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 5/2014 - Τεύχος 5/2014

TABLE OF CONTENTS

Electromotive Force: A Guide for the Perplexed

C.J. Papachristou[a], A.N. Magoulas[b]

[a] *Hellenic Naval Academy, Section of Natural Sciences, Piraeus 18539, Greece*
E-mail: papachristou@snd.edu.gr
[b] *Hellenic Naval Academy, Section of Electrotechnics and Computer Science, Piraeus 18539, Greece*
E-mail: aris@snd.edu.gr

Abstract. The concept of electromotive force (emf) may be introduced in various ways in an undergraduate course of theoretical electromagnetism. The multitude of alternate expressions for the emf is often the source of confusion to the student. We summarize the main ideas, adopting a pedagogical logic that proceeds from the general to the specific. The emf of a "circuit" is first defined in the most general terms. The expressions for the emf of some familiar electrodynamical systems are then derived in a rather straightforward manner. A diversity of physical situations is thus unified within a common theoretical framework.

1. INTRODUCTION

The difficulty in writing this article was not just due to the subject itself: we had to first overcome some almost irreconcilable differences in educational philosophy between an (opinionated) theoretical physicist and an (equally -if not more- opinionated) electrical engineer. At long last, a compromise was reached! This paper is the fruit of this "mutual understanding".

Having taught intermediate-level electrodynamics courses for several years, we have come to realize that, in the minds of many of our students, the concept of *electromotive force* (*emf*) is something of a mystery. What is an emf, after all? Is it the voltage of an ideal battery in a DC circuit? Is it work per unit charge? Or is it, in a more sophisticated way, the line integral of the electric field along a closed path? And what if a magnetic rather than an electric field is present?

Generally speaking, the problem with the emf lies in the diversity of situations where this concept applies, leading to a multitude of corresponding expressions for the emf. The subject is discussed in detail, of course, in all standard textbooks on electromagnetism, both at the intermediate [1-9] and at the advanced [10-12] level. Here we summarize the main ideas, choosing a pedagogical approach that proceeds from the general to the specific. We begin by defining the concept of emf of a "circuit" in the most general way possible. We then apply this definition to certain electrodynamic systems in order to recover familiar expressions for the emf. The main advantage of this approach is that a number of different physical situations are treated in a unified way within a common theoretical framework.

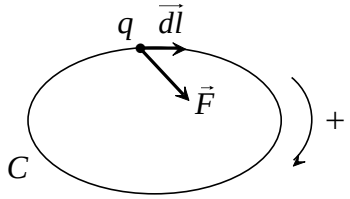
The general definition of the emf is given in Section 2. In subsequent sections (Sec.3-5) application is made to particular cases, such as motional emf, the emf due to a time-varying magnetic field, and the emf of a DC circuit consisting of an ideal battery and a resistor. In Sec.6, the connection between the emf and Ohm's law is discussed.

ISSN:1791-4469

2. THE GENERAL DEFINITION OF EMF

Consider a region of space in which an electromagnetic (e/m) field exists. In the most general sense, any *closed* path *C* (or *loop*) within this region will be called a *"circuit"* (whether or not the whole or parts of *C* consist of material objects such as wires, resistors, capacitors, batteries, or any other elements whose presence may contribute to the e/m field).

We *arbitrarily* assign a positive direction of traversing the loop *C*, and we consider an element $\vec{dl}$ of *C* oriented in the positive direction. Imagine now a test charge *q* located at the position of $\vec{dl}$, and let $\vec{F}$ be the force on *q* at time *t*:



This force is exerted by the e/m field itself, as well as, possibly, by additional energy *sources* (e.g., batteries) that can interact electrically with *q*. The *force per unit charge* at the position of $\vec{dl}$ at time *t*, is

$$\vec{f} = \frac{\vec{F}}{q} \tag{1}$$

Note that $\vec{f}$ is independent of *q*, since the force by the e/m field and/or the sources on *q* is proportional to the charge. In particular, reversing the sign of *q* will have no effect on $\vec{f}$ (although it will change the direction of $\vec{F}$).

We now define the *electromotive force* (*emf*) of the circuit *C* at time *t* as the line integral of $\vec{f}$ along *C*, taken in the *positive* sense of *C*:

$$\mathcal{E} = \oint_C \vec{f} \cdot \vec{dl} \tag{2}$$

Note that the sign of the emf is dependent upon our choice of the positive direction of circulation of *C*: by changing this convention, the sign of $\mathcal{E}$ is reversed.

We remark that, in the *non-relativistic* limit, the emf of a circuit *C* is the same for all inertial observers since *at this limit* the force $\vec{F}$ is invariant under a change of frame of reference.

In the following sections we apply the defining equation (2) to a number of specific electrodynamic situations that are certainly familiar to the student.

3. MOTIONAL EMF IN THE PRESENCE OF A STATIC MAGNETIC FIELD

Consider a circuit consisting of a closed wire *C*. The wire is moving inside a *static* magnetic field $\vec{B}(\vec{r})$. Let $\vec{\upsilon}$ be the velocity of the element $\overrightarrow{dl}$ of *C* relative to our inertial frame of reference. A charge *q* (say, a free electron) at the location of $\overrightarrow{dl}$ executes a composite motion, due to the motion of the loop *C* itself relative to our frame, as well as the motion of *q along C*. The total velocity of *q* relative to us is $\vec{\upsilon}_{tot} = \vec{\upsilon} + \vec{\upsilon}'$, where $\vec{\upsilon}'$ is the velocity of *q* in a direction parallel to $\overrightarrow{dl}$. The force from the magnetic field on *q* is

$$\vec{F} = q\,(\vec{\upsilon}_{tot} \times \vec{B}) = q\,(\vec{\upsilon} \times \vec{B}) + q\,(\vec{\upsilon}' \times \vec{B}) \;\Rightarrow$$
$$\vec{f} = \frac{\vec{F}}{q} = (\vec{\upsilon} \times \vec{B}) + (\vec{\upsilon}' \times \vec{B})$$

By (2), then, the emf of the circuit *C* is

$$\mathcal{E} = \oint_C \vec{f} \cdot \overrightarrow{dl} = \oint_C (\vec{\upsilon} \times \vec{B}) \cdot \overrightarrow{dl} + \oint_C (\vec{\upsilon}' \times \vec{B}) \cdot \overrightarrow{dl}$$

But, since $\vec{\upsilon}'$ is parallel to $\overrightarrow{dl}$, we have that $(\vec{\upsilon}' \times \vec{B}) \cdot \overrightarrow{dl} = 0$. Thus, finally,

$$\mathcal{E} = \oint_C (\vec{\upsilon} \times \vec{B}) \cdot \overrightarrow{dl} \qquad (3)$$

Note that the wire *need not maintain a fixed shape, size or orientation* during its motion! Note also that the velocity $\vec{\upsilon}$ may vary around the circuit.

By using (3), it can be proven (see Appendix) that

$$\mathcal{E} = -\frac{d\Phi}{dt} \qquad (4)$$

where $\Phi = \int \vec{B} \cdot \overrightarrow{da}$ is the magnetic flux through the wire *C* at time *t*. Note carefully that (4) does not express any novel physical law: it is simply a direct consequence of the definition of the emf!

4. EMF DUE TO A TIME-VARYING MAGNETIC FIELD

Consider now a closed wire *C* that is *at rest* inside a *time-varying* magnetic field $\vec{B}(\vec{r},t)$. As experiments show, as soon as $\vec{B}$ starts changing, a current begins to flow in the wire. This looks impressive, given that the free charges in the (stationary) wire were initially at rest. And, as everybody knows, a magnetic field exerts forces on *moving* charges only! It is also observed experimentally that, if the magnetic field $\vec{B}$ stops varying in time, the current in the wire

ISSN:1791-4469

disappears. The only field that can put an initially stationary charge in motion and keep this charge moving is an *electric* field.

