearization of the force term, valid in the small angle regime, we can easily solve the differential equation of motion. The solution of the later contains the well known frequency of oscillation, $\sqrt{g/l}$, in terms of gravity acceleration and the length of the pendulum, which is amplitude independent. This is what we call isochronism. On the other hand, when we proceed without making the linear approximation, the period is produced by integrating time through the energy conservation principle, in terms of the elliptic integral of the first kind, and depends on the amplitude. Unfortunately the period is not now given in a closed form.

As far as quantum pendulum is concerned we must first proceed to the construction of Schrödinger's equation. For this, we start from the classical Lagrangian and construct the classical Hamiltonian by finding the generalized momentum of the system. Then we apply the conventional rule by replacing the canonical momentum with the corresponding quantum operator. The solution of this Schrödinger equation is a difficult task, due to the non linear potential term, not only for the eigenfunctions but for the energy eigenvalues as well, and complicated arithmetic methods are concerned. Thus, the result lacks of any analytical relations that would demonstrate the qualitative involvement of tunneling to the swing period. However if we assume a parabolic form of the potential, (which of course is equivalent to the linearization of the classical force term), we reduce the complexity to the one of the harmonic oscillator problem. The classical frequency of oscillation is then explicitly contained in the quantum energy spectrum, equal to the photon frequency for transition between two adjacent states, and is energy independent. The later is equivalent to the forementioned classical isochronism.

Instead of trying to solve Schrödinger equation we can attack the problem in a different way. For this we construct the system's Green's function, as the Fourier transformed of its propagator, via the aid of semiclassical path integration. According to semiclassical path integration, the Green's function for one dimensional physical systems is fulfilled through the mutual contribution of all possible changes in phase of the wave-function due propagation or reflection. Again, this is not trivial at all since it demands the summation of all possible changes of phase, (real or imaginary), for the infinite set of possible paths, without omitting any of them, performing the tedious calculation while always keeping the correct order of regions interchange. This results in a complicated but compact formula for the Green's function. Thus, the great advantage comes from the fact that we can now reveal the analytic form of the energy spectrum in a closed form as the approximate complex poles of the fractional Green's function, keeping their real part. This is done after expanding the denominator, (constituted by transcendental functions), in terms of analytic tunneling parameters. Again, we extract the classical period from the difference of two adjacent energy eigenvalues. The energy spectrum turns to be discrete although still infinite. The spacing between adjacent energy levels is of the order of magnitude of Planck's constant ($\sim 10^{-34}$) times the frequency of classical oscillations. This corresponds to a truly negligible quantity either for a macroscopic pendulum or a microscopic. For example, taking the length of the pendulum to be comparable in magnitude with that of a chemical bond, (meaning some Angstroms), we get an energy spacing of about 10^{-9} eV. This means that even in the classical case where all values of energy are permitted, no one can in practice achieve such a

high accuracy in measuring the energy. This makes the spacing rather unimportant. In addition the arithmetic value of the quantum corrected swing period turns to be only negligibly different from the classical. Since this is the case, we do not proceed to any quantitative calculation but rather bring out the qualitative involvement of tunneling.

The structure of the present paper is as follows. In the next section we analyze in short the dynamics of the classical pendulum stressing the lack of isochronism for increasing amplitudes, caused by the non linear nature of the potential involved in the Langrangian. In the next section we introduce the Schrödinger equation for the pendulum and demonstrate its solutions through the Mathieu functions while pointing the quantum equivalent to classical isochronism. Following this we develop the semiclassical path integral method for the construction of the system's Green's function, which is the Fourier transformed of its propagator. We then proceed to the calculation of the energy spectrum and subsequently of its classical period of motion through the difference of two successive eigenvalues. We develop a prototype formula in a compact form. Finally we mark and further analyze the role of the derivatives with respect to the energy of the action phase-factors, as these appear in the above mentioned formula. In the last section we briefly conclude.

THE CLASSICAL PENDULUM

The simple pendulum governed by the laws of classical mechanics concerns a plane motion with only one degree of freedom. It consists of a point mass m attached to one end of a weightless chord of length l, moving without any resistance acted upon by gravity and the tension in the chord, while the other end turns without friction around a fixed point. Although simple, its exact solution is not trivial at all, due to the non linear force term that appears in its differential equation

$$ml\ddot{\theta} + mg\sin\theta = 0 \tag{1}$$

The exact solution of the above equation demands the aid of elliptic functions. However we can approach pendulum's vibrational motion through the linearized equation

$$ml\ddot{\theta} + mg\theta = 0 \tag{2}$$

valid for small only amplitudes where $\sin\theta \cong \theta$. The solution of (2) which is the differential equation of a harmonic oscillator is straight forward and gives the well known result for the period of swinging

$$T_o = 2\pi\left(l/g\right)^{1/2} \tag{3}$$

On the other hand the accurate result for the period is met by integrating time in the energy conservation principle, as far as Figure 1 is concerned according to which α is the angular amplitude and the minimum of potential energy is taken as zero for θ=0. Doing so we find

ISSN:1791-4469

$$E=\frac{1}{2}ml^2\dot{\theta}^2+mgl(1-\cos\theta)=mgl(1-\cos\alpha)\Rightarrow$$

$$\frac{d\theta}{dt}=\sqrt{2\frac{g}{l}(\cos\theta-\cos\alpha)}\Rightarrow dt=\frac{d\theta}{\sqrt{2\frac{g}{l}(\cos\theta-\cos\alpha)}}\Rightarrow$$

$$T(\alpha)=4\sqrt{\frac{l}{2g}}\int_0^{\alpha}\frac{d\theta}{\sqrt{\cos\theta-\cos\alpha}}=2\sqrt{\frac{l}{g}}\int_0^{\alpha}\frac{d\theta}{\sqrt{\sin^2(\alpha/2)-\sin^2(\theta/2)}}$$

(4)

Following Landau's prescription [10] we substitute $\sin\xi=\frac{\sin(\theta/2)}{\sin(\alpha/2)}$ to convert the above relation to

$$T(\alpha)=4\sqrt{\frac{l}{g}}\,\mathrm{K}\left(\sin(\alpha/2)\right)\quad\text{or}\quad T(E)=4\sqrt{\frac{l}{g}}\,\mathrm{K}\left(\sin\left[\cos^{-1}(1-E/mgl)/2\right]\right)\qquad(5)$$

where $K(k)=\int_0^{\pi/2}\frac{d\xi}{\sqrt{1-k^2\sin^2\xi}}$ is the complete elliptic integral of the first kind [11]. The two relations are connected through the $T(\alpha)=T_o\frac{\mathrm{K}\left(\sin(\alpha/2)\right)}{\pi/2}$ and of course coincide for $\alpha\to 0$.

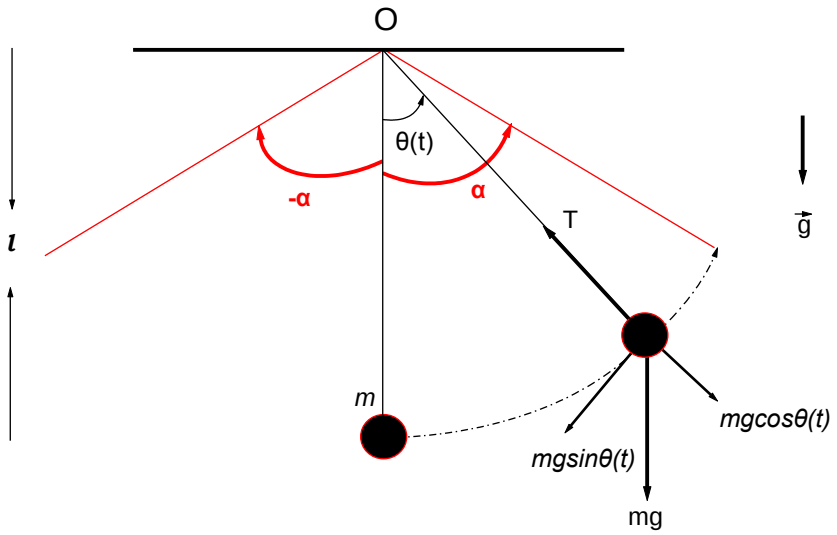


FIGURE 1. The physical pendulum. The vertical position of the chord corresponds to zero potential energy, while α stands for the angular amplitude of the swing.

The motion corresponding to (3) is isochronic, meaning that the period is independent of the amplitude and constant for a given pendulum. The reason for this comes from the Langrangian of the pendulum, which is then also linearized in the small angle regime, ($\sin\theta \cong \theta$), as $L(\theta,\dot{\theta}) = T(\dot{\theta}) - U(\theta) = \frac{1}{2}ml^2\dot{\theta}^2 - \frac{1}{2}mgl\theta^2$. Thus, if we carry out a transformation in which the angular displacement is changed by a factor λ while the time interval remains unchanged, then the Lagrangian is simply multiplied by the constant λ^2 and the equations of motion are unaltered. Thus for a given linear pendulum the equations of motion permit a series of geometrically similar paths where the times of motion between corresponding points are in unit ratio. This similarity vanishes for large amplitudes and this is clearly reflected on the dependence of the period on the amplitude and subsequently on the energy as well.

Since the elliptic integral is not easily calculated, many approximation schemes have been developed over the years for the calculation of the period of swing: see [12] and references therein. In this way many of these approximations lead to interesting applications, as is for example the use of the physical pendulum as part of the set up for measuring acoustic impedance, via the use of the Struve function [13].

THE QUANTUM PENDULUM

In order to construct the time independent Schrödinger equation for the pendulum we start from the Lagrangian $L(\theta,\dot{\theta}) = T(\dot{\theta}) - U(\theta) = \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1-\cos\theta)$ and find the system's generalized momentum [10] as

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} \Rightarrow p_\theta = ml^2\dot{\theta} \tag{6}$$

Thus the Hamiltonian can be expressed as a function of the generalized variables θ, p_θ

$$\begin{aligned} &H(\theta,\dot{\theta}) = T(\dot{\theta}) + U(\theta) = \frac{1}{2}ml^2\dot{\theta}^2 + mgl(1-\cos\theta) \Rightarrow \\ &H_g(\theta, p_\theta) = \frac{1}{2}ml^2\left(\frac{p_\theta}{ml^2}\right)^2 + mgl(1-\cos\theta) \Rightarrow \\ &H_g(\theta, p_\theta) = \frac{p_\theta^{\,2}}{2ml^2} + mgl(1-\cos\theta) \end{aligned} \tag{7}$$

Then the operator $\hat{H}$ corresponding to the classical Hamiltonian is found by applying the convention rule and replacing the canonical momentum p_θ by $-i\hbar\frac{\partial}{\partial\theta}$, to take

$$\begin{aligned} &\hat{H} = -\frac{\hbar^2}{2ml^2}\frac{\partial^2}{\partial\theta^2} + mgl(1-\cos\theta) \Rightarrow \\ &\hat{H} = -\frac{\hbar^2}{2I}\frac{\partial^2}{\partial\theta^2} + I\omega^2(1-\cos\theta) \end{aligned} \tag{8}$$

ISSN:1791-4469

where we have substituted $I = ml^2$ for the moment of inertia and $\omega^2 = \frac{g}{l}$ for the square of the classical frequency of swinging. Thus the Schrödinger equation reads

$$-\frac{\hbar^2}{2I}\frac{\partial^2\Psi(\theta)}{\partial\theta^2} + I\omega^2(1-\cos\theta)\Psi(\theta) = E\Psi(\theta) \tag{9}$$

The wavefunction $\Psi(\theta)$ has to be single-valued, meaning periodic in θ of period 2π: $\Psi(\theta+2\pi) = \Psi(\theta)$. In addition we can write (9) as

$$\frac{\partial^2\Psi(\theta)}{\partial\theta^2} + \left(\frac{2I}{\hbar^2}\left(E - I\omega^2\right) + 2\left(\frac{I\omega}{\hbar}\right)^2\cos\theta\right)\Psi(\theta) = 0 \tag{10}$$

and by further substituting $\theta=2\varphi$ as

$$\begin{aligned}&\frac{1}{4}\frac{\partial^2\Psi(\phi)}{\partial\phi^2} + \left(\frac{2I}{\hbar^2}\left(E - I\omega^2\right) + 2\left(\frac{I\omega}{\hbar}\right)^2\cos 2\phi\right)\Psi(\phi) = 0 \Rightarrow \\ &\frac{\partial^2\Psi(\phi)}{\partial\phi^2} + \left(\frac{8I}{\hbar^2}\left(E - I\omega^2\right) + 8\left(\frac{I\omega}{\hbar}\right)^2\cos 2\phi\right)\Psi(\phi) = 0\end{aligned} \tag{11}$$

where for simplicity we have used the same symbol for the wave-function. The last form is that of a Mathieu equation [14] whose standard form is

$$\frac{d^2y}{dx^2} + \left(a - 2q\cos 2x\right)y = 0 \tag{12}$$

with

$$a = \frac{8I}{\hbar^2}\left(E - I\omega^2\right) > 0 \text{ and } q = -4\left(\frac{I\omega}{\hbar}\right)^2 < 0 \tag{13}$$

The periodicity of the wave-function now demands $\Psi(2\phi+2\pi) = \Psi(2\phi) \Rightarrow \Psi\left(2\left(\phi+\pi\right)\right) = \Psi(2\phi)$ which means that the wave-function is periodic of period π as a function of φ. Such solutions are the Mathieu functions of even order [14], namely the $ce_{2m}(\phi)$ and $se_{2m+2}(\phi)$ for m=0,1,2,.. However, as was pointed out in [15], equation (11) corresponds to a Mathieu equation with negative q, and not positive as it should, and so we have to further perform a last change of variables of the form $\phi \to \frac{\pi}{2} - \theta$. The normalized solutions of the Schrödinger equation (11) are given by the following relations [15]:

$$\begin{aligned}\Psi_o^{(e)}(\theta,-q) &= \frac{1}{\sqrt{2\pi}}ce_o\left(\frac{\pi-\theta}{2},q\right) \\ \Psi_{2m}^{(e)}(\theta,-q) &= \frac{1}{\sqrt{\pi}}(-1)^m ce_{2m}\left(\frac{\pi-\theta}{2},q\right) \\ \Psi_{2m}^{(o)}(\theta,-q) &= \frac{1}{\sqrt{\pi}}(-1)^{m+1} se_{2m}\left(\frac{\pi-\theta}{2},q\right)\end{aligned} \tag{14}$$

where the superscripts *e* and *o* denote even and odd solutions in θ respectively. By assuming trigonometric expansions for the eigenfunctions we can compute both the eigenvalues (known as characteristic values) and the related eigenfunctions. However the calculations are rather tedious since they involve numerical calculations based on certain three-term recursion relations or appropriate definitions of quantities as continued fractions [15]. An initial guess of the wave function is also needed.

Things become extremely simple however, if we assume a parabolic potential ($\sin\theta \cong \theta$) to write (9) in the form

$$-\frac{\hbar^2}{2I}\frac{\partial^2\Psi(\theta)}{\partial\theta^2}+2I\omega^2\sin^2\left(\theta/2\right)\Psi(\theta)=E\Psi(\theta)$$

approximated as (15)

$$-\frac{\hbar^2}{2I}\frac{\partial^2\Psi(\theta)}{\partial\theta^2}+\frac{1}{2}I\omega^2\theta^2\Psi(\theta)=E\Psi(\theta)$$

The above correspond to the harmonic oscillator problem whose eigenfunctions are given in terms of the Hermite polynomials [16] while its eigenvalues take the simple form

$$E_n=\left(n+\frac{1}{2}\right)\hbar\omega \tag{16}$$

Since only the $\Delta n=\pm 1$ transitions are allowed [16], ω is equal to the corresponding photon frequency and is common for all the possible values of energy. This is the quantum equivalent to classical isochronism that was mentioned in the previous section. In this way we can calculate the classical frequency of swinging through the difference of two successive quantum energy levels.

THE SEMICLASSICAL PATH INTEGRAL APPROACH

The path integral formulation can be applied to the construction of a system's Green's function, the later being the Fourier transformed in time of the propagator [17]. In its semiclassical version which is based on a periodic orbit theory [18,19], the Green's function for one dimensional propagation is accomplished by taking account of all possible changes in phase of the wave-function. However this is not trivial at all since it encloses the contribution of all possible changes of phase when a turning point is met either in the allowed or in the forbidden regions of motion, for an infinite set of possible paths.

The above mentioned pioneering work inspired many others to further improve the method and analytically solve interesting one dimensional problems. Among those, Holstein and Swift [20] and Holstein alone [21] showed how *Gsc(E)* , which is the semiclassical fixed energy transmission amplitude, can be used to achieve analytic continuation of the propagator to forbidden regions, and furthermore established its connection to propagation and to reflection. For example when *Gsc(E)* is calculated for propagation between two points in a classically allowed region of motion, the perturbed complex energy spectrum of the later is revealed [22]. The central result for the calculation of the transmission amplitude via an infinite set of paths that the particle follows, can be written in compact form as

ISSN:1791-4469

$$Gsc(E) = \frac{\mathrm{m}}{2\pi\overline{k(r_1,r_2)}}\sum_{j=1}^{\infty}\left\{\prod_{i=1}^{N(j)}a_{ij}\right\} \qquad (17)$$

In the above equation $\overline{k(r_1,r_2)}$ is a non local wave number of the particle connecting the initial and the final point of propagation and defined by $\overline{k(r_1,r_2)}=\sqrt{k(r_1)k(r_2)}$ where $k(r)=\hbar^{-1}\sqrt{2m\left(E-V(r)\right)}$, with *E* standing for the energy and *V(r)* for the potential function. The index *j* corresponds to a particular path, while the index *i* corresponds to a certain event along the path. Therefore, the symbol a_{ij} represents each *i* event factor that contributes to the j^{th} path. Their total number is *N(j)* and depends on the path. These event factors are of two types. One type represents propagation and the other represents reflection. The a_{ij} propagation event factors describe propagation (from α to β) in an either allowed (given by $\exp[i\int_{\alpha}^{\beta}k(r)dr]\equiv e^{ik}$, where $k\neq k(r)$), or in a forbidden region of motion (given by $\exp\left[-\left|\int_{\alpha}^{\beta}\kappa(r)dr\right|\right]\equiv e^{-\kappa}$ with $\kappa(r)=\hbar^{-1}\sqrt{2m\left(V(r)-E\right)}$, where $\kappa\neq\kappa(r)$), while the a_{ij} reflection event factors describe reflections from turning points, (*-i* for reflection from a turning point in an allowed region, *+ i/2* for reflection in a forbidden region , and –1 for reflection from an infinite barrier). Both *k* and *κ* are dimensionless quantities and should not be confused with the corresponding position dependent wave numbers, although we have kept the same symbol in order to emphasize their origin. Thus the product $\prod_{i=1}^{N(j)}a_{ij}$ gives the unique amplitude for each possible path for going from r_1 to r_2 in constant energy. The sum in (17) extends to the infinity of possible paths connecting the space points r_1 and r_2. The above mentioned rules are depicted in the figure that follows.

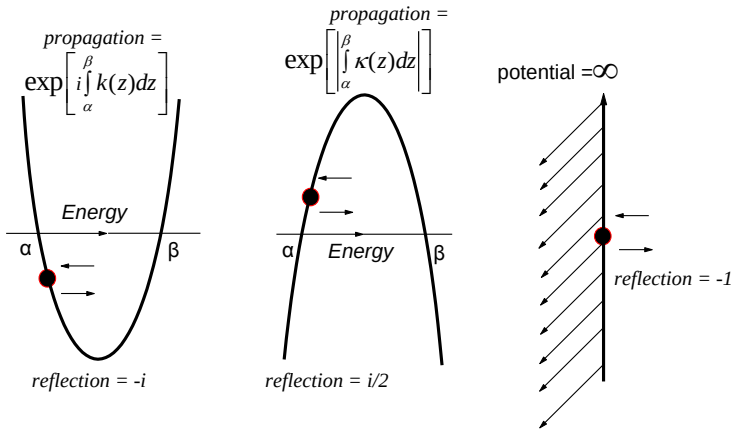


FIGURE 2. The rules for the construction of the path integral amplitudes through the α_{ij} event factors. The first motion is in a classically allowed region, the second in a classically forbidden region and the third case describes reflection from an infinite barrier. The direction before the reflection is always from right to left and is reversed after meeting the turning point.

The above described method can also be found in standard textbooks of path integrals, or quantum tunneling as well, [23,24].

It is readily seen that the calculation of the overall transmission amplitude depends on the topology of the potential function and on the nature of the turning points. Since there is an infinity of paths traversing both the allowed and the forbidden regions, it is very crucial to include all of them in the calculation by performing correct the rather complicated combinatorics. Then, and since each path repeats itself, the infinite class of paths can be summed to constitute geometric progressions, from which the analytic properties of *Gsc(E)* can be recognized directly. For the present requirement of computing the overall transmission amplitude, the points r_1 and r_2 are in the classically allowed region of motion of the pendulum, which is the area between angles –α and α shown in Figure 1.

THE CONSTRUCTION OF THE PENDULUM'S GREEN'S FUNCTION

Since the pendulum has only one degree of freedom it can be treated as an one dimensional physical system. Thus we can apply the formulation of the semiclassical Green's function that was described in the previous chapter, with only a few changes in notation. In particular the mass factor comes as the moment of inertia instead of the mass alone and we use the angle variable θ instead of the space variable r. In addition we should sketch the one dimensional potential as in Figure 3 that follows, in order to properly describe the hindered rotation of the pendulum:

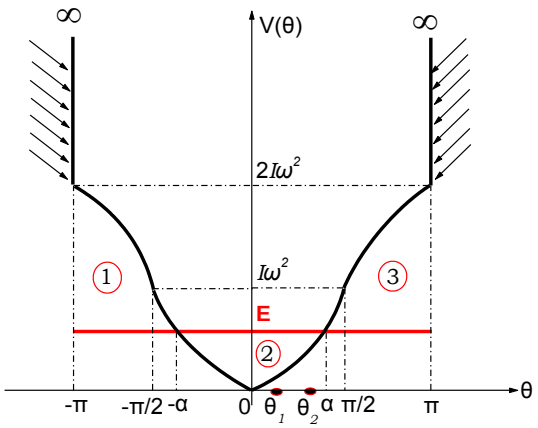


FIGURE 3. The potential of the quantum pendulum for hindered rotation. In order to exclude free rotation an infinite barrier is supposed at angle π (-π). There are three regions of motion: the classically accessible region 2, and the classically inaccessible regions 1 and 3, as these are structured by the specific value of the energy E and are separated from each other by the turning points α and –α.

ISSN:1791-4469

Figure 3 describes the induced topology of the potential as this is determined by the specific value of the energy. There are two turning points, meaning angles –α and α, while an infinite rotational barrier is supposed at angle π (-π), in order to forbid free rotation. For example we can imagine an elastic wall at this position. In this way we have three regions of motion, one classically allowed (region 2) and two classically forbidden (regions 1 and 3). In order to construct the overall transition amplitude for propagation between points θ_1 and θ_2 of region 2, we separate the problem in smaller ones. For this we write *Gsc(E)* as a sum of transition amplitudes involving specific regions of motion each time, of the form:

$$G_{sc}(E) = G^{2} + G^{2,1} + G^{2,3} + G^{2,1,3} \tag{18}$$

where $G^{\rho,\sigma}$ for example denotes the amplitude for propagation involving regions ρ and σ only, in all possible ways. In the paragraphs that follow we separately develop each amplitude providing the basic steps.

i) Amplitude G^2

It involves propagation inside region 2. It is constructed by fundamental amplitudes $\tilde{A}_{x,y}$ that connect points x and y in a straight path, and by amplitudes $A_{x,y}$, (we use the A letter for the classically accessible region of motion 2), that connect x and y with multiple repetitions (including reflections) in all possible ways, following the rules of the event factors that were given in the previous section. In this way we can write

$$\begin{aligned}
G^{2} &= \tilde{A}_{\theta_1,\theta_2} + \tilde{A}_{\theta_1,\alpha}\left\{A_{a,\alpha}\tilde{A}_{a,\theta_2} + A_{a,-\alpha}\tilde{A}_{-a,\theta_2}\right\} + \tilde{A}_{\theta_1,-\alpha}\left\{A_{-a,\alpha}\tilde{A}_{a,\theta_2} + A_{-a,-\alpha}\tilde{A}_{-a,\theta_2}\right\} \Rightarrow \\
G^{2} &= \tilde{A}_{\theta_1,\theta_2} + \tilde{A}_{\theta_1,\alpha}A_{a,\alpha}\left\{\tilde{A}_{a,\theta_2} - ie^{i\mu}\tilde{A}_{-a,\theta_2}\right\} + \tilde{A}_{\theta_1,-\alpha}A_{-a,-\alpha}\left\{\tilde{A}_{-a,\theta_2} - ie^{i\mu}\tilde{A}_{a,\theta_2}\right\} \Rightarrow \\
G^{2} &= \tilde{A}_{\theta_1,\theta_2} + A_{a,\alpha}\left\{\tilde{A}_{\theta_1,\alpha}\left(\tilde{A}_{a,\theta_2} - ie^{i\mu}\tilde{A}_{-a,\theta_2}\right) + \tilde{A}_{\theta_1,-\alpha}\left(\tilde{A}_{-a,\theta_2} - ie^{i\mu}\tilde{A}_{a,\theta_2}\right)\right\}
\end{aligned} \tag{19}$$

where we define the following quantities

$$\begin{aligned}
\tilde{A}_{x,y} &= \exp\left[i\frac{\sqrt{2\mathrm{I}}}{\hbar}\int_{x}^{y}\sqrt{E-V(\theta)}d\theta\right] \\
\tilde{A}_{-a,a} &= \exp\left[i\frac{\sqrt{2\mathrm{I}}}{\hbar}\int_{-a}^{a}\sqrt{E-V(\theta)}d\theta\right] \equiv e^{i\mu} \\
A_{-a,-a} &= A_{a,a} = -i + (-i)e^{i\mu}(-i)e^{i\mu}(-i) + (-i)e^{i\mu}(-i)e^{i\mu}(-i)e^{i\mu}(-i)e^{i\mu}(-i) + \ldots \Rightarrow \\
A_{-a,-a} &= A_{a,a} = -i\left[1 + (-ie^{i\mu})^{2} + (-ie^{i\mu})^{4} + \ldots\right] \Rightarrow \\
A_{-a,-a} &= A_{a,a} = -\frac{i}{1+e^{2i\mu}}
\end{aligned} \tag{20}$$

Thus the G^2 amplitude takes the following form

$$G^{2} = \tilde{A}_{\theta_1,\theta_2} - \frac{i}{1+e^{2i\mu}}\left\{\tilde{A}_{\theta_1,\alpha}\left(\tilde{A}_{a,\theta_2} - ie^{i\mu}\tilde{A}_{-a,\theta_2}\right) + \tilde{A}_{\theta_1,-\alpha}\left(\tilde{A}_{-a,\theta_2} - ie^{i\mu}\tilde{A}_{a,\theta_2}\right)\right\} \tag{21}$$

ii) Amplitudes $G^{2,1}$, $G^{2,3}$

It is quite obvious that the above amplitudes are equal to each other due to symmetry reasons. We use the letter B for the symbolism of the amplitudes in the forbidden region 1. In addition the star as a superscript denotes the fact that the amplitude includes at least one propagation event and not just a reflection factor. Since –α is the border between the two regions of motion the goal is to reach angle –α in all possible ways while staying in region 2, and then interchange the two regions in every possible way. According to these we can write

$$G^{2,1} = A_{\theta_1,-\alpha}\left\{B^{*}_{-a,-\alpha}A_{-a,-\alpha} + \left(B^{*}_{-a,-\alpha}A_{-a,-\alpha}\right)^2 + \ldots\right\}\left(\tilde{A}_{-a,\theta_2} + \tilde{A}_{-\alpha,\alpha}\tilde{A}_{a,\theta_2}\right) \tag{22}$$

which transforms to the following due to the geometric progression that appears

$$G^{2,1} = A_{\theta_1,-\alpha}\left\{\frac{B^{*}_{-a,-\alpha}A^{*}_{-a,-\alpha}}{1-B^{*}_{-a,-\alpha}A^{*}_{-a,-\alpha}}\right\}\left(\tilde{A}_{-a,\theta_2} + \tilde{A}_{-\alpha,\alpha}\tilde{A}_{a,\theta_2}\right) \tag{23}$$

and where we define the following quantities

$$\begin{aligned}
&A^{*}_{-a,-a} = e^{i\mu}(-i)e^{i\mu} + e^{i\mu}(-i)e^{i\mu}(-i)e^{i\mu}(-i)e^{i\mu} + \ldots \Rightarrow \\
&A^{*}_{-a,-a} = -\frac{ie^{2i\mu}}{1+e^{2i\mu}} \\
&\tilde{B}_{-\pi,-a} = \exp\left[-\frac{\sqrt{2\mathrm{I}}}{\hbar}\int_{-\pi}^{-a}\sqrt{V(\theta)-E}\,d\theta\right] \equiv e^{-\kappa} \\
&B^{*}_{-a,-a} = e^{-\kappa}(-1)e^{-\kappa} + e^{-\kappa}(-1)e^{-\kappa}(i/2)e^{-\kappa}(-1)e^{-\kappa} + \ldots \Rightarrow \\
&B^{*}_{-a,-a} = -\frac{2e^{-2\kappa}}{2+ie^{-2\kappa}}
\end{aligned} \tag{24}$$

Putting all these in equation (22) and performing the algebra we get the following result for the amplitude

$$G^{2,1} = \left\{\frac{2e^{4i\mu}e^{-2\kappa}\left(\tilde{A}_{-a,\theta_2} + \tilde{A}_{-\alpha,\alpha}\tilde{A}_{a,\theta_2}\right)\left(-ie^{i\mu}\tilde{A}_{\theta_1,\alpha} + \tilde{A}_{\theta_1,-\alpha}\right)}{\left(1+e^{2i\mu}\right)\left[\left(1+e^{2i\mu}\right)\left(2+ie^{-2\kappa}\right) - 2ie^{-2\kappa}e^{2i\mu}\right]}\right\} \tag{25}$$

The total amplitude for propagation in the classically allowed and one of the two classically forbidden regions of motion will be twice of the above for the reasons previously explained, and so we finally get

$$G^{2,(1or3)} = G^{2,1} + G^{2,3} = \left\{\frac{4e^{4i\mu}e^{-2\kappa}\left(\tilde{A}_{-a,\theta_2} + \tilde{A}_{-\alpha,\alpha}\tilde{A}_{a,\theta_2}\right)\left(-ie^{i\mu}\tilde{A}_{\theta_1,\alpha} + \tilde{A}_{\theta_1,-\alpha}\right)}{\left(1+e^{2i\mu}\right)\left[\left(1+e^{2i\mu}\right)\left(2+ie^{-2\kappa}\right) - 2ie^{-2\kappa}e^{2i\mu}\right]}\right\} \tag{26}$$