We are thus compelled to conclude that *a time-varying magnetic field is necessarily accompanied by an electric field.* (It is often said that "a changing magnetic field *induces* an electric field". This is somewhat misleading since it gives the impression that the "source" of an electric field could be a magnetic field. Let us keep in mind, however, that the true sources of any e/m field are the electric charges and the electric currents!)

So, let $\vec{E}(\vec{r},t)$ be the electric field accompanying the time-varying magnetic field $\vec{B}$. Consider again a charge *q* at the position of the element $\overrightarrow{dl}$ of the wire. Given that the wire is now at rest (relative to our inertial frame), the velocity of *q* will be due to the motion of the charge along the wire only, i.e., in a direction parallel to $\overrightarrow{dl}$: $\vec{\upsilon}_{tot} = \vec{\upsilon}'$ (since $\vec{\upsilon} = 0$). The force on *q* by the e/m field is

$$\vec{F} = q\,[\vec{E} + (\vec{\upsilon}_{tot} \times \vec{B})] = q\,[\vec{E} + (\vec{\upsilon}' \times \vec{B})] \;\Rightarrow$$

$$\vec{f} = \frac{\vec{F}}{q} = \vec{E} + (\vec{\upsilon}' \times \vec{B})$$

The emf of the circuit *C* is now

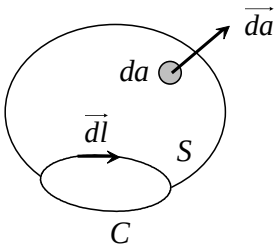
$$\mathcal{E} = \oint_C \vec{f} \cdot \overrightarrow{dl} = \oint_C \vec{E} \cdot \overrightarrow{dl} + \oint_C (\vec{\upsilon}' \times \vec{B}) \cdot \overrightarrow{dl}$$

But, as explained earlier, $(\vec{\upsilon}' \times \vec{B}) \cdot \overrightarrow{dl} = 0$. Thus, finally,

$$\mathcal{E} = \oint_C \vec{E} \cdot \overrightarrow{dl} \qquad (5)$$

Equation (4) is still valid. This time, however, it is not merely a mathematical consequence of the definition of the emf ; rather, it is a true physical law deduced from experiment! Let us examine it in some detail.

In a region of space where a time-varying e/m field $(\vec{E}, \vec{B})$ exists, consider an arbitrary open surface *S* bounded by the closed curve *C*:



(The *relative* direction of $\overrightarrow{dl}$ and the surface element $\overrightarrow{da}$, normal to *S*, is determined according to the familiar right-hand rule.) The loop *C* is assumed stationary relative to the inertial observer; hence the emf along *C* at time *t* is given by (5). The magnetic flux through *S* at this instant is

$$\Phi_m(t) = \int_S \vec{B} \cdot \overrightarrow{da}$$

(Note that the signs of $\mathcal{E}$ and Φ_m depend on the chosen positive direction of *C*.) Since the field $\vec{B}$ is *solenoidal*, the value of Φ_m for a given *C* is independent of the choice of the surface *S*. That is, the same magnetic flux will go through *any* open surface bounded by the closed curve *C*.

According to the *Faraday-Henry law*,

$$\mathcal{E} = -\frac{d\Phi_m}{dt} \tag{6}$$

or explicitly,

$$\oint_C \vec{E} \cdot \overrightarrow{dl} = -\frac{d}{dt}\int_S \vec{B} \cdot \overrightarrow{da} \tag{7}$$

(The negative sign on the right-hand sides of (6) and (7) expresses *Lenz's law*.)

Equation (7) can be re-expressed in differential form by using Stokes' theorem,

$$\oint_C \vec{E} \cdot \overrightarrow{dl} = \int_S (\vec{\nabla} \times \vec{E}) \cdot \overrightarrow{da}$$

and by taking into account that the surface *S* may be arbitrarily chosen. The result is

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{8}$$

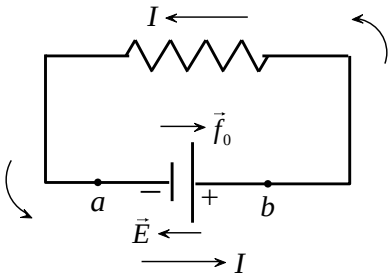
We note that if $\partial\vec{B}/\partial t \neq 0$, then necessarily $\vec{E} \neq 0$. Hence, as already mentioned, a time-varying magnetic field is always accompanied by an electric field. If, however, $\vec{B}$ is *static* ($\partial\vec{B}/\partial t = 0$), then $\vec{E}$ is *irrotational*: $\vec{\nabla} \times \vec{E} = 0 \Leftrightarrow \oint \vec{E} \cdot \overrightarrow{dl} = 0$, which allows for the possibility that $\vec{E} = 0$.

Corollary: The emf around a *fixed* loop *C* inside a *static* e/m field $\left(\vec{E}(\vec{r}), \vec{B}(\vec{r})\right)$ is $\mathcal{E} = 0$ (the student should explain this).

5. EMF OF A CIRCUIT CONTAINING A BATTERY AND A RESISTOR

Consider a circuit consisting of an ideal battery (i.e., one with no internal resistance) connected to an external resistor. As shown below, the emf of the circuit *in the direction of the current* is equal to the voltage *V* of the battery. Moreover, the emf in this case represents the *work per unit charge* done by the source (battery).

ISSN:1791-4469



We recall that, in general, the emf of a circuit *C* at time *t* is equal to the integral

$$\mathcal{E} = \oint_C \vec{f} \cdot \overrightarrow{dl}$$

where $\vec{f} = \vec{F}/q$ is the force per unit charge at the location of the element $\overrightarrow{dl}$ of the circuit, at time *t*. In essence, we assume that in every element $\overrightarrow{dl}$ we have placed a test charge *q* (this could be, e.g., a free electron of the conducting part of the circuit). The force $\vec{F}$ on each *q* is then measured *simultaneously* for all charges at time *t*. Since here we are dealing with a *static* (time-independent) situation, however, we can treat the problem somewhat differently: The measurements of the forces $\vec{F}$ on the charges *q* need not be made at the same instant, given that nothing changes with time, anyway. So, instead of placing several charges *q* around the circuit and measuring the forces $\vec{F}$ on each of them at a particular instant, we imagine *a single charge q* making a complete tour around the loop *C*. We may assume, e.g., that the charge *q* is one of the (*conventionally positive*) free electrons taking part in the constant current *I* flowing in the circuit. We then measure the force $\vec{F}$ on *q* at each point of *C*.

We thus assume that *q* is a *positive* charge moving *in the direction of the current I*. We also assume that the direction of circulation of *C* is the *same as the direction of the current* (counterclockwise in the figure). During its motion, *q* is subject to two forces: (1) the force $\vec{F}_0$ by the source (battery) that carries *q* from the negative pole *a* to the positive pole *b through the source*, and (2) the electrostatic force $\vec{F}_e = q\vec{E}$ due to the electrostatic field $\vec{E}$ at each point of the circuit *C* (both inside and outside the source). The total force on *q* is

$$\vec{F} = \vec{F}_0 + \vec{F}_e = \vec{F}_0 + q\vec{E} \;\Rightarrow\; \vec{f} = \frac{\vec{F}}{q} = \frac{\vec{F}_0}{q} + \vec{E} \equiv \vec{f}_0 + \vec{E}$$

Then,

$$\mathcal{E} = \oint_C \vec{f} \cdot \overrightarrow{dl} = \oint_C \vec{f}_0 \cdot \overrightarrow{dl} + \oint_C \vec{E} \cdot \overrightarrow{dl} = \oint_C \vec{f}_0 \cdot \overrightarrow{dl} \qquad (9)$$

since $\oint_C \vec{E} \cdot \overrightarrow{dl} = 0$ for an electrostatic field. However, the action of the source on *q* is limited to the region between the poles of the battery, that is, the section of the circuit from *a* to *b*. Hence, $\vec{f}_0 = 0$ *outside* the source, so that (9) reduces to

$$\mathcal{E} = \int_a^b \vec{f}_0 \cdot \overrightarrow{dl} \qquad (10)$$

Now, since the current *I* is constant, the charge *q* moves at constant speed along the circuit. This means that the *total* force on *q* in the direction of the path *C* is zero. In the interior of the resistor, the electrostatic force $\vec{F}_e = q\vec{E}$ is counterbalanced by the force on *q* due to the collisions of the charge with the positive ions of the metal (this latter force does *not* contribute to the emf and is *not* counted in its evaluation!). In the interior of the (ideal) battery, however, where there is no resistance, the electrostatic force $\vec{F}_e$ must be counterbalanced by the *opposing* force $\vec{F}_0$ exerted by the source. Thus, in the section of the circuit between *a* and *b*,

$$\vec{F} = \vec{F}_0 + \vec{F}_e = 0 \;\Rightarrow\; \vec{f} = \frac{\vec{F}}{q} = \vec{f}_0 + \vec{E} = 0 \;\Rightarrow\; \vec{f}_0 = -\vec{E}$$

Equation (10) then takes the final form,

$$\mathcal{E} = -\int_a^b \vec{E}\cdot\overrightarrow{dl} = V_b - V_a = V \tag{11}$$

where V_a and V_b are the electrostatic potentials at *a* and *b*, respectively. This is, of course, what every student knows from elementary e/m courses!

The work done by the source on *q* upon transferring the charge from *a* to *b* is

$$W = \int_a^b \vec{F}_0\cdot\overrightarrow{dl} = q\int_a^b \vec{f}_0\cdot\overrightarrow{dl} = q\,\mathcal{E} \tag{12}$$

[where we have used (10)]. So, the *work of the source per unit charge* is $W/q = \mathcal{E}$. This work is converted into heat in the resistor, so that the source must again supply energy in order to carry the charges once more from *a* to *b*. This is something like the torture of Sisyphus in Greek mythology!

6. EMF AND OHM'S LAW

Consider a closed wire *C* inside an e/m field. The circuit may contain sources (e.g., a battery) and may also be in motion relative to our inertial frame of reference. Let *q* be a test charge at the location of the element $\overrightarrow{dl}$ of *C*, and let $\vec{F}$ be the total force on *q* (due to the e/m field and/or the sources) at time *t*. (As mentioned in Sec.2, this force is, classically, a frame-independent quantity.) The force per unit charge at the location of $\overrightarrow{dl}$ at time *t*, then, is $\vec{f} = \vec{F}/q$. According to our general definition, the emf of the circuit at time *t* is

$$\mathcal{E} = \oint_C \vec{f}\cdot\overrightarrow{dl} \tag{13}$$

Now, if σ is the *conductivity* of the wire, then, by *Ohm's law* in its general form (see, e.g., p. 285 of [1]) we have:

$$\vec{J} = \sigma\vec{f} \tag{14}$$

ISSN:1791-4469

where $\vec{J}$ is the *volume current density* at the location of $\overrightarrow{dl}$ at time *t*. (Note that the more common expression $\vec{J}=\sigma\vec{E}$, found in most textbooks, is a special case of the above formula. Note also that $\vec{J}$ is measured *relative to the wire*, thus is the same for all inertial observers.) By combining (13) and (14) we get:

$$\mathcal{E}=\frac{1}{\sigma}\oint_C \vec{J}\cdot\overrightarrow{dl} \tag{15}$$

Taking into account that $\vec{J}$ is in the direction of $\overrightarrow{dl}$ at each point of *C*, we write:

$$\vec{J}\cdot\overrightarrow{dl}=J\,dl=\frac{I}{S}dl$$

where *S* is the *constant* cross-sectional area of the wire. If we make the additional assumption that, at each instant *t*, the current *I* is constant around the circuit (although *I* may vary with time), we finally get:

$$\mathcal{E}=\frac{l}{\sigma S}I=\frac{\rho l}{S}I=IR \tag{16}$$

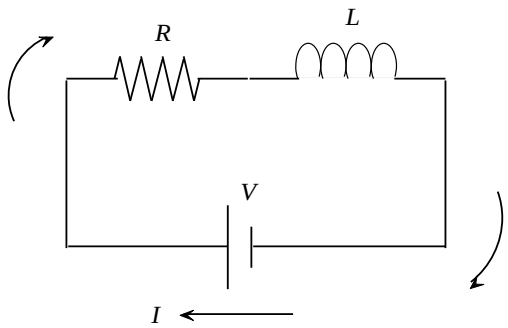
where *l* is the total length of the wire, $\rho=1/\sigma$ is the *resistivity* of the material, and *R* is the total resistance of the circuit. Equation (16) is the familiar special form of Ohm's law.

As an example, let us return to the circuit of Sec.5, this time assuming a *non-ideal* battery with internal resistance *r*. Let R_0 be the external resistance connected to the battery. The total resistance of the circuit is $R=R_0+r$. As before, we call $V=V_b-V_a$ the potential difference between the terminals of the battery, which is equal to the voltage across the external resistor. Hence, $V=IR_0$, where *I* is the current in the circuit. The emf of the circuit (in the direction of the current) is

$$\mathcal{E}=IR=I\,(R_0+r)=V+Ir$$

Note that the potential difference *V* between the terminals *a* and *b* equals the emf only when no current is flowing ($I=0$).

As another example, consider a circuit *C* containing an ideal battery of voltage *V* and having total resistance *R* and total *inductance L*:



In this case, the emf of *C in the direction of the current flow* is

$$\mathcal{E}(t) = V + V_L = V - L\frac{dI}{dt} = I(t)R$$

To understand why the total emf of the circuit is $V+V_L$, we think as follows: On its tour around the circuit, a test charge q is subject to two forces (ignoring collisions with the positive ions in the interior of the wire): a force *inside* the source, and a force by the *non-conservative* electric field accompanying the time-varying magnetic flux through the circuit. Hence, the total emf will be the sum of the emf due to the (ideal) battery alone and the emf expressed by the Faraday-Henry law (6). The latter emf is precisely V_L ; it has a nonzero value for as long as the current I is changing.

Some interesting energy considerations are here in order. The total power supplied to the circuit by the battery at time t is

$$P = IV = I^2R + LI\frac{dI}{dt}$$

The term I^2R represents the power *irreversibly lost* as heat in the resistor (energy, per unit time, spent in moving the electrons through the crystal lattice of the conductor and transferred to the ions that make up the lattice). Thus, this power must necessarily be supplied back by the source in order to *maintain* the current against dissipative losses in the resistor. On the other hand, the term LI (dI/dt) represents the energy per unit time required to *build up* the current against the "back emf" V_L. This energy is *retrievable* and is given back to the source when the current decreases. It may also be interpreted as energy per unit time required in order to *establish the magnetic field* associated with the current. This energy is "stored" in the magnetic field surrounding the circuit.

7. CONCLUDING REMARKS

In concluding this article, let us highlight a few points of importance:

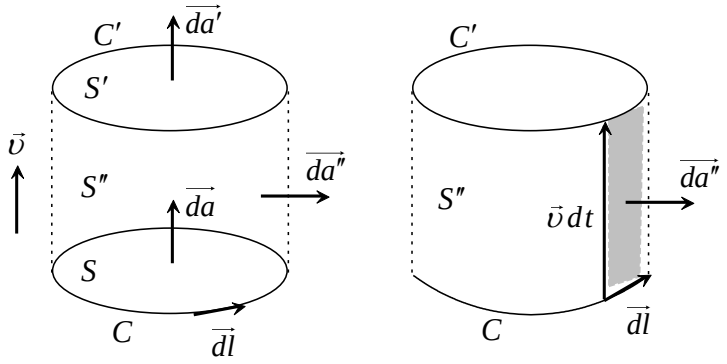
1. The emf was defined as a line integral of force per unit charge around a loop (or "circuit") in an e/m field. The loop may or may not consist of a real conducting wire, and it may contain sources such as batteries.
2. In the classical (non-relativistic) limit, the emf is independent of the inertial frame of reference with respect to which it is measured.
3. In the case of *purely motional* emf, Faraday's "law" (4) is in essence a mere consequence of the definition of the emf. On the contrary, when a time-dependent magnetic field is present, the similar-looking equation (6) is a true physical law (the Faraday-Henry law).
4. In a DC circuit with a battery, the emf in the direction of the current equals the voltage of the battery and represents work per unit charge done by the source.
5. If the loop describing the circuit represents a conducting wire of finite resistance, Ohm's law can be expressed in terms of the emf by equation (16).