ISSN:1791-4469

iii) Amplitude $G^{2,1,3}$

It order to properly construct the amplitude for paths involving all regions of motion we think as follows: first of all we should reach angle –α in every possible way while constantly staying at region 2, which constitutes amplitude $A_{\theta_1,-\alpha}$, and then interchange regions 1 and 2 in every possible way starting and ending up to angle –α, this is the amplitude $X^{-\alpha}(1,2)$ to be defined. We reach angle α in a straight path and then interchange regions 3 and 2 in every possible way starting and ending up to angle α, this is amplitude $X^{\alpha}(3,2)$, and we infinitely repeat this process. We interchange for the last time regions 1 and 2, and finally reach angle θ_2 while remaining in region 2. Finally we should multiply the above overall amplitude with a factor of two, since we could equally well start with the interchange of regions 2 and 3, and proceed analogously. The above are mathematically described by the following equation

$$G^{2,1,3}=2A_{\theta_1,-\alpha}\left\{\begin{matrix}X^{-\alpha}(1,2)(-i)\tilde{A}_{-a,\alpha}X^{\alpha}(3,2)(-i)\tilde{A}_{a,-\alpha}+\\ \left[X^{-\alpha}(1,2)(-i)\tilde{A}_{-a,\alpha}X^{\alpha}(3,2)(-i)\tilde{A}_{a,-\alpha}\right]^2+..\end{matrix}\right\}X^{-\alpha}(1,2)(-i)A_{-a,\theta_2}\Rightarrow$$

$$G^{2,1,3}=2iA_{-a,\theta_2}A_{\theta_1,-\alpha}\frac{\left(X^{-\alpha}(1,2)\tilde{A}_{-a,\alpha}\right)^2}{1+\left(X^{-\alpha}(1,2)\tilde{A}_{-a,\alpha}\right)^2}X^{-\alpha}(1,2)\Rightarrow \quad (27)$$

$$G^{2,1,3}=2iA_{-a,\theta_2}A_{\theta_1,-\alpha}\frac{X^{-\alpha}(1,2)}{1+\left(X^{-\alpha}(1,2)\tilde{A}_{-a,\alpha}\right)^{-2}}$$

where we have used the fact that the two "doublet" amplitudes $X^{-\alpha}(1,2)$ and $X^{\alpha}(3,2)$ are equal to each other due to symmetry reasons. In fact each doublet is given by the following formula

$$X^{-\alpha}(1,2)=X^{\alpha}(3,2)=B^{*}_{-a,-\alpha}A_{-a,-\alpha}+\left(B^{*}_{-a,-\alpha}A_{-a,-\alpha}\right)^2+...\Rightarrow$$

$$X^{-\alpha}(1,2)=X^{\alpha}(3,2)=\frac{B^{*}_{-a,-\alpha}A^{*}_{-a,-\alpha}}{1-B^{*}_{-a,-\alpha}A^{*}_{-a,-\alpha}} \quad (28)$$

Putting equations (24) and (28) in (27) and performing the tedious algebra we get the following result for the amplitude $G^{2,1,3}$

$$G^{2,1,3}=\frac{2ie^{-6\kappa}e^{12i\mu}\left(\tilde{A}_{-a,\theta_2}-ie^{i\mu}\tilde{A}_{a,\theta_2}\right)\left(\tilde{A}_{\theta_1,-a}-ie^{i\mu}\tilde{A}_{\theta_1,\alpha}\right)}{\left(1+e^{2i\mu}\right)^2\left[(1+e^{2i\mu})(2+ie^{-2\kappa})-2ie^{-2\kappa}e^{2i\mu}\right]\left\{\left[(1+e^{2i\mu})(2+ie^{-2\kappa})-2ie^{-2\kappa}e^{2i\mu}\right]^2-4e^{-4\kappa}e^{6i\mu}\right\}} \quad (29)$$

CALCULATION OF THE QUANTUM CORRECTED CLASSICAL SWING PERIOD

The total propagation amplitude as this is calculated by the sum of equations (21), (26) and (29), is constructed by a regular (non pole) and three pole (fractional) terms. It is clear that the pole terms contain all the bits of interesting information for the problem under study, while the regular term provides the general background of propagation. In fact the major contribution to propagation is achieved when the denominator of a pole term approaches (if not equal) to zero. In this way the energy spectrum of the system is revealed, since it corresponds to the energy poles of the propagation amplitude, and the Green's function itself can be written as a sum over the resonance energy states that the pendulum potential can support, [19].

Amplitude G^2 involves propagation inside classically accessible region 2 only, and so its energy spectrum resembles the one of a bound problem, like the harmonic oscillator (H.O.S.). For narrow states, which are states close to the bottom of the well, we get a closer approximation to the H.O.S. In order to calculate the spectrum we seek for the zeros of $1+e^{2i\mu}$, which is the denominator of the pole part of the amplitude G^2. Thus, in this framework, energy poles E_n will appear whenever

$$1+e^{2i\mu}=0 \Rightarrow \mu = n\pi + \pi/2 \tag{30}$$

Let us calculate μ for the harmonic potential of the form $V(\theta)=\frac{1}{2}I\omega^2\theta^2$. According to equation (20) we have $\mu(E)=\frac{\sqrt{2I}}{\hbar}\int_{-\alpha}^{\alpha}\sqrt{E-\frac{1}{2}I\omega^2\theta^2}\,d\theta$. The integral is elementary and we get $\mu(E)=\frac{\pi E}{\hbar\omega}$.

Imposing the pole condition of equation (30) we get $E_n=\left(n+\frac{1}{2}\right)\hbar\omega$, and so we reveal the H.O.S.

The classical swing period can be extracted from the spectrum in two ways. First, it is given in terms of the derivative of $\mu(E)$ with respect to energy as

$$T=2\hbar\frac{\partial\mu(E)}{\partial E}\equiv 2\breve{\tau}(\mathrm{E}) \tag{31}$$

where we have defined the time scale $\breve{\tau}(\mathrm{E})$ for the classically accessible region of motion. The reason for eq. (31) to happen comes from the fact that according to the physics of path integration $\mu(E)$ is the action phase factor for propagation in the classically accessible region of motion with constant energy *E*, and so is equal to $E\cdot\tau/\hbar$, where τ is the time of propagation, [21]. Since τ is half the classical period, equation (31) follows. Second and since the energy spectrum depends only on quantum number *n*, it can be given in terms of the energy derivative with respect to quantum number n,

$$T=\frac{2\pi\hbar}{\partial E_n/\partial n} \tag{32}$$

where equation (32) comes from the expansion of the energy eigenvalues about the central value (when a wave packet is considered), in terms of which we can take the classical period, as well as the revival, and superrevival times of the wave packet, [25]. Equations (31) and (32) become

ISSN:1791-4469

equivalent to each other under the action of condition (30), since then μ becomes a function of the quantum number n, and so

$$(\breve{\tau}(\mathrm{E}) \equiv)\hbar\frac{\partial\mu(E)}{\partial E} = \hbar\frac{\partial\mu(E)}{\partial n}\frac{\partial n}{\partial E} = \pi\hbar\frac{\partial n}{\partial E} \tag{33}$$

For the case under study equation (32) is simply transformed to

$$T(E_n) = \frac{2\pi\hbar}{E_{n+1} - E_n} \tag{34}$$

As far as the harmonic potential is concerned, equations (31) and (32) give the classical result, meaning equation (3) for the classical swing period, where the quantum effects are absent. For the case of the real trigonometric potential, equation (31) can be written as

$$T = 2\hbar\frac{\partial\mu(E)}{\partial E} = 2\hbar\left[\frac{\sqrt{2\mathrm{I}}}{\hbar}\frac{1}{2\sqrt{mgl}}\int_{-a}^{a}\frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}}\right] = \left[2\sqrt{\frac{l}{2g}}\int_{-a}^{a}\frac{d\theta}{\sqrt{\cos\theta - \cos\alpha}}\right] \tag{35}$$

which of course coincides with the classical result of eq. (4) and (5), and gives the swing period in terms of the elliptic integral. Again the quantum effects are absent and this is due to the fact that the amplitude G^2 is constructed by taking in account only the classically accessible region of motion. However the spectrum is no more continuous since condition (30) quantizies the spectrum, in a way equivalent to the W.K.B. approximation, [16].

In order to calculate the classical swing period as this is perturbed from its classical value due to the quantum tunneling effect, we should look at the pole terms of the amplitudes $G^{2,(1or3)}$ and $G^{2,1,3}$ that emerge through propagation in the classically inaccessible regions 1 and 3. Both of these maximally increase their magnitude when their denominator approaches zero. Besides equation (30) this happens when the following condition is fulfilled

$$(1+e^{2i\mu})(2+ie^{-2\kappa}) - 2ie^{-2\kappa}e^{2i\mu} = 0 \tag{36}$$

However, equation (36) has not real solutions, meaning that there do not exist real values of the energy to satisfy it. Condition (30) makes (36) almost zero, besides the remaining term $2ie^{-2\kappa}$, which is exponentially small in magnitude. If we expand equation (36) in first order about the energy poles E_n of the classical accessible region 2, we get

$$Z_n \cong E_n + \frac{\hbar e^{-2\kappa}\left(\breve{\tau}(E_n) + e^{-2\kappa}\hat{\tau}(E_n)\right)/2}{\breve{\tau}^2(E_n)e^{-4\kappa} + \left(\breve{\tau}(E_n) + e^{-2\kappa}\hat{\tau}(E_n)\right)^2} + \frac{i}{2}\frac{\hbar e^{-4\kappa}\breve{\tau}(E_n)}{\breve{\tau}^2(E_n)e^{-4\kappa} + \left(\breve{\tau}(E_n) + e^{-2\kappa}\hat{\tau}(E_n)\right)^2} \tag{37}$$

where we have defined time scale $\hat{\tau}(\mathrm{E})$ for the classically inaccessible region of motion in an analogous to equation (31) way, meaning

$$\hat{\tau}(\mathrm{E}) \equiv \hbar\frac{\partial\kappa(E)}{\partial E} \tag{38}$$

The perturbed energy poles Z_n have been calculated by taking in account the tunneling effect in the classically forbidden regions 1 and 3. In this way and according to our previous analysis, the difference between the real part of two successive perturbed energy poles, will provide the quantum corrected swing frequency. Writing the real part of the perturbed energy poles as a function of the unperturbed energy poles, meaning

$$\mathrm{Re}\, Z_n = E_n + f\left(E_n\right) \tag{39}$$

and introducing the perturbed swing frequency Ω as

$$\mathrm{Re}\, Z_{n+1} - \mathrm{Re}\, Z_n \equiv \hbar\Omega \tag{40}$$

we get the following equation

$$\mathrm{Re}\, Z_{n+1} - \mathrm{Re}\, Z_n = E_{n+1} - E_n + \left(f(E_{n+1}) - f(E_n) \right) \Rightarrow$$
$$\hbar\Omega = \hbar\omega(E_n)\left(1 + \frac{\partial f(E_n)}{\partial E_n} \right) \tag{41}$$

By keeping only the dominant terms in equation (37), equation (41) gives the following result for the quantum shift of the frequency

$$\Delta\omega = -\frac{\omega(E_n)}{T(E_n)} 2e^{-2\kappa}\left(\hat{\tau}(E_n) + \hbar \frac{\partial\left(\ln T(E_n)\right)}{\partial E_n} \right) \tag{42}$$

Thus, the quantum corrected period of swing can be easily calculated. Using the index *quan* for quantum and *cl* for classical in order to emphasize our result, we can write the following formula

$$T_{quan}(E_n) = T_{cl}(E_n) \exp\left\{ \frac{2e^{-2\kappa(E_n)}}{T_{cl}(E_n)} \left(2\hat{\tau}(E_n) + \hbar \frac{\partial\left(\ln T_{cl}(E_n)\right)}{\partial E_n} \right) \right\} \tag{43}$$

which is the central result of this work. Due to the presence of both the exponentially small term $e^{-2\kappa(E_n)}$ and Planck's constant, quantum and classical period are nearly equal in magnitude. Quantum corrections, as these are contained in the above relation, come from three different sources. First the existence of the quantum tunneling effect through the exponential $e^{-2\kappa(E_n)}$ term, second the dependence of the tunneling effect on energy through the induced time scale $\hat{\tau}(E_n)$, and third the dependence of the classical period on energy (present in the classical analysis as well), through its logarithmic derivative $\frac{\partial\left(\ln T_{cl}(E_n)\right)}{\partial E_n}$.

Let us look in some detail the physical meaning of the tunneling time scale $\hat{\tau}(E_n)$. The energy dependence of the classical action inside the potential barrier, meaning quantity $\hbar\kappa(E) = \sqrt{2I}\int_{-\pi}^{-\alpha(E)} d\theta\sqrt{V(\theta) - E}$, comes not only from the integrand but from the upper limit of integration as well. The action phase factor may be written as

$$\kappa(E) = \frac{\sqrt{2I}}{\hbar}\int_{-\pi}^{-\alpha(E)} T(\theta, E) d\theta =$$
$$\frac{\sqrt{2I}}{\hbar}\left\{ \Phi[-\alpha(E), E] - \Phi[-\pi, E] \right\} \tag{44}$$

where of course

$$\frac{\partial\Phi(\theta, E)}{\partial\theta} = T(\theta, E) \tag{45}$$

However, and since the upper limit of integration is the classical turning point, the following result holds,

$$T\left[-\alpha(E), E\right] = 0 \tag{46}$$

which leads to the

$$\frac{d\kappa(E)}{dE} = -\sqrt{\frac{I}{2\hbar^2}} \int_{-\pi}^{-\alpha(E)} d\theta \frac{1}{\sqrt{V(\theta)-E}} \tag{47}$$

which can be further transformed to the following

$$\frac{d\kappa(E)}{dE} = \sqrt{\frac{I}{2\hbar^2}} \left\{ \int_{E}^{2I\omega^2} \frac{d\theta}{dV} \frac{1}{\sqrt{V-E}} dV \right\} \tag{48}$$

In order to eliminate E from the right term we introduce the weight function $(E-\lambda)^{-1/2}$ and perform integration over E in both terms. It is clear that the parameter λ has energy dimensions. Equation (48) is then transformed to:

$$\int_{\lambda}^{2I\omega^2} dE \frac{d\kappa(E)}{dE} \frac{1}{\sqrt{E-\lambda}} = \sqrt{\frac{I}{2\hbar^2}} \int_{\lambda}^{2I\omega^2} dE \int_{E}^{2I\omega^2} \left(\frac{d\theta}{dV}\right) \frac{1}{\sqrt{V-E}} \frac{1}{\sqrt{E-\lambda}} dV \tag{49}$$

The double integral can be easily handled via a change in the order of integration to take the

$$\int_{\lambda}^{2I\omega^2} dE \frac{d\kappa(E)}{dE} \frac{1}{\sqrt{E-\lambda}} = \sqrt{\frac{\mu}{2\hbar^2}} \int_{\lambda}^{2I\omega^2} dV \int_{\lambda}^{V} \left(\frac{d\theta}{dV}\right) \frac{1}{\sqrt{V-E}\sqrt{E-\lambda}} dE \tag{50}$$

The integral over E is then elementary and equal to π, and so we take:

$$\int_{\lambda}^{2I\omega^2} dE \frac{d\kappa(E)}{dE} \frac{1}{\sqrt{E-\lambda}} = \pi \sqrt{\frac{I}{2\hbar^2}} \int_{\lambda}^{2I\omega^2} dV \left(\frac{d\theta}{dV}\right) \tag{51}$$

while the last integral is trivial and gives

$$\begin{aligned} \int_{\lambda}^{2I\omega^2} dE \frac{d\kappa(E)}{dE} \frac{1}{\sqrt{E-\lambda}} &= \pi \sqrt{\frac{I}{2\hbar^2}} \{\pi - \theta(\lambda)\} \Rightarrow \\ \int_{\lambda}^{2I\omega^2} \hat{\tau}(\mathrm{E}) \frac{dE}{\sqrt{E-\lambda}} &= \pi \sqrt{\frac{I}{2}} \{\pi - \theta(\lambda)\} \end{aligned} \tag{52}$$

Since $\hat{\tau}(\mathrm{E})$ is defined only inside the classically inaccessible region (1 or 3), it is clear that $\lambda \geq V(\alpha)$. Equation (52) then, describes the way that the knowledge of the function $\hat{\tau}(\mathrm{E})$ permits the knowledge of the angular position inside the potential barrier as a function of the potential. In addition $\hat{\tau}(\mathrm{E})$ appears in both the real and imaginary energy shift of the perturbed energy poles in equation (37). In a previous work of us [26], the later was connected with what was defined as energy indeterminacies during the tunneling process.

CONCLUDING REMARKS

In the present work we derived an improved and prototype formula for the swing period of the classical pendulum, by taking in account the pure quantum mechanical tunneling effect via semi-classical path integration. Our central result is contained in equation (43) where the tunneling effect is present through the action phase factor $\kappa(E)$ and its derivative with respect to energy. Both the $\mu(E)$ and $\kappa(E)$ action phase factors in eq. (43), defined in the classically accessible and classically forbidden regions of motion respectively, introduce a different time scale of the physical system under study. The first order energy derivative of $\mu(E)$ reproduces the classical formula for the swing period in terms of the elliptic integral, while its second order derivative reflects the quantum equivalent to classical isochronism. On the other hand the energy derivative of the $\kappa(E)$ action phase factor introduces the time scale $\hat{\tau}(E_n)$ whose knowledge is equivalent to the knowledge of the position of the particle during tunneling, as a function of the potential.

REFERENCES

1. G. L. Baker and J. A. Blackburn, The Pendulum: A Case Study in Physics, Oxford: Oxford University Press, 2005, pp. 2-7.

2. J. Milne, Nature 37, 570-571 (1988).

3. G. F. West and D. J. Dunlop, Journal of Physics E: Scientific Instruments 4, 37-40 (1971).

4. C. H. Bagley and G. G. Luther, Physical Review Letters 78, 3047-3050 (1997).

5. M. K. Bantel and R. D. Newman, Classical and Quantum Gravity, 17, 2313-2318 (2000).

6. G. D'Anna, Europhysics Letters, 51, 293-299 (2000).

7. E. U. Condon, Physical Review, 31, 891-894 (1928).

8. G. L. Baker and J. A. Blackburn and H. J. T. Smith, American Journal of Physics, 70, 525-531 (2002).

9. D. B. Sullivan and J. E. Zimmerman, American Journal of Physics, 39, 1504-1517 (1971).

10. L. D. Landau and E. M. Lifshitz, Mechanics, Oxford: Butterworth Heinemann, 1981, 3rd ed., pp. 26.

11. M. Abramowitz and I. Stegun, Handbook of Mathematical Functions, 9th edn, New York: Dover, 1980, pp. 590.

12. Q. X. Yuan and P. Ding P, European Journal of Physics 30, L79-82, (2009).

13. T. G. Douvropoulos, European Journal of Physics 33, 207-217, (2012).

ISSN:1791-4469

14. I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Series and Products, London: Academic Press, 1994, pp. 1006.

15. R. Aldrovaldi and P. L. Ferreira, American Journal of Physics 48, 660-664, (1980).

16. L. D. Landau and E. M. Lifshitz, Quantum Mechanics, Oxford: Butterworth Heinemann, 1977, 3rd edition, pp. 67.

17. R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals, New York: Wiley, 1961, pp. 111.

18. W. H. Miller, Journal of Chemical Physics 83, 960-963, (1979).

19. M. C. Gutzwiller, Journal of Mathematical Physics 12, 343-358, (1971).

20. B.R. Holstein and A. R. Swift, American Journal of Physics 50, 833-839, (1982).

21. B. R. Holstein, American Journal of Physics 51, 897-901, (1983).

22. T. G. Douvropoulos and C. A. Nicolaides, Journal of Physics B 35, 4453-4473, (2002).

23. H. Kleinert, Path integrals in Quantum Mechanics, Statistics and Polymer Physics, Singapore: World Scientific, 1995, 2nd edition, pp. 215-222.

24. M. Razavy, Quantum Theory of Tunneling, Singapore: World Scientific, 2003, pp. 219-240.

25. D. J. Tannor, Introduction to Quantum Mechanics: A Time Dependent Perspective, California: University Science Books, 2007, pp 81-84.

26. T. G. Douvropoulos, "Energy Indeterminacies and Inequalities During Tunneling", in Energy Conservation: New Research, edited by G. Spadoni, New York: Nova Publishers, 2009, pp. 103-122.

Infinitesimal symmetry transformations of matrix-valued differential equations: An algebraic approach

C. J. Papachristou

Department of Physical Sciences, Hellenic Naval Academy, Piraeus 18539, Greece
E-mail: papachristou@snd.edu.gr

Abstract. The study of symmetries of partial differential equations (PDEs) has been traditionally treated as a geometrical problem. Although geometrical methods have been proven effective with regard to finding infinitesimal symmetry transformations, they present certain conceptual difficulties in the case of matrix-valued PDEs; for example, the usual differential-operator representation of the symmetry-generating vector fields is not possible in this case. An algebraic approach to the symmetry problem of PDEs is described, based on abstract operators (characteristic derivatives) which admit a standard differential-operator representation in the case of scalar-valued PDEs.

Keywords: Matrix-valued differential equations, symmetry transformations, Lie algebras, recursion operators

PACS: 02.30.Jr, 02.10.Yn, 02.20.Sv

1. Introduction

The problem of symmetries of a system of partial differential equations (PDEs) has been traditionally treated as a geometrical problem in the jet space of the independent and the dependent variables (including a sufficient number of partial derivatives of the latter variables with respect to the former ones). Two more or less equivalent approaches have been followed: (*a*) invariance of the system of PDEs itself, under infinitesimal transformations generated by corresponding vector fields in the jet space [1]; (*b*) invariance of a differential ideal of differential forms representing the system of PDEs, under the Lie derivative with respect to the vector fields representing the symmetry transformations [2-6].

Although effective with regard to calculating symmetries, these geometrical approaches suffer from a certain drawback of conceptual nature when it comes to matrix-valued PDEs. The problem is related to the inevitably mixed nature of the coordinates in the jet space (scalar independent variables versus matrix-valued dependent ones) and the need for a differential-operator representation of the symmetry vector fields. How does one define differentiation with respect to matrix-valued variables? Moreover, how does one calculate the Lie bracket of two differential operators in which some (or all) of the variables, as well as the coefficients of partial derivatives with respect to these variables, are matrices?

Although these difficulties were handled in some way in [4-6], it was eventually realized that an alternative, purely algebraic approach to the symmetry problem would be more appropriate in the case of matrix-valued PDEs. Elements of this approach were presented in [7] and later

ISSN:1791-4469

applied in particular problems [8-10]; no formal theoretical framework was fully developed, however.

An attempt to develop such a framework is made in this article. In Sections 2 and 3 we introduce the concept of the *characteristic derivative* – an abstract operator generalization of the Lie derivative used in scalar-valued problems – and we demonstrate the Lie-algebraic nature of the set of these derivatives.

The general symmetry problem for matrix-valued PDEs is presented in Sec. 4, and the Lie-algebraic property of symmetries of a PDE is proven in Sec. 5. In Sec. 6 we discuss the concept of a *recursion operator* [1,8-14] by which an infinite set of symmetries may in principle be produced from any known symmetry.

Finally, an application of these ideas is made in Sec. 7 by using the chiral field equation as an example.

To simplify our formalism, we restrict our analysis to the case of a single matrix-valued PDE in one dependent variable. Generalization to systems of PDEs is straightforward and is left to the reader (see, e.g., [1] for scalar-valued problems).

2. Total and characteristic derivative operators

A PDE for the unknown function $u=u(x^1, x^2, \dots) \equiv u(x^k)$ [where by (x^k) we collectively denote the independent variables $x^1, x^2, \dots$] is an expression of the form $F[u]=0$, where $F[u] \equiv F(x^k, u, u_k, u_{kl}, \dots)$ is a function in the *jet space* [1] of the independent variables (x^k), the dependent variable u, and the partial derivatives of various orders of u with respect to the x^k, which derivatives will be denoted by using subscripts: u_k, u_{kl}, u_{klm}, etc. A *solution* of the PDE is any function $u=\varphi(x^k)$ for which the relation $F[u]=0$ is satisfied at each point (x^k) in the domain of φ.

We assume that u, as well as all functions $F[u]$ in the jet space, are square-matrix-valued of fixed matrix dimensions. In particular, we require that, in its most general form, a function $F[u]$ in the jet space is expressible as a finite or an infinite sum of products of alternating x-dependent and u-dependent terms, of the form

$$F[u] = \sum a(x^k)\Pi[u]\, b(x^k)\Pi'[u]\, c(x^k)\cdots \tag{2.1}$$

where the $a(x^k)$, $b(x^k)$, $c(x^k)$, etc., are matrix-valued, and where the matrices $\Pi[u]$, $\Pi'[u]$, etc., are products of variables u, u_k, u_{kl}, etc., of the "fiber" space (or, more generally, products of powers of these variables). The set of all functions (2.1) is thus a (generally) non-commutative algebra.

If u is a scalar quantity, a total derivative operator can be defined in the usual way [1] as

$$D_i = \frac{\partial}{\partial x^i} + u_i \frac{\partial}{\partial u} + u_{ij} \frac{\partial}{\partial u_j} + u_{ijk} \frac{\partial}{\partial u_{jk}} + \cdots \tag{2.2}$$

where the summation convention over repeated up-and-down indices (such as j and k in this equation) has been adopted and will be used throughout. If, however, u is matrix-valued, the above expression is obviously not valid. A generalization of the definition of the total derivative is thus necessary for matrix-valued PDEs.

Definition 2.1. The *total derivative operator* with respect to the variable x^i is a linear operator D_i acting on functions $F[u]$ of the form (2.1) in the jet space and having the following properties:

1. On functions $f(x^k)$ in the base space, $D_i\, f(x^k) = \partial f / \partial x^i \equiv \partial_i\, f(x^k)$.

2. For $F[u]=u$, u_i, u_{ij}, etc., we have: $D_i\, u = u_i$, $D_i\, u_j = D_j\, u_i = u_{ij} = u_{ji}$, etc.

3. The operator D_i is a *derivation* on the algebra of all matrix-valued functions of the form (2.1) in the jet space; i.e., the *Leibniz rule* is satisfied:

$$D_i\big(F[u]G[u]\big) = \big(D_iF[u]\big)G[u] + F[u]D_iG[u] \tag{2.3}$$

Higher-order total derivatives $D_{ij}=D_iD_j$ may similarly be defined but they no longer possess the derivation property. Given that $\partial_i\partial_j=\partial_j\partial_i$ and that $u_{ij}=u_{ji}$, it follows that $D_iD_j = D_jD_i \Leftrightarrow D_{ij} = D_{ji}$; that is, total derivatives commute. We write: $[D_i, D_j]=0$, where, in general, $[A, B] \equiv AB–BA$ will denote the *commutator* of two operators or two matrices, as the case may be.

If u^{-1} is the inverse of u, such that $uu^{-1}= u^{-1}u = \mathbf{1}$, then we can define

$$D_i\left(u^{-1}\right) \equiv -u^{-1}\left(D_iu\right)u^{-1} \tag{2.4}$$

Moreover, for any functions $A[u]$ and $B[u]$ in the jet space, it can be shown that

$$D_i[A, B] = \big[D_iA, B\big] + \big[A, D_iB\big] \tag{2.5}$$

As an example, let $(x^1, x^2) \equiv (x, t)$ and let $F[u]=xtu_x^{\,2}$, where u is matrix-valued. Writing $F[u]=xtu_xu_x$, we have: $D_t\,F[u]=xu_x^{\,2} + xt\,(u_{xt}\,u_x + u_x\,u_{xt})$.

Let now $Q[u] \equiv Q\,(x^k, u, u_k, u_{kl}, ...)$ be a function in the jet space. We will call this a *characteristic function* (or simply a *characteristic*) of a certain derivative, defined as follows:

Definition 2.2. The *characteristic derivative* with respect to $Q[u]$ is a linear operator Δ_Q acting on functions $F[u]$ in the jet space and having the following properties:

1. On functions $f(x^k)$ in the base space,

$$\Delta_Q\, f(x^k) = 0 \tag{2.6}$$

(that is, Δ_Q acts only in the fiber space).

2. For $F[u]=u$,

$$\Delta_Q u = Q[u] \tag{2.7}$$

3. Δ_Q commutes with total derivatives:

$$\Delta_QD_i = D_i\Delta_Q \;\Leftrightarrow\; [\Delta_Q, D_i] = 0 \quad (\text{all } i) \tag{2.8}$$

ISSN:1791-4469

4. The operator Δ_Q is a *derivation* on the algebra of all matrix-valued functions of the form (2.1) in the jet space (the Leibniz rule is satisfied):

$$\Delta_Q\left(F[u]G[u]\right)=\left(\Delta_Q F[u]\right)G[u]+F[u]\Delta_Q G[u] \tag{2.9}$$

Corollary: By (2.7) and (2.8) we have:

$$\Delta_Q u_i=\Delta_Q D_i u=D_i Q[u] \tag{2.10}$$

We note that the operator Δ_Q is a well-defined quantity, in the sense that the action of Δ_Q on u uniquely determines the action of Δ_Q on any function $F[u]$ of the form (2.1) in the jet space. Moreover, since, by assumption, u and $Q[u]$ are matrices of equal dimensions, it follows from (2.7) that Δ_Q preserves the matrix character of u, as well as of any function $F[u]$ on which this operator acts.

We also remark that we have imposed conditions (2.6) and (2.8) having a certain property of symmetries of PDEs in mind; namely, *every* symmetry of a PDE can be represented by a transformation of the dependent variable u alone, i.e., can be expressed as a transformation in the fiber space (see [1], Chap. 5).

The following formulas, analogous to (2.4) and (2.5), may be written:

$$\Delta_Q\left(u^{-1}\right)\equiv -u^{-1}\left(\Delta_Q u\right)u^{-1} \tag{2.11}$$

$$\Delta_Q[A,B]=\left[\Delta_Q A,\,B\right]+\left[A,\,\Delta_Q B\right] \tag{2.12}$$

As an example, let $(x^1, x^2)\equiv(x, t)$ and let $F[u]=a(x,t)u^2b(x,t)+[u_x\,,\,u_t]$, where a, b and u are matrices of equal dimensions. Writing $u^2=uu$ and using properties (2.7), (2.9), (2.10) and (2.12), we find: $\Delta_Q F[u]=a(x,t)(Qu+uQ)b(x,t)+[D_x Q,\,u_t]+[u_x\,,\,D_t Q]$.

In the case where u is scalar-valued (thus so is $Q[u]$), the characteristic derivative Δ_Q admits a differential-operator representation of the form

$$\Delta_Q=Q[u]\frac{\partial}{\partial u}+\left(D_i Q[u]\right)\frac{\partial}{\partial u_i}+\left(D_i D_j Q[u]\right)\frac{\partial}{\partial u_{ij}}+\cdots \tag{2.13}$$

(See [1], Chap. 5, for an analytic proof of property (2.8) in this case.)

3. The Lie algebra of characteristic derivatives

The characteristic derivatives Δ_Q acting on functions $F[u]$ of the form (2.1) in the jet space constitute a *Lie algebra of derivations* on the algebra of the $F[u]$. The proof of this statement is contained in the following three Propositions.