ISSN:1791-4469

APPENDIX

Here is an analytical proof of equation (4) of Sec.3:

Assume that, at time *t*, the wire describes a closed curve *C* that is the boundary of a plane surface *S*. At time *t´*= *t*+*dt*, the wire (which has moved in the meanwhile) describes another curve *C´* that encloses a surface *S´*. Let $\vec{dl}$ be an element of *C* in the direction of circulation of the curve, and let $\vec{\upsilon}$ be the velocity of this element relative to an inertial observer (the velocity of the elements of *C* may vary along the curve):



The direction of the surface elements $\vec{da}$ and $\vec{da'}$ is consistent with the chosen direction of $\vec{dl}$, according to the right-hand rule. The element of the side ("cylindrical") surface *S´´* formed by the motion of *C*, is equal to

$$\vec{da''} = \vec{dl} \times (\vec{\upsilon}\, dt) = (\vec{dl} \times \vec{\upsilon})\, dt$$

Since the magnetic field is static, we can view the situation in a somewhat different way: Rather than assuming that the curve *C* moves within the time interval *dt* so that its points coincide with the points of the curve *C´* at time *t´*, we consider two *constant* curves *C* and *C´ at the same instant t*. In the case of a *static* field $\vec{B}$, the magnetic flux through *C´* at time *t´*= *t*+*dt* (according to our original assumption of a moving curve) is the same as the flux through this same curve at time *t*, given that no change of the magnetic field occurs within the time interval *dt*. Now, we note that the open surfaces S_1=*S* and S_2= *S´* ∪ *S´´* share a common boundary, namely, the curve *C*. Since the magnetic field is *solenoidal*, the same magnetic flux Φ_m passes through S_1 and S_2 at time *t*. That is,

$$\int_{S_1} \vec{B} \cdot \vec{da}_1 = \int_{S_2} \vec{B} \cdot \vec{da}_2 \;\Rightarrow\; \int_S \vec{B} \cdot \vec{da} = \int_{S'} \vec{B} \cdot \vec{da'} + \int_{S''} \vec{B} \cdot \vec{da''}$$

But, returning to our initial assumption of a *moving* curve, we note that

$$\int_S \vec{B} \cdot \vec{da} = \Phi_m(t) = \textit{magnetic flux through the wire at time } t$$

and

$\int_{S'} \vec{B} \cdot \overrightarrow{da'} = \Phi_m(t+dt) =$ *magnetic flux through the wire at time t+dt*

Hence,

$$\Phi_m(t) = \Phi_m(t+dt) + \int_{S''} \vec{B} \cdot \overrightarrow{da''} \Rightarrow$$

$$d\Phi_m = \Phi_m(t+dt) - \Phi_m(t) = -\int_{S''} \vec{B} \cdot \overrightarrow{da''} = -dt \oint_C \vec{B} \cdot (\overrightarrow{dl} \times \vec{\upsilon}) \Rightarrow$$

$$-\frac{d\Phi_m}{dt} = \oint_C \vec{B} \cdot (\overrightarrow{dl} \times \vec{\upsilon}) = \oint_C (\vec{\upsilon} \times \vec{B}) \cdot \overrightarrow{dl} = \mathcal{E}$$

in accordance with (3) and (4).

REFERENCES

1. D. J. Griffiths, *Introduction to Electrodynamics*, 3rd Edition (Prentice-Hall, 1999).
2. R. K. Wangsness, *Electromagnetic Fields*, 2nd Edition (Wiley, 1986).
3. J. R. Reitz, F. J. Milford, R. W. Christy, *Foundations of Electromagnetic Theory*, 4th Edition (Addison-Wesley, 1993).
4. W. N. Cottingham, D. A. Greenwood, *Electricity and Magnetism* (Cambridge, 1991).
5. P. Lorrain, D. R. Corson, F. Lorrain, *Electromagnetic Fields and Waves*, 3rd Edition (Freeman, 1988).
6. V. Rojansky, *Electromagnetic Fields and Waves* (Dover, 1979).
7. A. Shadowitz, *The Electromagnetic Field* (McGraw-Hill, 1975).
8. C. J. Papachristou, *Introduction to Electromagnetic Theory and the Physics of Conducting Solids*, 2nd Edition, in Greek (Naval Academy of Greece Publications, 2010).
9. A. N. Magoulas, *Applied Electromagnetism*, 2nd Edition, in Greek (Naval Academy of Greece Publications, 2012).
10. J. D. Jackson, *Classical Electrodynamics*, 3rd Edition (Wiley, 1999)[1].
11. W. K. H. Panofsky, M. Phillips, *Classical Electricity and Magnetism,* 2nd Edition (Addison-Wesley, 1962).
12. W. Greiner, *Classical Electrodynamics* (Springer, 1998).

[1] One of us (C.J.P.) strongly feels that the 2nd Edition of 1975 (unfortunately out of print) was a much better edition!

Quantum Transmission via Resonant Tunneling in a Double Barrier Structure: a Path Integral Approach.

Theodosios G. Douvropoulos

Hellenic Naval Academy, Section of Natural Sciences,
Hatzikyriakou Ave. Piraeus, 18539, Greece
e-mail: douvrotheo@snd.edu.gr, douvrotheo@yahoo.com

Abstract. We study both the symmetric and asymmetric one-dimensional double barrier potential which describes the band profile of a quantum tunneling diode in the absence or the application respectively, of a constant electrical field. The semiclassical path integral theory is employed to determine the transmission amplitude, which is the Green's function for a single charge transport. The Green's function is given in an analytical form and some attributes of the transmission amplitude due to resonant tunneling are established while it is shown to exhibit maxima at the position of the resonant states. The complex spectrum of the heterostructure is revealed and the time scale for charge transport is given, both in an analytical form.

Keywords: double barrier structure, resonant tunneling diode, quasi bound states, Green's function, transmission amplitude, charge transport, transport time

PACS: 03.65.Sq, 03.65.Xp, 05.60.Gg, 72.10.Bg, 72.23.Hk

INTRODUCTION

Quantum transmission in nanostructures is in general dominated by the tunneling phenomenon. This term describes a particle transport through a classically forbidden region of motion, (barrier), meaning a region in which the potential energy exceeds the total particle energy. There is no classical analogue, since in classical mechanics the particle is totally reflected at the so called classical turning points, that is, points where the total energy equals the potential energy, and therefore no transmission ever occurs. The quantum mechanical wavefunction instead, does not vanish inside and after the barrier. Thus, according to the laws of quantum mechanics a particle incident on a potential barrier has a finite probability of appearing on the other side.

Tunneling in solids was first studied by Fowler and Nordheim [1] in the thermionic emission of electrons from metal into vacuum. Later, interest was taken in the study of tunneling through thin insulating layers, separating two metals, or a metal and a semiconductor. Zener [2] introduced the interband tunneling, describing electrons that tunnel from one energy band to another through the forbidden energy gap. The outstanding breakthroughs in the area of semiconductor device technology that followed, made possible the experimental observation of Zener

ISSN:1791-4469

tunneling in p-n junction diodes. For example, Esaki [3] introduced the so-called Esaki diode, in which the interband Zener tunneling, produces negative differential resistance in the I-V characteristics.