Proposition 3.1. Let Δ_Q be a characteristic derivative with respect to the characteristic $Q[u]$; i.e., $\Delta_Q u=Q[u]$ [cf. Eq. (2.7)]. Let λ be a constant (real or complex). We define the operator $\lambda\Delta_Q$ by the relation

$$(\lambda\Delta_Q)F[u] \equiv \lambda\,(\Delta_Q F[u])\,.$$

Then, $\lambda\Delta_Q$ is a characteristic derivative with characteristic $\lambda Q[u]$. That is,

$$\lambda\,\Delta_Q = \Delta_{\lambda Q} \tag{3.1}$$

Proof. (*a*) The operator $\lambda\Delta_Q$ is linear, since so is Δ_Q .

(*b*) For $F[u]=u$, $(\lambda\Delta_Q)u= \lambda(\Delta_Q u)=\lambda Q[u]$.

(*c*) $\lambda\Delta_Q$ commutes with total derivatives D_i , since so does Δ_Q .

(*d*) Given that Δ_Q satisfies the Leibniz rule (2.9), it is easily shown that so does $\lambda\Delta_Q$.

Comment: Condition (*c*) would not be satisfied if we allowed λ to be a function of the x^k, instead of being a constant, since $\lambda(x^k)$ generally does not commute with the D_i. Therefore, relation (3.1) is not valid for a non-constant λ.

Proposition 3.2. Let Δ_1 and Δ_2 be characteristic derivatives with respect to the characteristics $Q_1[u]$ and $Q_2[u]$, respectively; i.e., $\Delta_1 u=Q_1[u]$, $\Delta_2 u=Q_2[u]$. We define the operator $\Delta_1+\Delta_2$ by

$$(\Delta_1+\Delta_2)\,F[u] \equiv \Delta_1\,F[u] + \Delta_2\,F[u]\,.$$

Then, $\Delta_1+\Delta_2$ is a characteristic derivative with characteristic $Q_1[u]+Q_2[u]$. That is,

$$\Delta_1 + \Delta_2 = \Delta_Q \quad \text{with} \quad Q[u] = Q_1[u] + Q_2[u] \tag{3.2}$$

Proof. (*a*) The operator $\Delta_1+\Delta_2$ is linear, as a sum of linear operators.

(*b*) For $F[u]=u$, $(\Delta_1+\Delta_2)u = \Delta_1 u+\Delta_2 u = Q_1[u]+Q_2[u]$.

(*c*) $\Delta_1+\Delta_2$ commutes with total derivatives D_i , since so do Δ_1 and Δ_2 .

(*d*) Given that each of Δ_1 and Δ_2 satisfies the Leibniz rule (2.9), it is not hard to show that the same is true for $\Delta_1+\Delta_2$.

Proposition 3.3. Let Δ_1 and Δ_2 be characteristic derivatives with respect to the characteristics $Q_1[u]$ and $Q_2[u]$, respectively; i.e., $\Delta_1 u=Q_1[u]$, $\Delta_2 u=Q_2[u]$. We define the operator $[\Delta_1\,,\,\Delta_2]$ (*Lie bracket* of Δ_1 and Δ_2) by

$$[\Delta_1\,,\,\Delta_2]\,F[u] \equiv \Delta_1\,(\Delta_2 F[u]) - \Delta_2\,(\Delta_1 F[u])\,.$$

Then, $[\Delta_1\,,\,\Delta_2]$ is a characteristic derivative with characteristic $\Delta_1 Q_2[u]–\Delta_2 Q_1[u]$. That is,

$$[\Delta_1,\,\Delta_2] = \Delta_Q \quad \text{with} \quad Q[u] = \Delta_1 Q_2[u] - \Delta_2\,Q_1[u] \equiv Q_{1,2}[u] \tag{3.3}$$

ISSN:1791-4469

Proof. (*a*) The linearity of $[\Delta_1, \Delta_2]$ follows from the linearity of Δ_1 and Δ_2.

(*b*) For $F[u]=u$, $[\Delta_1, \Delta_2]u = \Delta_1(\Delta_2 u) - \Delta_2(\Delta_1 u) = \Delta_1 Q_2[u] - \Delta_2 Q_1[u] \equiv Q_{1,2}[u]$.

(*c*) $[\Delta_1, \Delta_2]$ commutes with total derivatives D_i, since so do Δ_1 and Δ_2 .

(*d*) Given that each of Δ_1 and Δ_2 satisfies the Leibniz rule (2.9), one can show (after some algebra and cancellation of terms) that the same is true for $[\Delta_1, \Delta_2]$.

In the case where u (thus the Q's also) is scalar-valued, the Lie bracket admits a standard differential-operator representation [1]:

$$[\Delta_1, \Delta_2] = Q_{1,2}[u]\frac{\partial}{\partial u} + \left(D_i Q_{1,2}\right)\frac{\partial}{\partial u_i} + \left(D_i D_j Q_{1,2}\right)\frac{\partial}{\partial u_{ij}} + \cdots \tag{3.4}$$

where $Q_{1,2}[u] = [\Delta_1, \Delta_2]u = \Delta_1 Q_2[u] - \Delta_2 Q_1[u]$.

The Lie bracket $[\Delta_1, \Delta_2]$ has the following properties:

1. $[\Delta_1, a\Delta_2 + b\Delta_3] = a[\Delta_1, \Delta_2] + b[\Delta_1, \Delta_3]$;
   $[a\Delta_1 + b\Delta_2, \Delta_3] = a[\Delta_1, \Delta_3] + b[\Delta_2, \Delta_3]$ (a, b = const.)
2. $[\Delta_1, \Delta_2] = -[\Delta_2, \Delta_1]$ (*antisymmetry*)
3. $[\Delta_1, [\Delta_2, \Delta_3]] + [\Delta_2, [\Delta_3, \Delta_1]] + [\Delta_3, [\Delta_1, \Delta_2]] = 0$;
   $[[\Delta_1, \Delta_2], \Delta_3] + [[\Delta_2, \Delta_3], \Delta_1] + [[\Delta_3, \Delta_1], \Delta_2] = 0$ (*Jacobi identity*)

4. The symmetry problem for PDEs

Let $F[u]=0$ be a PDE in the independent variables $x^k \equiv x^1, x^2, \ldots$, and the (generally) matrix-valued dependent variable u. A transformation $u(x^k) \to u'(x^k)$ from the function u to a new function u' represents a *symmetry* of the PDE if the following condition is satisfied: $u'(x^k)$ is a solution of $F[u]=0$ *when* $u(x^k)$ is a solution; that is, $F[u']=0$ when $F[u]=0$.

We will restrict our attention to *continuous symmetries*, which can be expressed as infinitesimal transformations. Although such symmetries may involve transformations of the independent variables (x^k), they may equivalently be expressed as transformations of u alone (see [1], Chap. 5), i.e., as transformations in the fiber space.

An infinitesimal symmetry transformation is written symbolically as

$$u \to u' = u + \delta u$$

where δu is an infinitesimal quantity, in the sense that all powers $(\delta u)^n$ with $n>1$ may be neglected. The *symmetry condition* is thus written

$$F[u+\delta u] = 0 \quad \text{when} \quad F[u] = 0 \tag{4.1}$$

An infinitesimal change δu of u induces a change $\delta F[u]$ of $F[u]$, where

$$\delta F[u] = F[u+\delta u] - F[u] \quad \Leftrightarrow \quad F[u+\delta u] = F[u] + \delta F[u] \tag{4.2}$$

Now, if δu is an infinitesimal symmetry and if u is a solution of $F[u]=0$, then $u+\delta u$ also is a solution; that is, $F[u+\delta u]=0$. This means that $\delta F[u]=0$ when $F[u]=0$. The symmetry condition (4.1) is thus written as follows:

$$\delta F[u] = 0 \mod F[u] \tag{4.3}$$

A symmetry transformation (we denote it M) of the PDE $F[u]=0$ produces a one-parameter family of solutions of the PDE from any given solution $u(x^k)$. We express this by writing

$$M: u(x^k) \to \bar{u}(x^k;\alpha) \quad \text{with} \quad \bar{u}(x^k;0) = u(x^k) \tag{4.4}$$

For infinitesimal values of the parameter α,

$$\bar{u}(x^k;\alpha) \simeq u(x^k) + \alpha Q[u] \quad \text{where} \quad Q[u] = \left.\frac{d\bar{u}}{d\alpha}\right|_{\alpha=0} \tag{4.5}$$

The function $Q[u] \equiv Q(x^k, u, u_k, u_{kl}, ...)$ in the jet space is called the *characteristic* of the symmetry (or, the *symmetry characteristic*). Putting

$$\delta u = \bar{u}(x^k;\alpha) - u(x^k) \tag{4.6}$$

we write, for infinitesimal α,

$$\delta u = \alpha Q[u] \tag{4.7}$$

We notice that the infinitesimal operator δ has the following properties:

1. According to its definition (4.2), δ is a linear operator:

$$\delta(F[u]+G[u]) = (F[u+\delta u]+ G[u+\delta u]) - (F[u]+G[u]) = \delta F[u]+\delta G[u] .$$

2. By assumption regarding the nature of our symmetry transformations, δ produces changes in the fiber space while it doesn't affect functions $f(x^k)$ in the base space [this is implicitly stated in (4.6)].

3. Since δ represents a difference, it commutes with all total derivatives D_i:

$$\delta(D_i A[u]) = D_i(\delta A[u]) .$$

In particular, for $A[u]=u$,

ISSN:1791-4469

$$\delta u_i = \delta(D_i u) = D_i(\delta u) = \alpha D_i Q[u],$$

where we have used (4.7).

4. Since δ expresses an infinitesimal change, it may be approximated to a differential; in particular, it satisfies the Leibniz rule:

$$\delta(A[u]B[u]) = (\delta A[u])B[u] + A[u]\delta B[u].$$

For example, $\delta(u^2) = \delta(uu) = (\delta u)u + u\delta u = \alpha(Qu + uQ)$.

Now, consider the characteristic derivative Δ_Q with respect to the symmetry characteristic $Q[u]$. According to (2.7),

$$\Delta_Q u = Q[u] \tag{4.8}$$

We observe that the infinitesimal operator δ and the characteristic derivative Δ_Q share common properties. From (4.7) and (4.8) it follows that the two linear operators are related by

$$\delta u = \alpha \Delta_Q u \tag{4.9}$$

and, by extension,

$$\delta u_i = \alpha D_i Q[u] = \alpha \Delta_Q u_i, \text{ etc.}$$

[see (2.10)]. Moreover, for scalar-valued u and by the infinitesimal character of the operator δ, we may write:

$$\delta F[u] = \frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial u_i}\delta u_i + \cdots = \alpha\left(\frac{\partial F}{\partial u}Q[u] + \frac{\partial F}{\partial u_i}D_i Q[u] + \frac{\partial F}{\partial u_{ij}}D_i D_j Q[u] + \cdots\right)$$

while, by (2.13),

$$\Delta_Q F[u] = \frac{\partial F}{\partial u}Q[u] + \frac{\partial F}{\partial u_i}D_i Q[u] + \frac{\partial F}{\partial u_{ij}}D_i D_j Q[u] + \cdots \tag{4.10}$$

The above observations lead us to the conclusion that, in general, the following relation is true:

$$\delta F[u] = \alpha \Delta_Q F[u] \tag{4.11}$$

The symmetry condition (4.3) is thus written:

$$\Delta_Q F[u] = 0 \mod F[u] \tag{4.12}$$

In particular, if u is scalar-valued, the above condition is written

$$\frac{\partial F}{\partial u}Q[u]+\frac{\partial F}{\partial u_i}D_iQ[u]+\frac{\partial F}{\partial u_{ij}}D_iD_jQ[u]+\cdots=0 \mod F[u] \tag{4.13}$$

which is a linear PDE for $Q[u]$. More generally, for matrix-valued u and for a function $F[u]$ of the form (2.1), the symmetry condition for the PDE $F[u]=0$ is a *linear* PDE for the symmetry characteristic $Q[u]$. We write this PDE symbolically as

$$S(Q;u)\equiv\Delta_Q F[u]=0 \mod F[u] \tag{4.14}$$

where the function $S(Q;u)$ is linear in Q and all total derivatives of Q. (The linearity of S in Q follows from the Leibniz rule and the specific form (2.1) of $F[u]$.)

Below is a list of formulas that may be useful in calculations:

- $\Delta_Q u_i = D_i Q[u]$, $\Delta_Q u_{ij} = D_i D_j Q[u]$, etc.
- $\Delta_Q u^2 = \Delta_Q (uu) = Q[u]u+uQ[u]$ (etc.)
- $\Delta_Q (u^{-1}) = -u^{-1}(\Delta_Q u)u^{-1} = -u^{-1}Q[u]u^{-1}$
- $\Delta_Q [A[u], B[u]] = [\Delta_Q A, B] + [A, \Delta_Q B]$

Comment: According to (4.12), $\Delta_Q F[u]$ vanishes if $F[u]$ vanishes. Given that Δ_Q is a linear operator, the reader may wonder whether this condition is identically satisfied (a linear operator acting on a zero function always produces a zero function!). Note, however, that the function $F[u]$ is *not identically* zero; it becomes zero only *for solutions* of the given PDE. What we need to do, therefore, is to first evaluate $\Delta_Q F[u]$ for *arbitrary* u and *then* demand that the result vanish *when* u is a solution of the PDE $F[u]=0$.

An alternative – and perhaps more transparent – version of the symmetry condition (4.12) is the requirement that the following relation be satisfied:

$$\Delta_Q F[u]=\hat{L}F[u] \tag{4.15}$$

where $\hat{L}$ is a linear operator acting on functions in the jet space (see, e.g., [1], Chap. 2 and 5, for a rigorous justification of this condition in the case of scalar-valued PDEs). For example, one may have

$$\Delta_Q F[u]=\sum_i \beta_i(x^k)D_iF[u]+\sum_{i,j}\gamma_{ij}(x^k)D_iD_jF[u]+A(x^k)F[u]+F[u]B(x^k)$$

where the β_i and γ_{ij} are scalar-valued, while A and B are matrix-valued. Let us see some examples, restricting for the moment our attention to scalar PDEs.

Example 4.1. The *sine-Gordon (s-G) equation* is written

$$F[u]\equiv u_{xt}-\sin u=0\ .$$

ISSN:1791-4469

Here, $(x^1, x^2) \equiv (x, t)$. Since sin$u$ can be expanded into an infinite series in powers of u, we see that $F[u]$ has the required form (2.1). Moreover, since u is a scalar function, we can write the symmetry condition by using (4.13):

$$S(Q;u) \equiv Q_{xt} - (\cos u)\, Q = 0 \mod F[u]$$

where $S(Q;u) = \Delta_Q F[u]$ and where by subscripts we denote total differentiations with respect to the indicated variables. Let us verify the solution $Q[u] = u_x$. This characteristic corresponds to the symmetry transformation [cf. Eq. (4.4)]

$$M:\ u(x,t) \to \bar{u}(x,t;\alpha) = u(x+\alpha,\, t) \tag{4.16}$$

which implies that, if $u(x,t)$ is a solution of the s-G equation, then $\bar{u}(x,t) = u(x+\alpha,\, t)$ also is a solution. We have:

$$Q_{xt} - (\cos u)\, Q = (u_x)_{xt} - (\cos u)\, u_x = (u_{xt} - \sin u)_x = D_x F[u] = 0 \mod F[u]\,.$$

Notice that $\Delta_Q F[u]$ is of the form (4.15), with $\hat{L} \equiv D_x$. Similarly, the characteristic $Q[u] = u_t$ corresponds to the symmetry

$$M:\ u(x,t) \to \bar{u}(x,t;\alpha) = u(x,\, t+\alpha) \tag{4.17}$$

That is, if $u(x,t)$ is a solution of the s-G equation, then so is $\bar{u}(x,t) = u(x,\, t+\alpha)$. The symmetries (4.16) and (4.17) reflect the fact that the s-G equation does not contain the variables x and t explicitly. (Of course, this equation has many more symmetries which are not displayed here; see, e.g., [1].)

Example 4.2. The *heat equation* is written

$$F[u] \equiv u_t - u_{xx} = 0\,.$$

The symmetry condition (4.13) reads

$$S(Q;u) \equiv Q_t - Q_{xx} = 0 \mod F[u]$$

where $S(Q;u) = \Delta_Q F[u]$. As is easy to show, the symmetries (4.16) and (4.17) are valid here, too. Let us now try the solution $Q[u] = u$. We have:

$$Q_t - Q_{xx} = u_t - u_{xx} = F[u] = 0 \mod F[u]\,.$$

This symmetry corresponds to the transformation

$$M:\ u(x,t) \to \bar{u}(x,t;\alpha) = e^{\alpha} u(x,t) \tag{4.18}$$

and is a consequence of the linearity of the heat equation.

Example 4.3. One form of the *Burgers equation* is

$$F[u] \equiv u_t - u_{xx} - u_x{}^2 = 0 \ .$$

The symmetry condition (4.13) is written

$$S(Q;u) \equiv Q_t - Q_{xx} - 2u_x Q_x = 0 \mod F[u]$$

where, as always, $S(Q;u)=\Delta_Q F[u]$. Putting $Q= u_x$ and $Q= u_t$, we verify the symmetries (4.16) and (4.17):

$$Q_t - Q_{xx} - 2u_x Q_x = u_{xt} - u_{xxx} - 2u_x u_{xx} = D_x F[u] = 0 \mod F[u]$$
$$Q_t - Q_{xx} - 2u_x Q_x = u_{tt} - u_{xxt} - 2u_x u_{xt} = D_t F[u] = 0 \mod F[u]$$

Note again that $\Delta_Q F[u]$ is of the form (4.15), with $\hat{L} \equiv D_x$ and $\hat{L} \equiv D_t$. Another symmetry is $Q[u]=1$, which corresponds to the transformation

$$M: \ u(x,t) \to \bar{u}(x,t;\alpha) = u(x,t) + \alpha \tag{4.19}$$

and is a consequence of the fact that u enters $F[u]$ only through its derivatives.

Example 4.4. The *wave equation* is written

$$F[u] \equiv u_{tt} - c^2 u_{xx} = 0 \quad (c = \text{const.})$$

and its symmetry condition reads

$$S(Q;u) \equiv Q_{tt} - c^2 Q_{xx} = 0 \mod F[u] \ .$$

The solution $Q[u] = x u_x + t u_t$ corresponds to the symmetry transformation

$$M: \ u(x,t) \to \bar{u}(x,t;\alpha) = u(e^{\alpha}x, e^{\alpha}t) \tag{4.20}$$

expressing the invariance of the wave equation under a scale change of x and t. [The reader may show that the transformations (4.16) – (4.19) also express symmetries of the wave equation.]

It is remarkable that each of the above PDEs admits an infinite set of symmetry transformations [1]. An effective method for finding such infinite sets is the use of a *recursion operator*, which produces a new symmetry characteristic every time it acts on a known characteristic. More will be said on recursion operators in Sec. 6.

5. The Lie algebra of symmetries

As is well known [1], the set of symmetries of a PDE $F[u]=0$ has the structure of a Lie algebra. Let us demonstrate this property in the context of our abstract formalism.

ISSN:1791-4469

Proposition 5.1. Let $\mathcal{L}$ be the set of characteristic derivatives Δ_Q with respect to the symmetry characteristics $Q[u]$ of the PDE $F[u]=0$. The set $\mathcal{L}$ is a (finite- or infinite-dimensional) Lie subalgebra of the Lie algebra of characteristic derivatives acting on functions $F[u]$ in the jet space (cf. Sec. 3).

Proof. (*a*) Let $\Delta_Q \in \mathcal{L} \Leftrightarrow \Delta_Q F[u]=0 \pmod{F[u]}$. If λ is a constant (real or complex, depending on the nature of the problem) then $(\lambda\Delta_Q)F[u] \equiv \lambda\Delta_Q F[u]=0$, which means that $\lambda\Delta_Q \in \mathcal{L}$. Given that $\lambda\Delta_Q = \Delta_{\lambda Q}$ [see Eq. (3.1)] we conclude that, if $Q[u]$ is a symmetry characteristic of $F[u]=0$, then so is $\lambda Q[u]$.

(*b*) Let $\Delta_1 \in \mathcal{L}$ and $\Delta_2 \in \mathcal{L}$ be characteristic derivatives with respect to the symmetry characteristics $Q_1[u]$ and $Q_2[u]$, respectively. Then, $\Delta_1 F[u]=0$, $\Delta_2 F[u]=0$, and hence, $(\Delta_1+\Delta_2)F[u] \equiv \Delta_1 F[u]+\Delta_2 F[u]=0$; therefore, $(\Delta_1+\Delta_2)\in\mathcal{L}$. It also follows from Eq. (3.2) that, if $Q_1[u]$ and $Q_2[u]$ are symmetry characteristics of $F[u]=0$, then so is their sum $Q_1[u]+Q_2[u]$.

(*c*) Let $\Delta_1 \in \mathcal{L}$ and $\Delta_2 \in \mathcal{L}$, as above. Then, by (4.15),

$$\Delta_1 F[u] = \hat{L}_1 F[u], \quad \Delta_2 F[u] = \hat{L}_2 F[u].$$

Now, by the definition of the Lie bracket and the linearity of both Δ_i and $\hat{L}_i$ (i=1,2) we have:

$$[\Delta_1, \Delta_2]F[u] = \Delta_1(\Delta_2 F[u]) - \Delta_2(\Delta_1 F[u]) = \Delta_1(\hat{L}_2 F[u]) - \Delta_2(\hat{L}_1 F[u])$$
$$\equiv (\Delta_1\hat{L}_2 - \Delta_2\hat{L}_1)F[u] = 0 \mod F[u]$$

We thus conclude that $[\Delta_1, \Delta_2]\in\mathcal{L}$. Moreover, it follows from Eq. (3.3) that, if $Q_1[u]$ and $Q_2[u]$ are symmetry characteristics of $F[u]=0$, then so is the function

$$Q_{1,2}[u] = \Delta_1 Q_2[u] - \Delta_2 Q_1[u].$$

Assume now that the PDE $F[u]=0$ has an n-dimensional symmetry algebra $\mathcal{L}$ (which may be a finite subalgebra of an infinite-dimensional symmetry Lie algebra). Let $\{\Delta_1, \Delta_2, \ldots, \Delta_n\}\equiv\{\Delta_k\}$, with corresponding symmetry characteristics $\{Q_k\}$, be a set of n linearly independent operators that constitute a basis of $\mathcal{L}$, and let Δ_i, Δ_j be any two elements of this basis. Given that $[\Delta_i, \Delta_j]\in\mathcal{L}$, this Lie bracket must be expressible as a linear combination of the $\{\Delta_k\}$, with constant coefficients. We write

$$[\Delta_i, \Delta_j] = \sum_{k=1}^{n} c_{ij}^{k} \Delta_k \tag{5.1}$$

where the coefficients of the Δ_k in the sum are the antisymmetric *structure constants* of the Lie algebra $\mathcal{L}$ in the basis $\{\Delta_k\}$.

The operator relation (5.1) can be expressed in an equivalent, characteristic form by allowing the operators on both sides to act on u and by using the fact that $\Delta_k u = Q_k[u]$:

$$[\Delta_i\,,\,\Delta_j]u = \left(\sum_{k=1}^{n} c_{ij}^{k}\,\Delta_k\right)u = \sum_{k=1}^{n} c_{ij}^{k}\,(\Delta_k u) \;\Rightarrow$$

$$\Delta_i\,Q_j[u] - \Delta_j\,Q_i[u] = \sum_{k=1}^{n} c_{ij}^{k}\,Q_k[u] \tag{5.2}$$

Example 5.1. One of the several forms of the *Korteweg-de Vries (KdV) equation* is

$$F[u] \equiv u_t + uu_x + u_{xxx} = 0\,.$$

The symmetry condition (4.14) is written

$$S(Q;u) \equiv Q_t + Q\,u_x + u\,Q_x + Q_{xxx} = 0 \mod F[u] \tag{5.3}$$

where $S(Q;u) = \Delta_Q F[u]$. The KdV equation admits a symmetry Lie algebra of infinite dimensions [1]. This algebra has a finite, 4-dimensional subalgebra $\mathcal{L}$ of *point transformations*. A symmetry operator (characteristic derivative) Δ_Q is determined by its corresponding characteristic $Q[u] = \Delta_Q u$. Thus, a basis $\{\Delta_1,...,\Delta_4\}$ of $\mathcal{L}$ corresponds to a set of four independent characteristics $\{Q_1,...,Q_4\}$. Such a basis of characteristics is the following:

$$Q_1[u] = u_x\,,\quad Q_2[u] = u_t\,,\quad Q_3[u] = tu_x - 1\,,\quad Q_4[u] = xu_x + 3tu_t + 2u$$

The $Q_1,...,Q_4$ satisfy the PDE (5.3), since, as we can show,

$$S(Q_1;u) = D_x F[u]\,,\quad S(Q_2;u) = D_t F[u]\,,\quad S(Q_3;u) = tD_x F[u]\,,$$
$$S(Q_4;u) = (5 + xD_x + 3tD_t)F[u]$$

[Note once more that $\Delta_Q F[u]$ is of the form (4.15) in each case.] Let us now see two examples of calculating the structure constants of $\mathcal{L}$ by application of (5.2). We have:

$$\Delta_1 Q_2 - \Delta_2 Q_1 = \Delta_1 u_t - \Delta_2 u_x = (\Delta_1 u)_t - (\Delta_2 u)_x = (Q_1)_t - (Q_2)_x = (u_x)_t - (u_t)_x = 0$$
$$\equiv \sum_{k=1}^{4} c_{12}^{k}\,Q_k$$

Since the Q_k are linearly independent, we must necessarily have $c_{12}^{k} = 0\,,\ k = 1,2,3,4$. Also,

$$\Delta_2 Q_3 - \Delta_3 Q_2 = \Delta_2(tu_x - 1) - \Delta_3 u_t = t(\Delta_2 u)_x - (\Delta_3 u)_t = t(Q_2)_x - (Q_3)_t$$
$$= tu_{tx} - (u_x + tu_{xt}) = -u_x = -Q_1 \equiv \sum_{k=1}^{4} c_{23}^{k}\,Q_k$$

ISSN:1791-4469

Therefore, $c_{23}^1 = -1$, $c_{23}^2 = c_{23}^3 = c_{23}^4 = 0$.

6. Recursion operators

Let $\delta u = \alpha Q[u]$ be an infinitesimal symmetry of the PDE $F[u]=0$, where $Q[u]$ is the symmetry characteristic. For any solution $u(x^k)$ of this PDE, the function $Q[u]$ satisfies the linear PDE

$$S(Q;u) \equiv \Delta_Q F[u] = 0 \tag{6.1}$$

Because of the linearity of (6.1) in Q, the sum $Q_1[u]+Q_2[u]$ of two solutions of this PDE, as well as the multiple $\lambda Q[u]$ of any solution by a constant, also are solutions of (6.1) for a given u. Thus, for any solution u of $F[u]=0$, the solutions $\{Q[u]\}$ of the PDE (6.1) form a linear space, which we call S_u.

A *recursion operator* $\hat{R}$ is a linear operator that maps the space S_u into itself. Thus, if $Q[u]$ is a symmetry characteristic of $F[u]=0$ (i.e., a solution of (6.1) for a given u) then so is $\hat{R}Q[u]$:

$$S(\hat{R}Q;u) = 0 \quad \text{when} \quad S(Q;u) = 0 \tag{6.2}$$

Obviously, any power of a recursion operator also is a recursion operator. Thus, starting with any symmetry characteristic $Q[u]$, one may in principle obtain an infinite set of such characteristics by repeated application of the recursion operator.

A new approach to recursion operators was suggested in the early 1990s [11,15-17] (see also [8-10]) according to which a recursion operator may be viewed as an *auto-Bäcklund transformation* (BT) [18] for the symmetry condition (6.1) of the PDE $F[u]=0$. By integrating the BT, one obtains new solutions $Q'[u]$ of the linear PDE (6.1) from known ones, $Q[u]$. Typically, this type of recursion operator produces *nonlocal* symmetries in which the symmetry characteristic depends on *integrals* (rather than derivatives) of u. This approach proved to be particularly effective for matrix-valued PDEs such as the nonlinear self-dual Yang-Mills equation, of which new infinite-dimensional sets of "potential symmetries" were discovered [9,11,15].

7. An example: The chiral field equation

Let us consider the *chiral field equation*

$$F[g] \equiv (g^{-1}g_x)_x + (g^{-1}g_t)_t = 0 \tag{7.1}$$

where, in general, subscripts x and t denote total derivatives D_x and D_t, respectively, and where g is a $GL(n,C)$-valued function of x and t, i.e., a complex, non-singular $(n\times n)$ matrix function, differentiable for all x and t. Let $\delta g = \alpha Q[g]$ be an infinitesimal symmetry transformation for the PDE (7.1), with symmetry characteristic $Q[g]=\Delta_Q g$. It is convenient to put

$$Q[g] = g\,\Phi[g] \quad \Leftrightarrow \quad \Phi[g] = g^{-1}Q[g]\ .$$

The symmetry condition for (7.1) is

$$\Delta_Q F[g] = 0 \mod F[g]\ .$$

This condition will yield a linear PDE for Q or, equivalently, a linear PDE for Φ. By using the properties of the characteristic derivative, we find the latter PDE to be

$$S(\Phi; g) \equiv D_x\left(\Phi_x + [g^{-1}g_x, \Phi]\right) + D_t\left(\Phi_t + [g^{-1}g_t, \Phi]\right) = 0 \mod F[g] \qquad (7.2)$$

where, as usual, square brackets denote commutators of matrices.