In most artificially engineered structures, quantum confinement may seriously reduce the dimensions of the system under study. Such confinement is usually caused by a heterojunction, (MOSFET), or simply a semiconductor- air interface, (quantum wire). For example if one thin layer of material is grown on top of another, such as the simple AlGaAs/GaAS heterostructure, the change in potential is in only the vertical growth direction, and therefore the problem practically becomes one dimensional. The ability of constructing well controlled heterostructure layers enabled Tsu and Esaki to predict at first and observe shortly after [4,5], that when bias is applied across a double barrier heterostructure we get similar to the Esaki diode current voltage characteristics. However in this case it is the resonant tunneling, which is tunneling through the barriers within the same band, that is responsible for the I/V characteristics, and not the interband tunneling. Resonant tunneling refers to the case where the transmission amplitude, which is the Green's function for electron propagation through the heterostructure, is sharply peaked about certain complex values of the energy. In fact the real parts of the above mentioned complex values are very close to these of the bound states associated with the quantum well formed between the two confining barriers. The resonant energies of such a heterostructure support a complex spectrum, due to the fact that the electron may escape away from the quantum well in either direction. Thus, there is a finite lifetime associated with the bound state. This is why these states are known as quasi-bound states. The Green's function for the lowest resonant energy may approach unity in some cases and so semiclassical methods can be applied with high accuracy.

In our days the numerical calculation of the Green's function, for systems consisting of heterostructures, can be done with a relative ease, with the aid of modern computers. However, an analytical solution is always desirable and of an instructive value. For example the analytic solution provides a direct comparison between the properties of different systems, even coming from different branches of science, as long as they can be described by the same type of potential function. In addition the application of the path integral analytic formalism seems to lack of any previous experience on these systems. Thus, our work was motivated by the need of fulfilling both the above requirements.

The purpose of this paper is to describe and further produce analytic relations for a double barrier heterostructure, via an analytic path integral formalism. Doing so we first demonstrate the importance of the double barrier structure through its close resemblance to the structure of a resonant tunneling diode. We present a widely used structure which consists of two AlGaAs barriers, (speaking more accurately we should write AlxGa1-xAs where x~0,3), separated by a thin GaAs quantum well, surrounded by heavily doped GaAs layers. Then the semiclassical path integral approach is developed and the transmission amplitude for electron transport through the heterostructure is analytically calculated. We also produce analytic relations for the complex energy spectrum supported by the double barrier structure, as well as for the time needed for charge transport.

2. THE DOUBLE BARRIER POTENTIAL AS A RESONANT TUNNELING DIODE

We consider the potential of figure 1 that follows. Potentials with such a form constitute models for resonant tunneling diodes and tunneling processes in various systems of physics and chemistry. As shown in figure 1, points ε and ζ correspond to the maximum of the barriers which they do not need to be symmetric. There are four turning points: at α and β (finite barrier,

tunnelling is allowed) and at γ and *δ*, (again finite barrier, tunnelling is allowed), while β and γ also define the limits of the classically allowed region of motion.

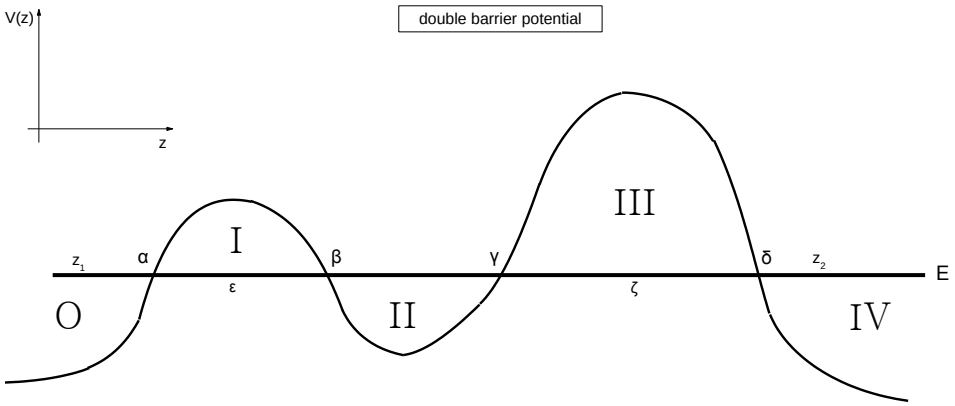


FIGURE 1: The double barrier potential *V(z)* that couples a quantum well to the continuum via two channels.

The topology of the above potential, meaning the specific permutation of classi-cally allowed and forbidden regions of motion, is the same with the one of the effective potential (band profile) of a resonant tunneling diode as this is pictured in figure 2 that follows. The barrier height is due to the conduction band offset, ΔE_c, while E_f refers to the Fermi energy of the heavily doped GaAs layers and E_o refers to the lowest resonant energy or quasibound state of the GaAs quantum well. The thickness of the AlGaAS barriers is such that allows tunneling to be significant.

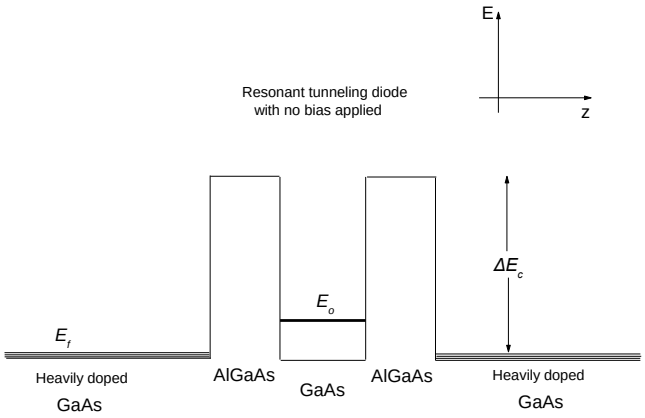


FIGURE 2: Effective potential or band profile of a resonant tunneling diode structure.

When a positive bias *V* is applied to the right contact relative to the left, the Fermi energy on the left is raised to the resonant energy E_o and a large current flows from left to right due to the maximization of the transmission amplitude. Opposite charge flow is strongly suppressed, since the carriers at the Fermi energy on the right, feel a large potential barrier, as shown in figure 3 that follows:

ISSN:1791-4469

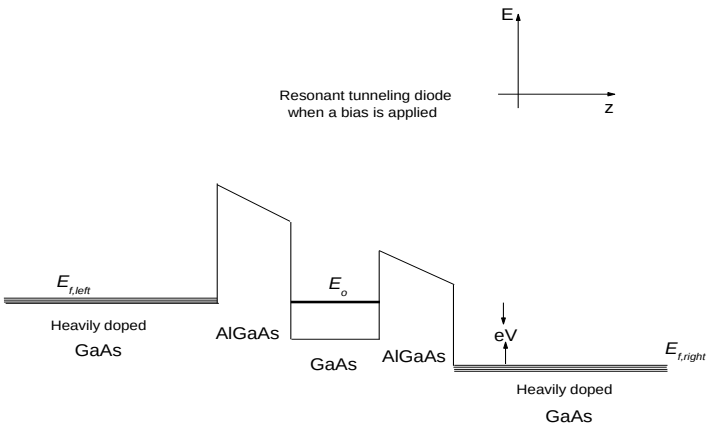


FIGURE 3: Effective potential or band profile for a resonant tunneling diode structure when a positive bias is applied to the right contact.

However, if we further increase the applied bias, the bottom of the conduction band of the left contact is raised to the resonant energy E_o, and therefore they will not remain anymore available electrons for tunneling. This is why the current is decreased with increasing bias, which results in a region of negative differential resistance in the I/V characteristics. Comparing figures 2 and 3 it is readily seen that the application of a bias destroys the symmetry between the two potential barriers.