A useful identity that will be needed in the sequel is the following:

$$(g^{-1}g_t)_x - (g^{-1}g_x)_t + [g^{-1}g_x,\ g^{-1}g_t] = 0 \qquad (7.3)$$

Let us first consider symmetry transformations in the base space, i.e., coordinate transformations of x, t. An obvious symmetry is x-translation, $x'=x+\alpha$, given that the PDE (7.1) does not contain the independent variable x explicitly. For infinitesimal values of the parameter α, we write $\delta x=\alpha$. The symmetry characteristic is $Q[g]=g_x$, so that $\Phi[g]=g^{-1}g_x$. By substituting this expression for Φ into the symmetry condition (7.2) and by using the identity (7.3), we can verify that (7.2) is indeed satisfied:

$$S(\Phi; g) = D_x F[g] = 0 \mod F[g]\ .$$

Similarly, for t-translation, $t'=t+\alpha$ (infinitesimally, $\delta t=\alpha$) with $Q[g]=g_t$, we find

$$S(\Phi; g) = D_t F[g] = 0 \mod F[g]\ .$$

Another obvious symmetry of (7.1) is a scale change of both x and t: $x'=\lambda x$, $t'=\lambda t$. Setting $\lambda=1+\alpha$, where α is infinitesimal, we write $\delta x=\alpha x$, $\delta t=\alpha t$. The symmetry characteristic is $Q[g]=xg_x+tg_t$, so that $\Phi[g]=xg^{-1}g_x+tg^{-1}g_t$. Substituting for Φ into the symmetry condition (7.2) and using the identity (7.3) where necessary, we find that

$$S(\Phi; g) = \left(2 + xD_x + tD_t\right) F[g] = 0 \mod F[g]\ .$$

Let us call $Q_1[g]=g_x$, $Q_2[g]=g_t$, $Q_3[g]=xg_x+tg_t$, and let us consider the corresponding characteristic derivative operators Δ_i defined by $\Delta_i\, g=Q_i$ (i=1,2,3). It is then straightforward to verify the following commutation relations:

$$[\Delta_1, \Delta_2]\, g = \Delta_1 Q_2 - \Delta_2 Q_1 = 0 \quad \Leftrightarrow \quad [\Delta_1, \Delta_2] = 0\ ;$$

$$[\Delta_1, \Delta_3]\, g = \Delta_1 Q_3 - \Delta_3 Q_1 = -g_x = -Q_1 = -\Delta_1 g \quad \Leftrightarrow \quad [\Delta_1, \Delta_3] = -\Delta_1\ ;$$

ISSN:1791-4469

$$[\Delta_2\,,\Delta_3]\,g = \Delta_2 Q_3 - \Delta_3 Q_2 = -g_t = -Q_2 = -\Delta_2 g \quad \Leftrightarrow \quad [\Delta_2\,,\Delta_3] = -\Delta_2\,.$$

Next, we consider the "internal" transformation (i.e., transformation in the fiber space) $g'=g\Lambda$, where Λ is a non-singular constant matrix. Then,

$$F[g'] = \Lambda^{-1} F[g]\,\Lambda = 0 \mod F[g]\,,$$

which indicates that this transformation is a symmetry of (7.1). Setting $\Lambda=1+\alpha M$, where α is an infinitesimal parameter while M is a constant matrix, we write, in infinitesimal form, $\delta g=\alpha g M$. The symmetry characteristic is $Q[g]=gM$, so that $\Phi[g]=M$. Substituting for Φ into the symmetry condition (7.2), we find:

$$S(\Phi\,;g) = [F[g]\,,M] = 0 \mod F[g]\,.$$

Given a matrix function $g(x,t)$ satisfying the PDE (7.1), consider the following system of PDEs for two functions $\Phi[g]$ and $\Phi'[g]$:

$$\begin{aligned} \Phi'_x &= \Phi_t + [g^{-1}g_t\,,\,\Phi] \\ -\Phi'_t &= \Phi_x + [g^{-1}g_x\,,\,\Phi] \end{aligned} \tag{7.4}$$

The *integrability condition* (or consistency condition) $(\Phi'_x)_t = (\Phi'_t)_x$ of this system requires that Φ satisfy the symmetry condition (7.2); i.e., $S(\Phi\,;g)=0$. Conversely, by applying the integrability condition $(\Phi_t)_x = (\Phi_x)_t$ and by using the fact that g is a solution of $F[g]=0$, one finds that Φ' must also satisfy (7.2); i.e., $S(\Phi'\,;g) = 0$.

We conclude that, for any function $g(x,t)$ satisfying the PDE (7.1), the system (7.4) is an *auto-Bäcklund transformation* (BT) [18] relating solutions Φ and Φ' of the symmetry condition (7.2) of this PDE; that is, relating different symmetries of the chiral field equation. Thus, if a symmetry characteristic $Q=g\Phi$ of the PDE (7.1) is known, a new characteristic $Q'=g\Phi'$ may be found by integrating the BT (7.4); the converse is also true. Since the BT (7.4) produces new symmetries from old ones, it may be regarded as a *recursion operator* for the PDE (7.1) [8-11,15-17].

As an example, consider the internal-symmetry characteristic $Q[g]=gM$ (where M is a constant matrix) corresponding to $\Phi[g]=M$. By integrating the BT (7.4) for Φ', we get $\Phi'=[X,M]$ and thus $Q'=g[X,M]$, where X is the "potential" of the PDE (7.1), defined by the system of PDEs

$$X_x = g^{-1}g_t \;, \quad -X_t = g^{-1}g_x \tag{7.5}$$

Note the *nonlocal* character of the BT-produced symmetry Q', due to the presence of the potential X. Indeed, as seen from (7.5), in order to find X one has to *integrate* the chiral field g with respect to the independent variables x and t. The above process can be continued indefinitely by repeated application of the recursion operator (7.4), leading to an infinite sequence of increasingly nonlocal symmetries.

Unfortunately, as the reader may check, no new information is furnished by the BT (7.4) in the case of coordinate symmetries (for example, by applying the BT for $Q=g_x$ we get the known

symmetry $Q'=g_t$). A recursion operator of the form (7.4), however, does produce new nonlocal symmetries from coordinate symmetries in related problems with more than two independent variables, such as the self-dual Yang-Mills equation [8-11,15].

8. Concluding remarks

The algebraic approach to the symmetry problem of PDEs, presented in this article, is, in a sense, an extension to matrix-valued problems of the ideas contained in [1], in much the same way as [4] and [5] constitute a generalization of the Harrison-Estabrook geometrical approach [2] to matrix-valued (as well as vector-valued and Lie-algebra-valued) PDEs. The main advantage of the algebraic approach is the bypassing of the difficulty associated with the differential-operator representation of the symmetry-generating vector fields that act on matrix-valued functions in the jet space.

It should be noted, however, that the standard methods [1,4,5] are still most effective for *calculating* symmetries of PDEs. In this regard, one needs to enrich the ideas presented in this article by describing a systematic process for evaluating (not just recognizing) symmetries, in the spirit of [4,5]. This will be the subject of a subsequent article.

References

1. P. J. Olver, *Applications of Lie Groups to Differential Equations* (Springer-Verlag, 1993).
2. B. K. Harrison, F. B. Estabrook, *Geometric approach to invariance groups and solutions of partial differential equations*, J. Math. Phys. 12 (1971) 653.
3. D. G. B. Edelen, *Applied Exterior Calculus* (Wiley, 1985).
4. C. J. Papachristou, B. K. Harrison, *Isogroups of differential ideals of vector-valued differential forms: Application to partial differential equations*, Acta Appl. Math. 11 (1988) 155.
5. C. J. Papachristou, B. K. Harrison, *Some aspects of the isogroup of the self-dual Yang-Mills system*, J. Math. Phys. 28 (1987) 1261.
6. C. J. Papachristou, B. K. Harrison, *Nonlocal symmetries and Bäcklund transformations for the self-dual Yang-Mills system*, J. Math. Phys. 29 (1988) 238.
7. C. J. Papachristou, *Symmetry and integrability of classical field equations*, arXiv:0803.3688 (http://arxiv.org/abs/0803.3688).
8. C. J. Papachristou, *Symmetry, conserved charges, and Lax representations of nonlinear field equations: A unified approach*, Electron. J. Theor. Phys. 7, No. 23 (2010) 1 (http://www.ejtp.com/articles/ejtpv7i23p1.pdf).
9. C. J. Papachristou, B. K. Harrison, *Bäcklund-transformation-related recursion operators: Application to the self-dual Yang-Mills equation*, J. Nonlin. Math. Phys., Vol. 17, No. 1 (2010) 35.
10. C. J. Papachristou, *Symmetry and integrability of a reduced, 3-dimensional self-dual gauge field model*, Electron. J. Theor. Phys. 9, No. 26 (2012) 119. (http://www.ejtp.com/articles/ejtpv9i26p119.pdf).

ISSN:1791-4469

11. C. J. Papachristou, *Lax pair, hidden symmetries, and infinite sequences of conserved currents for self-dual Yang-Mills fields*, J. Phys. A 24 (1991) L 1051.
12. M. Marvan, A. Sergyeyev, *Recursion operators for dispersionless integrable systems in any dimension*, Inverse Problems 28 (2012) 025011.
13. S. Igonin, M. Marvan, *On construction of symmetries and recursion operators from zero-curvature representations and the Darboux–Egoroff system*, J. Geom. Phys. 85 (2014) 106.
14. A. Sergyeyev, *A simple construction of recursion operators for multidimensional dispersionless integrable systems*, J. Math. Anal. Appl. 454 (2017) 468.
15. C. J. Papachristou, *Potential symmetries for self-dual gauge fields*, Phys. Lett. A 145 (1990) 250.
16. G. A. Guthrie, *Recursion operators and non-local symmetries*, Proc. R. Soc. Lond. A 446 (1994) 107.
17. M. Marvan, *Another look on recursion operators*, in *Differential Geometry and Applications* (Proc. Conf., Brno, 1995) 393.
18. C. J. Papachristou, A. N. Magoulas, *Bäcklund transformations: Some old and new perspectives*, Nausivios Chora Vol. 6 (2016) pp. C3-C17 (http://nausivios.snd.edu.gr/docs/2016C.pdf).

A Decision Aiding System Based On Multiple Criteria For The Support Of The Hellenic Coast Guard in Maritime Search And Rescue Operations

Paraskevas-Marios Chavatzopoulos

HCG Lieutenant Junior Grade, Hellenic Army Academy - Technical University of Crete
parischavatzopoulos@yahoo.gr

Abstract. Search and Rescue maritime operations depend on the suitable choice of ships that can complete satisfactorily their mission and rescue every person in danger, despite the circumstances. To achieve that, one must possess the proper type of ships for this kind of missions to endure weather and sea conditions and deploy at maximum speed. This paper proposes the necessary criteria to be assessed by the UTAstar method to achieve optimal selection.

Keywords: Search and Rescue, SAR, Hellenic Coast Guard, UTAstar method, multiple criteria decision making.

I. INTRODUCTION

Decision making is admittedly a complex procedure aiming at the study and extensive analysis of the critical impact of all alternatives. It then tries to assemble all the demanding requirements of everything involved at the procedure. [7] The efficient use of such key tools will optimize the necessary quality of accurate information typically linked to the decision making and will enable the decision makers to accurately analyze and decide more precisely the possible alternatives. The possible selection of specific types of ships involved at maritime operations for Search and Rescue is undoubtedly in need of such a procedure.

The specific type of boats used at maritime Search and Rescue operations naturally have to be carefully selected with a complex decision making procedure in order to take under consideration the time response, the work load, the unpredictable weather and sea conditions and the continuous economic demands.

This paper describes and discusses the various criteria that need to be carefully considered, both quantitatively and qualitatively, and the used methodology that has to be followed for the aforementioned procedure.

To address the problem of choosing Search and Rescue (SAR) ships, we use a multicriteria method called UTAstar [2] (Siskos and Yannacopoulos, 1985) which is an optimization of UTA[3] method (Lagreze and Siskos, 1982).

Solutions are obtained for the existing fleet of the Hellenic Coast Guard. For security reasons, none of the ships' capabilities will be presented.

The main goal is to provide an executive committee with a practical tool to delegate the choosing of new coast guard ships for specific missions.

ISSN:1791-4469

II. METHODOLOGY

As aforementioned, the method used to assess the alternative SAR ships is UTAstar [2]. UTAstar presents a low structural indicator and the capability to compare the alternative ships pairwise. It can also handle effectively both qualitative and quantitative criteria. UTAstar method is characterized as a monotonic regression method for analyzing the decision makers' a priori preferences (Matsatsinis, 2005) [7].

UTAstar is a set of utility functions that are models consistent with the decision maker's a priori preferences. In order to assess this set of utility function, the method uses ordinal regression method. Using linear programming, it adjusts optimally additive non-linear utility functions so that they fit data which consist of multicriteria evaluations of some alternatives and a subjective ranking of these alternatives given by the decision maker.

Concluding the procedure, the UTAstar method will have ranked the types of ships from the most suitable to the least suitable one for maritime SAR operations.

UTAstar Method [2, 4, 5, 7]

Overview

This method consists of three things: a set of decision makers, a set of quantitative and qualitative criteria and a set of alternatives. At first, a questionnaire must be completed by each decision maker so as to evaluate each alternative over each criterion. Each criterion can be evaluated with a value within the boundaries (best and worst value) that have been a priori set and given to the decision maker. After that, the decision maker ranks all the alternatives. There is predefined structure of preferences (>, ~) with which one declares either absolute preference (>) or indifference (~) over a set of alternatives; this means that one can rank two or more alternatives at the same place.

Once the decision maker expresses his judgment in a form of a ranking, the method estimates an additive utility function that is as consistent as possible with the decision maker's opinion. This allows the decision makers to do an empirical evaluation and rank the ships regardless of the quantitative criteria.

Then, all questionnaires are fed to the method in order to process them following four simple steps that are defined below. The result is a table of alternatives with a specific value that defines the final rank of each alternative.

Definitions

Assume $A = \{a_1, a_2, \ldots, a_m\}$ is the set of the alternative ships offered for evaluation by the set $J = \{1, 2, \ldots, p\}$ of decision makers over the set of criteria $g: g_1, g_2, \ldots, g_n$. Each criterion g represents a quantitative/qualitative monotonic variable. For each $a_i \in A$, $g(a_i) = [g_1(a_i), g_2(a_i), \ldots, g_n(a_i)]$ depicts the multicriteria judgment of the i^{th} alternative ship expressed by each decision maker.[7]

FIGURE 1. Depiction of vector g(a$_i$) for quantitative criteria.[7]

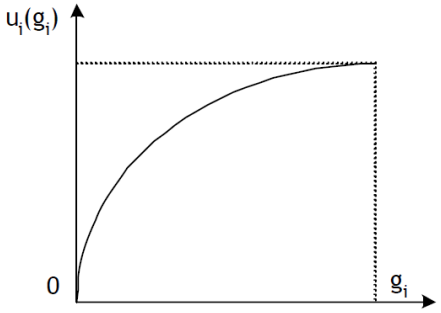
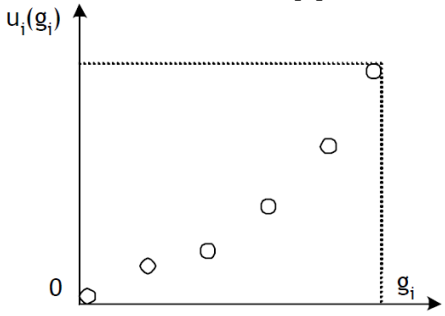


FIGURE 2. Depiction of vector g(a$_i$) for qualitative criteria.[7]



It is vital to clarify the scales of measurement for each criterion. Thus, g_{i*} is the worst value for the criterion and g_i^* is the best value. All the values in between are put in $\alpha_i - 1$ spacing. And so the scales for the i criterion is as follows: $G_i = \{g_{i*} = g_i^1, g_i^2 = g_i^l, \dots, g_i^{a_i} = g_i^*\}$. The value for the best and worst value, the monotony, the spacing a_i for each criterion is given beforehand by the system.

TABLE 1. Spacing, value range, monotony of criteria.

Best Value	$r^* = 1$	$g_1 *$	$g_2 *$	...	g_{n*}
Worst Value	$r *$	$g_1{}^*$	$g_2{}^*$	...	$g_n{}^*$
Monotony	0 if Best Value > Worst Value, else 1				
Spacing		s_1	s_2	...	s_n

The k^{th} decision maker expresses his judgement and ranks each type of ship, which is then introduced at the vector $r_k(a) = [r_k(a_1), r_k(a_2), \dots, r_k(a_n)]$. The alternatives with small numbers signify better suitability and the alternative with $r_1^* = 1$ is the most suitable. The set of alternatives $A_k = \{a_1, \dots, a_k\}$ is sorted according to the ranking of the decision maker. The judgment for each alternative over each criterion for the k^{th} decision maker produces the table below.

TABLE 2. Criteria and alternatives for $\boldsymbol{k^{th}}$ decision maker.