During the last decades, there has been a considerable volume of research into resonant tunneling diodes. Besides the fundamental physics included into this simple structure, interest also stems from the various applications in microwave systems and digital logic circuits. In these systems we may be interested in the fundamental time associated with the tunneling process, which is often taken as the lifetime of the quasi – bound state. Besides this time constant, we should also have in mind the RC time constant due to the capacitance of the structure and the transit time across the nontunneling regions of the device as well. However, when the device is properly designed these two time constants can be minimized. According to the literature the transfer- matrix method based on the discretized form of the one dimensional *Schrödinger* equation, seems to be the most popular for analyzing the double barrier structure, see [6-11] and references therein. However our approach will be different, following the semiclassical path integral method which does not involve the *Schrödinger* equation at all.

3. THE SEMICLASSICAL PATH INTEGRAL APPROACH

The path integral construction of a system's Green's function for one dimensional propagation between two points z_1 and z_2 , is accomplished by taking account of all possible changes in phase of the wavefunction. These may be due to motion inside allowed and forbidden region or due to reflection from turning points. Miller [12] developed a semiclassical periodic orbit theory, based on Gutzwiiller's trace formula [13], in his pioneering work on the application of path integrals to tunneling. This work inspired many others to improve his method and analytically solve interesting one dimensional problems.

Among those, Holstein and Swift [14] and Holstein alone [15,16] 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. Holstein's [15] central result for the calculation of the

transmission amplitude via an infinite set of paths that the particle follows, connecting the initial point z_1 to the final point z_2, can be written in compact form as

$$Gsc(E) = \frac{1}{\overline{p(z_1,z_2)}} \sum_{j=1}^{\infty} \left\{ \prod_{i=1}^{N(j)} a_{ij} \right\} \quad (1)$$

In the above equation $\overline{p(z_1,z_2)}$ is a non local momentum of the particle defined by $\overline{p(z_1,z_2)} = 2\pi\sqrt{k^*(z_1)k^*(z_2)}$ where $k^*(z) = \sqrt{2(E-V(z))}$, with *E* standing for the energy and *V(z)* for the potential function, (atomic units employed throughout). In equation (1) each path is uniquely identified by an index, *j*, with *j* = 1,2,3,…..∞ , and is constructed by a set of factors α_{ij} corresponding to propagation (from α to β) in allowed regions (given by $\exp[i\int_{\alpha}^{\beta} k^*(z)dz] \equiv e^{ik}$), or in forbidden regions (given by $\exp\left[-\left|\int_{\alpha}^{\beta} \kappa^*(z)dz\right|\right] \equiv e^{-\kappa}$ where $\kappa^*(z) = \sqrt{2(V(z)-E)}$), and on the corresponding 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). The product $\prod_{i=1}^{N(j)} a_{ij}$ gives the unique amplitude for each possible path for going from z_1 to z_2. For each path, *j*, *N(j)* is the number of possible factors a_{ij} present in this path. In the picture that follows we schematically give the rules for the propagation-reflection factors α_{ij} that constitute the fundamental cells for constructing each individual path amplitude.

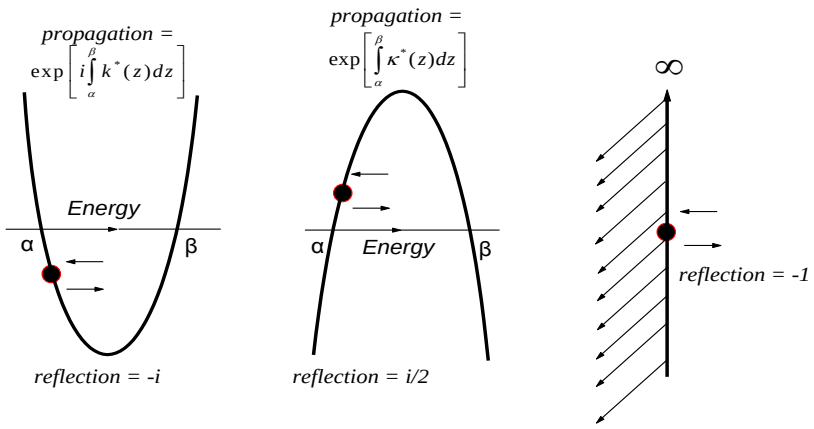


FIGURE 4: Rules for the construction of the path integral amplitudes through the α_{ij} factors. The first motion is in a classically allowed region of motion, the second in a classically forbidden region and the third motion includes reflection from an infinite barrier.

The above described method can also be found in standard textbooks of path integrals, or quantum tunneling, [17,18].

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

ISSN:1791-4469

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 z_1 and z_2 lie on the heavily doped GaAs layers, on each free side of the two barriers, as shown in figure 1.

When the potential is known explicitly, then *Gsc(E)* can be obtained analytically or numerically and its simple pole structure is revealed. Once *Gsc(E)* is known, its complex '*self-energy*' correction, namely the energy shift and width, can be extracted. This is possible by comparing the result representing the corresponding bound state problem for the inner quantum well (real energies) to the result for the resonance state problem (complex energies). As will be seen, *Gsc(E)* is a finite sum of complex poles, each representing a resonance state of the potential. The calculation leads naturally to the result that the pole of interest has a negative imaginary part, i.e. it corresponds to a decaying state, associated with the time needed for charge transport.

4. TRANSMISSION AMPLITUDE AND COMPLEX ENERGY SPECTRUM OF A DOUBLE BARRIER STRUCTURE

Considering the potential of figure 1, our objective is the calculation of the Greens function for propagation between points z_1 and z_2 via the semiclassical path integral theory. The calculation of Gsc(E) according to equation (1) entails the consideration of the phases of all possible paths, for the given total energy E. For each path, the overall phase is determined by the manner in which allowed and forbidden regions succeed one another, and by the nature of the corresponding turning points when reflection occurs. With respect to figure 1, the motion of the particle for E > 0 starts at (z_1, t_1) and ends at (z_2, t_2), at each side of the two potential barriers. So there are five regions: O (allowed), I (forbidden), II (allowed), III(forbidden), IV(allowed) and motion in O and IV is free: once the particle moves from point α to the left or from point δ to the right, it cannot be reflected. In this way we are actually interested in the calculation of the amplitude for travelling from α at the beginning of region I, to δ at the end of region III in all possible ways.

Since the sum over histories consists of calculating all possible paths connecting these points, we should spend a few lines explaining the symbolism that follows: for example $A^{III}_{\gamma\to\delta}$ means the amplitude for all the possible paths connecting γ and δ, while always staying inside region III, and $\left(I-II^{*}\right)^{\alpha\to\gamma}$ means the amplitude for all the possible paths connecting points α and γ by interchanging regions I and II, with the motion always ending inside region II, (where the asterisk goes). We can now proceed to the calculation by first dividing the problem up into smaller parts as follows:

we reach point γ without ever passing through region III by interchanging only regions I and II in all possible ways. We call this contribution C_1

$$C_1=\left(I-II^{*}\right)^{\alpha\to\gamma} \tag{2}$$

Let us give a graphical presentation of a typical path of the above contribution C_1, for the potential of figure 1. In all the graphs that follow the solid lines correspond to propagation in a classically allowed region of motion while the dash- dot lines correspond to propagation in a classically forbidden region of motion. Reflections are described by small curved lines and the succession of lines moves downwards as time passes.

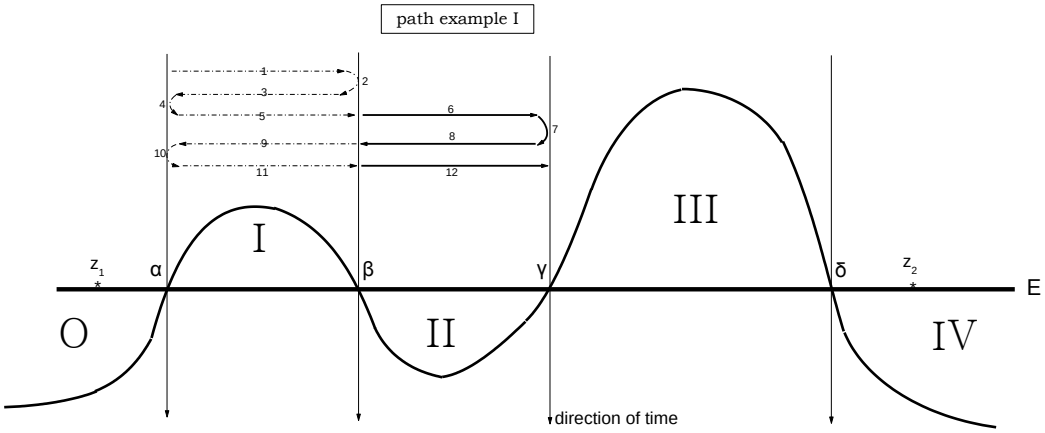


FIGURE 5: Graph representation of a typical path of contribution C_1 for the potential of figure 1. Propagation starts at α and ends at γ.

The above path is constituted by N(j)=12 propagation-reflection factors, and its contribution to the construction of the transmission amplitude is analytically given by the following time ordered product

$$a_j^{12} = e^{-\kappa_1}\frac{i}{2}e^{-\kappa_1}\frac{i}{2}e^{-\kappa_1}e^{ik}(-i)e^{ik}e^{-\kappa_1}\frac{i}{2}e^{-\kappa_1}e^{ik} = -\frac{1}{8}e^{-5\kappa_1}e^{3ik} \tag{3}$$

where $e^{ik} = \exp[i\int_{\beta}^{\gamma} k^*(z)dz]$ and where and where $e^{-\kappa_{1,2}} = \exp\left[-\left|\int_{\alpha,\gamma}^{\beta,\delta}\kappa^*(z)dz\right|\right]$ and j stands for the index of the path.

with γ as the starting point, we have the following alternatives: doing nothing, this is just a unity factor, or combine the three regions in all possible ways by keeping as last interchange i) that of regions I and II, C_2, or ii) that of regions III and II, C_3 , getting respectively

$$C_2 = \left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma} + \left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}\left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma} + \ldots\ldots \tag{4a}$$

and

$$C_3 = \left(III - II^*\right)^{\gamma\to\gamma} + \left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}\left(III - II^*\right)^{\gamma\to\gamma} + \left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}\left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}\left(III - II^*\right)^{\gamma\to\gamma} + \ldots\ldots \tag{4b}$$

It is obvious that each of the above contributions forms the infinite sum of the terms of a geometric progression, and so we can actually reduce the above sums to the following compact formulae

$$C_2 + 1 = \frac{1}{1 - \left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}} \tag{4c}$$

and

$$C_3 = \frac{\left(III - II^*\right)^{\gamma\to\gamma}}{1-\left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}} \tag{4d}$$

In fact we can also sum together C_2 +1 and C_3 to get

$$C_2 + 1 + C_3 = C_{2,3} = \frac{1+\left(III - II^*\right)^{\gamma\to\gamma}}{1-\left(III - II^*\right)^{\gamma\to\beta}\left(I - II^*\right)^{\beta\to\gamma}} \tag{5}$$

Let us again give a graphical presentation of a typical path of the above contribution $C_{2,3}$, for the potential of figure 1.

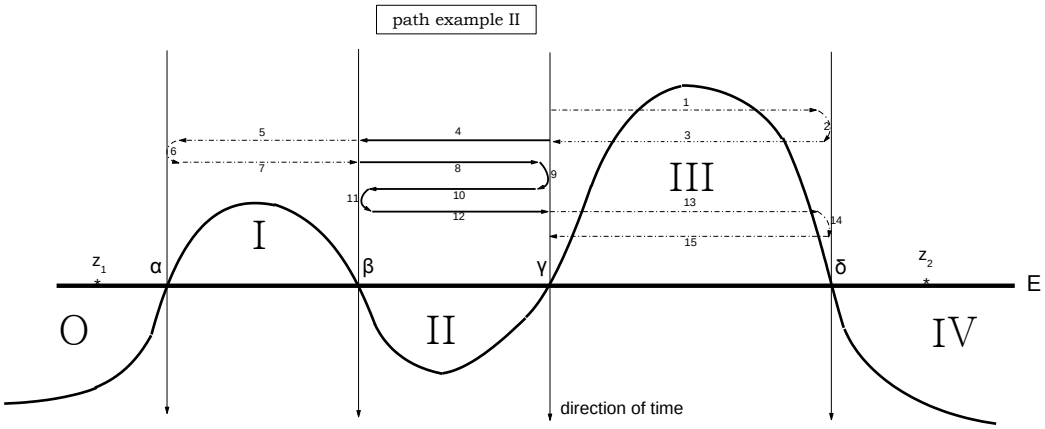


FIGURE 6: Graph representation of a typical path of contribution $C_{2,3}$ for the potential of figure 1. Propagation starts at γ and ends at γ.

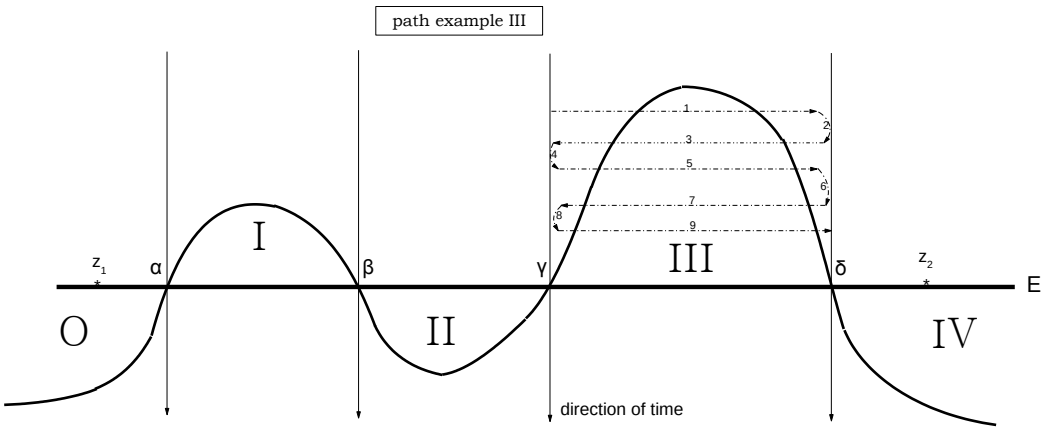


FIGURE 7: Graph representation of a typical path of contribution C_4 for the potential of figure 1. Propagation starts at γ and ends at δ.

The above path is constituted by *N(j)*=15 propagation-reflection factors, and its contribution to the construction of the transmission amplitude is analytically given by the following time ordered product

$$b_j^{15} = e^{-\kappa_2}\frac{i}{2}e^{-\kappa_2}e^{ik}e^{-\kappa_1}\frac{i}{2}e^{-\kappa_1}e^{ik}(-i)e^{ik}(-i)e^{ik}e^{-\kappa_2}\frac{i}{2}e^{-\kappa_2} = \frac{i}{8}e^{-2\kappa_1}e^{-4\kappa_2}e^{4ik} \tag{6}$$

the last contribution involves only region *III* and no other, so it is simply the amplitude

$$C_4 = A^{III}_{\gamma\to\delta} \tag{7}$$

We give a last graphical presentation of a typical path of the contribution C_4, for the potential of figure 1.