Alternative\Criterion	$\boldsymbol{g_1}$	$\boldsymbol{g_2}$	...	$\boldsymbol{g_n}$	**Ranking**
$\boldsymbol{a_1}$	$g_{1_k}(a_1)$	$g_{2_k}(a_1)$	...	$g_{n_k}(a_1)$	$r_{1_k}(a_1)$
$\boldsymbol{a_2}$	$g_{1_k}(a_2)$	$g_{2_k}(a_2)$	...	$g_{n_k}(a_2)$	$r_{2_k}(a_2)$
...	...	...	...	...	...
$\boldsymbol{a_m}$	$g_{1_k}(a_m)$	$g_{2_k}(a_m)$	...	$g_{n_k}(a_m)$	$r_{m_k}(a_m)$

The first thing one must do is to sort the alternatives according to the ranking for each decision maker. In this method, one can rank the same two or more alternatives. Hence, there are two cases compared pairwise; either one alternative is preferred (>) or both alternatives are ranked the same (~).

Next step is the aggregation of the n criteria and the errors of underestimation $\sigma^+(a_i)$ and overestimation $\sigma^-(a_i)$ in one global value $u\big(g(a_i)\big)$ as follows:

ISSN:1791-4469

$$u\big(g(a_i)\big) = u_1\big(g_1(a_i)\big) + u_2\big(g_2(a_i)\big) + \cdots + u_n\big(g_n(a_i)\big) - \sigma^+(a_i) + \sigma^-(a_i)$$
$$= \sum_{j=1}^{n} u_j\left(g_j(a_i)\right) - \sigma^+(a_i) + \sigma^-(a_i), i = 1..m$$

This global value is broken into additive utility functions $u_j\left(g_j(a_i)\right)$. For the evaluation of the partial utility functions, all values of u must be always expressed in terms of the values of the boundaries of each spacing of the i^{th} criterion,

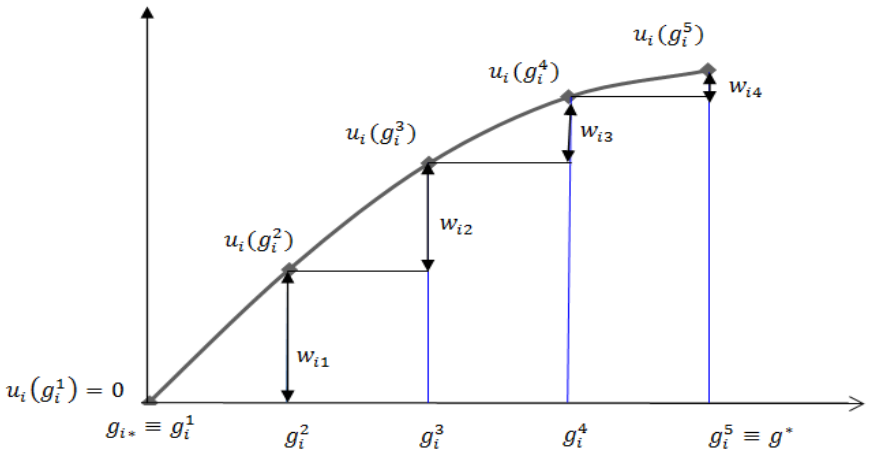
$$u_i(g_i(\alpha_l)) = u_i\big(g_i^j\big) + \frac{g_i(x) - g_i^j}{g_i^{j+1} - g_i^j}\left[u_i(g_i^{j+1}) - u_i(g_i^j)\right]$$

These partial utility functions represent the preferences of the decision maker from the least desirable alternative, where $u_i(g_i^1) = 0$, to the most desirable one $u_i(g_i^*)$. This is clearly shown in Figure 4.

The restrictions of monotony are modeled with the variables w_{ij}

$$w_{ij} = u_i\big(g_i^{j+1}\big) - u_i\big(g_i^j\big) \geq 0, \forall i = 1,2,\dots,n, \qquad j = 1,2,\dots,s_i - 1$$

FIGURE 4. Additive utility function versus criteria. [9]



The total value of the alternatives $u(g(\alpha_k)),\ k = 1,2,\dots,m$ is expressed as the sum of all weights and is zero for the worst value.

$$\begin{cases} u_i(g_i^1) = 0, \ \ \forall i = 1,2,\dots,n \\ u_i\big(g_i^j\big) = \sum_{i=1}^{j-1} w_{ij}\,, \forall i = 1,2,\dots,n \ \ j = 2,3,\dots s_i - 1 \end{cases}$$

The process of comparing pairwise follows. It is symbolized with Δ. It takes into account the initial ranking and the utility functions.

$$\Delta(a_k, a_{k+1}) = u_i(g(\alpha_k)) - \sigma^+(\alpha_k) + \sigma^-(\alpha_k) - u_i(g(\alpha_{k+1})) + \sigma^+(\alpha_{k+1}) - \sigma^+(\alpha_{k+1})$$

with the following restrictions where δ is the threshold and defined by the system. In this particular scenario is set to 0.05.

$$\Delta(a_k, a_{k+1}) \begin{cases} \geq \delta, & if\ \ a_k > a_{k+1} \\ = 0, & if\ \ a_k \sim a_{k+1} \end{cases}$$

Linear Program

To complete the process one must solve the Linear Program of minimizing the sum of the underestimating and overestimating errors of all the alternatives.

$$Minimize\ z = \sum_{k=1}^{m} (\sigma^+(\alpha_k) + \sigma^-(\alpha_k))$$

Subject to

$$\Delta(a_k, a_{k+1}) \geq \delta, \qquad if\ \ a_k > a_{k+1}\ \forall k = 1,2, \dots m$$

$$\Delta(a_k, a_{k+1}) = 0, \qquad if\ \ a_k \sim a_{k+1}\ \ \forall k = 1,2, \dots m$$

$$\sum_{i=1}^{n} \sum_{j=1}^{s_i - 1} w_{ij} = 1$$

$$w_{ij} \geq 0, \sigma^+(\alpha_k) \geq 0, \sigma^-(\alpha_k) \geq 0, \forall i, j, k$$

Final Step[9]

In the final step, one must test the multiple or near optimal solutions of the linear program and in case of non uniqueness one should find the mean additive value function of those (near) optimal solutions which maximize the objective functions:

$$\sum_{i=1}^{m} [\sigma^+(\alpha_i) + \sigma^-(\alpha_i)] \leq z^* + \varepsilon$$

where z^* is the optimal value of the linear program and ε a very small positive number.

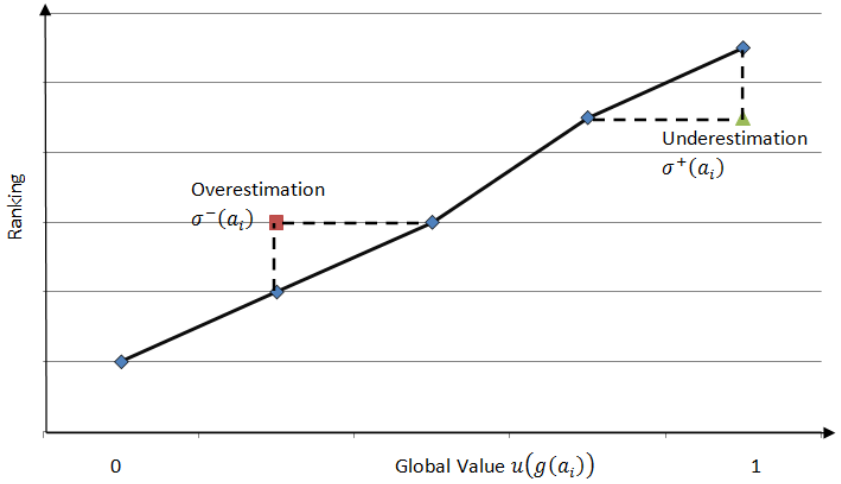
Advantages

The greatest advantage in using the UTAstar method is the use of a double error equation, both for underestimation (σ^+) and overestimation (σ^-), leading to an optimized decision (figure 3). The double error equation is used to make the alternative regain its position in the predefined ranking. It is an amount of utility that will be added or subtracted depending on the position of the alternative on the curve as shown in the following figure.

Another advantage is the ability to take into account both qualitative and quantitative criteria; the experience one has is very important in search and rescue as the numbers cannot depict fully the sea conditions.

ISSN:1791-4469

FIGURE 3. Ranking versus global value. Double error graphically explained. [9]



Criteria and Alternatives

Criteria Used for the Decision Making

To correctly decide upon the types of the ships considered, one must carefully look upon many a criterion to judge wisely and derive scientifically to a conclusion. Therefore, each decision maker judges over the following 20 criteria, 13 quantitative and 7 qualitative. They cover as many key aspects as possible of a SAR ship and sufficiently examine the ability to be engaged in a SAR operation and the economic cost to do so.

TABLE 3. Description and characteristics of the criteria used for the decision making. Values with (-) cannot be disclosed for security reasons.

	Criteria	Description	g_{i*} (Worst Value)	g_i^* (Best Value)	s_i	Mono-tonicity
General Characteristics	Maximum speed (miles/hr)	Top speed for the type	18	45	3	0
	Autonomy	Distance that the boat can travel and return to port without refueling.	-	-	3	0
	Transportation capability	Number of persons (survivors, etc) that can carry safely back to shore/port	-	-	3	0
	Required personnel	Minimum number of officers needed	-	-	3	0
Engine Characte-ristics	Engine suitability	Check if the engines are proper and can withstand the stress of maritime operations at all weathers	1	10	10	0
	Horsepower	Total horsepower of all the engines on the ship	-	-	5	0
ra ct eri	Technical	The specialization required	1	10	10	0

	Specialization	for a service to the ship or to address major problems on the ship				
	Ease of finding spare parts	The extent of specificity of the parts needed	1	10	10	0
	Fuel Tanks	The total amount of fuels that the ship can carry	-	-	4	0
	Fuel consumption	The amount of liters consumed in an hour	-	-	4	1
	Lubricant consumption	The amount of lubricant oils consumed in an hour	-	-	4	1
Hull Characteristics	Hull suitability	The hull's design (e.g. V-shaped) suitable for extreme weather conditions	1	10	10	0
	Sailing in extreme weather conditions	The seakeeping of the ship and how it responds to extreme sea conditions (>8BF)	1	10	10	0
	Maximum wind speed(BF)	The maximum wind speed that the ship stays seaworthy	1	10	10	0
	Material strength	The material of the hull defines the weather conditions the ship can stay seaworthy	1	10	10	0
	Self-righting	If the ship has the capability to self-right or not	0	1	1	0
Costs	Warranty	If the manufacturer gives a few years warranty	0	1	1	0
	Engine maintenance	The cost of one engine service	-	-	4	1
	Maintenance cost	The cost of all systems service	-	-	4	1
	Buy/Restore	The cost to buy this type of ship or to restore a seized/ decommissioned one	20.000.000	10.000	5	1

Alternatives – SAR Boats Assessed

The types of ships assessed by the decision makers comprise of the specific types of boats that the Hellenic Coast Guard currently uses for maritime SAR operations. For security reasons, only what can be found in open sources will be presented [8] and none other value of the criteria will be disclosed.

ISSN:1791-4469

TABLE 4. Type of Hellenic Coast Guard SAR boats assessed.

Model	Role	Length (meters)	Displacement (tones)	Number of boats in use
Lambro Halmatic 60	Salvage Boat	18	37	10
Sa'ar 4.5	OPV	58	450	3
Stan Patrol 5509	OPV	58.5	700	1
Vosper Europatrol 250 MkI	OPV	47.3	300	1
Class Dilos	Patrol Boat	29	86	6
Class Faiakas	Patrol Boat	24.6	-	2
CB-90 HCG	Patrol Boat - Combat	15.9	20	3
LCS-57 (Lambro 57) Mk I	Patrol Boat	18.2	28	19
LCS-57 (Lambro-57) Mk II	Patrol Boat	19.2	27	16

Decision Making Process

A standard questionnaire for the mentioned types of boats was created. The quantitative criteria were precisely defined from the manual specification of each type. 10 decision makers with vast experience in these types of boats were asked to express their judgment over the qualitative criteria. After that, they ranked the considered 9 types of boats from the most suitable to the least suitable one to participate in a maritime SAR operation.

The 10 questionnaires were input to a MATLAB-based UTAstar fully customizable program created for the purpose of this project.

III. RESULTS

The MATLAB-based program solved the linear problem of UTAstar for the 10 evaluations of the 20 criteria for the 9 types of boats and resulted in the final weights of each type of boat. The final rank is presented below.

TABLE 5. Resulting ranking of the existing fleet of HCG boats engaged in SAR operations.

Suitability (Best to Least)	Type
1	CB-90 HCG
2	LCS-57 (Lambro-57) Mk I
3	Class Dilos
4	Vosper Europatrol 250 Mk I
5	Sa'ar 4.5
6	Class Faiakas
7	Stan Patrol 5509
8	Lambro Halmatic 60
9	LCS-57 (Lambro-57) Mk II

IV. CONCLUSION AND PERSPECTIVES

In this paper, a set of criteria for evaluating various types of boats using UTAstar[2] is presented. UTAstar has significant advantages to the evaluation. Firstly, it can be customized accordingly and add more criteria if needed. Moreover, it takes into account, not only the facts and the economic costs, but also the personal opinion of a decision maker.

The result shows that all factors were taken into account. A ship of the Hellenic Coast Guard that costs 20.000.000, 00€ did not rank first, although it is the fastest. There are necessary modifications to be done that will enable a committee to focus on a set of the abovementioned criteria which are more relevant to the nature of the intended use. The possible use of a weighted UTAstar method would benefit the decision making process.

Concluding, the method presented is an extremely effective tool for any committee deciding over which type of boat/ship must be acquired. It can be progressively extended to more than SAR operations. One can accurately evaluate either the helicopters participating in SAR missions or even the types of ships of a naval fleet assigned to different kind of tasks.

ACKNOWLEDGMENTS

This is a work carried out in the framework of the Postgraduate Studies Program "Applied Operational Research and Analysis" at the Hellenic Army Academy – Technical University of Crete.

The opinions of the writer are purely for academic purposes and do not depict at any case those of the Hellenic Ministry of Maritime Affairs and Island Policy or the Hellenic Coast Guard.

REFERENCES

1. P.M. Chavatzopoulos, "Maritime Operations' Design for Search and Rescue", Technical University of Crete-Hellenic Army Academy, 2019.
2. Y. Siskos, D. Yannacopoulos, "UTASTAR: An ordinal regression method for building additive value functions'', Investigacao Operational, vol. 5, no. 1, pp. 39-53, 1985.
3. E. Jacquet-Lagreze, J. Siskos, "Assessing a set of additive utility functions for multicriteria decision-making, the UTA method'', European journal of operational research, vol. 10, no. 2, pp. 151-164, 1982.
4. N. Matsatsinis, Multicriteria Decision Analysis – UTAstar Method, 2017.(lectures for "Decision Analysis" course at Hellenic Army Academy)
5. E. Krassadaki, "Methodology to estimate non-typical and non-certified knowledge acquired during university studies", Technical University of Crete, 2013.
6. E. Krassadaki, K. Lakiotaki, N.F. Matsatsinis, "Students' behaviour in peer assessment: a multi-criteria clustering approach'', European Journal of Engineering Education, vol. 39, no. 3, pp. 233-246, 2014.
7. D. Trehas, "Develop web-based decision support system for analysis and simulate consumer behavior", Technical University of Crete, 2005.
8. Hellenic Coast Guard, https://de.wikipedia.org/wiki/Limeniko_Soma.
9. UTA methods and applications Part I, Y. Siskos, 13th MCDA/MCDM Summer School, 2018.