The above path is constituted by *N(j)*=9 propagation-reflection factors, and its contribution to the construction of the transmission amplitude is analytically given by the following time ordered product

$$c_j^9 = e^{-\kappa_2}\frac{i}{2}e^{-\kappa_2}\frac{i}{2}e^{-\kappa_2}\frac{i}{2}e^{-\kappa_2}\frac{i}{2}e^{-\kappa_1} = -\frac{1}{16}e^{-5\kappa_2} \tag{8}$$

We can now proceed to the calculation of the above contributions in terms of simpler amplitudes and transmission reflection factors.

$$C_1 = \left(I - II^*\right)^{\alpha\to\gamma} = A^I_{\alpha\to\beta}A^{II}_{\beta\to\gamma}\begin{Bmatrix}1+(-i)e^{ik}A^I_{\beta\to\beta}A^{II}_{\beta\to\gamma}+\\(-i)e^{ik}A^I_{\beta\to\beta}A^{II}_{\beta\to\gamma}(-i)e^{ik}A^I_{\beta\to\beta}A^{II}_{\beta\to\gamma}+....\end{Bmatrix} \tag{9}$$

Again we recognize the sum of a geometric progression and we can write the above expression in the closed form that follows

$$\begin{aligned} C_1 = \left(I - II^*\right)^{\alpha\to\gamma} &= A^I_{\alpha\to\beta}A^{II}_{\beta\to\gamma}\frac{1}{1+ie^{ik}A^I_{\beta\to\beta}A^{II}_{\beta\to\beta}} \\ &= \frac{A^I_{\alpha\to\beta}A^{II}_{\beta\to\gamma}}{1-A^I_{\beta\to\beta}A^{II}_{\beta\to\beta}} \end{aligned} \tag{10}$$

It is then straightforward to prove that the following relation also holds

$$\left(I - II^*\right)^{\beta\to\gamma} = \frac{A^I_{\beta\to\beta}A^{II}_{\beta\to\gamma}}{1-A^I_{\beta\to\beta}A^{II}_{\beta\to\beta}} \tag{11}$$

We can now proceed to the combination of regions II and III and calculate the following contributions

ISSN:1791-4469

$$\left(III-II^{*}\right)^{\gamma\to\gamma}=A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}\left\{\begin{matrix}1+A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}+\\ A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}+....\end{matrix}\right\}$$
$$=\frac{A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}}{1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}} \quad (12)$$

where again it is straightforward to also calculate

$$\left(III-II^{*}\right)^{\gamma\to\beta}=\frac{A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\beta}}{1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}} \quad (13)$$

We now take in account of all of the above contributions by multiplying the factors C_1 $C_{2,3}$ and C_4, since they correspond to statistically independent events, and get the final result

$$C_{total}=C_1\ \mathrm{x}\ C_{2,3}\ \mathrm{x}\ C_4=$$

$$\frac{A^{I}_{\alpha\to\beta}A^{II}_{\beta\to\gamma}}{1-A^{I}_{\beta\to\beta}A^{II}_{\beta\to\beta}}\ \mathrm{x}\ \frac{1+\dfrac{A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}}{1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}}}{1-\dfrac{A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\beta}}{1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}}\dfrac{A^{I}_{\beta\to\beta}A^{II}_{\beta\to\gamma}}{1-A^{I}_{\beta\to\beta}A^{II}_{\beta\to\beta}}}\ \mathrm{x}\ A^{III}_{\gamma\to\delta}= \quad (14)$$

$$\frac{A^{I}_{\alpha\to\beta}A^{II}_{\beta\to\gamma}A^{III}_{\gamma\to\delta}}{\left(1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}\right)\left(1-A^{I}_{\beta\to\beta}A^{II}_{\beta\to\beta}\right)-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\beta}A^{I}_{\beta\to\beta}A^{II}_{\beta\to\gamma}}$$

In order to complete the above formula we should also include the initial propagation from z_1 to α which is simply the $A^{O}_{z_1\to a}$ and the last propagation from δ to z_2 which is the $A^{IV}_{\delta\to z_2}$. In this way equation (1) of the Green's function is written as follows

$$G_{sc}(E)\ =\ \frac{A^{O}_{z_1\to a}A^{IV}_{\delta\to z_2}}{p(z_1,z_2)}\frac{A^{I}_{\alpha\to\beta}A^{II}_{\beta\to\gamma}A^{III}_{\gamma\to\delta}}{\left(1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}\right)\left(1-A^{I}_{\beta\to\beta}A^{II}_{\beta\to\beta}\right)-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\beta}A^{I}_{\beta\to\beta}A^{II}_{\beta\to\gamma}} \quad (15)$$

Since we have given the general expression of the Green's function we can now discuss the type of symmetry between the two potential barriers. We should note that independently of the type of symmetry it is always true that $A^{II}_{\gamma\to\gamma}=A^{II}_{\beta\to\beta}$, $A^{II}_{\beta\to\gamma}=A^{II}_{\gamma\to\beta}$ and $A^{II}_{\beta\to\beta}=A^{II}_{\gamma\to\gamma}$.

Symmetric barriers case:

It is clear that if we are dealing with symmetric potential barriers then it is true that $A^{III}_{\gamma\to\gamma}=A^{I}_{\beta\to\beta}$ and $A^{I}_{\alpha\to\beta}=A^{III}_{\gamma\to\delta}$. The total path contribution from α to δ can then be written as

$$C_{sym}=\frac{A^{I}_{\alpha\to\beta}A^{II}_{\beta\to\gamma}A^{III}_{\gamma\to\delta}}{\left(1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}\right)^{2}-\left(A^{III}_{\gamma\to\gamma}A^{II}_{\beta\to\gamma}\right)^{2}} \tag{16a}$$

Let us define $1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}\equiv e$ and $A^{III}_{\gamma\to\gamma}A^{II}_{\beta\to\gamma}\equiv d$. We then have

$$A^{I}_{\alpha\to\beta}A^{II}_{\beta\to\gamma}A^{III}_{\gamma\to\delta}=A^{I}_{\alpha\to\beta}\left[A^{II}_{\beta\to\gamma}A^{III}_{\gamma\to\delta}\frac{i}{2}e^{-\kappa}\right]\frac{2}{i}e^{\kappa}=-2ie^{\kappa}A^{I}_{\alpha\to\beta}d \tag{16b}$$

where

$$e^{-\kappa}=\exp\left[-\left|\int_{\alpha}^{\beta}\kappa^{*}(z)dz\right|\right]=\exp\left[-\left|\int_{\gamma}^{\delta}\kappa^{*}(z)dz\right|\right] \tag{16c}$$

In this way the overall contribution can be written as

$$C_{sym.}=-ie^{\kappa}A^{I}_{\alpha\to\beta}\left(\frac{1}{e-d}-\frac{1}{e+d}\right) \tag{17}$$

In order to further construct analytic relations we should calculate the fundamental amplitudes that appear in the above formulae in terms of the transmission factors.

$$\begin{aligned}A^{I}_{\alpha\to\beta}&=e^{-\kappa}\left(1+\frac{i}{2}e^{-\kappa}\frac{i}{2}e^{-\kappa}+\ldots\right)\\&=\frac{e^{-\kappa}}{1+e^{-2\kappa}/4}\end{aligned} \tag{18a}$$

$$\begin{aligned}A^{II}_{\beta\to\gamma}&=e^{ik}\left(1+(-i)e^{ik}(-i)e^{ik}+\ldots\right)\\&=\frac{e^{ik}}{1+e^{2ik}}\end{aligned} \tag{18b}$$

$$\begin{aligned}A^{II}_{\gamma\to\gamma}&=e^{ik}(-i)e^{ik}\left(1+(-i)e^{ik}(-i)e^{ik}+\ldots\right)\\&=-i\frac{e^{2ik}}{1+e^{2ik}}\end{aligned} \tag{18c}$$

$$\begin{aligned}A^{III}_{\gamma\to\gamma}&=e^{-\kappa}\frac{i}{2}e^{-\kappa}\left(1+\frac{i}{2}e^{-\kappa}\frac{i}{2}e^{-\kappa}+\ldots\right)\\&=\mathrm{i}/2\frac{e^{-2\kappa}}{1+e^{-2\kappa}/4}\end{aligned} \tag{18d}$$

We can now proceed to the calculation the denominators of equation (17). We will have

$$\begin{aligned}&e\mp d=1-A^{III}_{\gamma\to\gamma}A^{II}_{\gamma\to\gamma}\mp A^{III}_{\gamma\to\gamma}A^{II}_{\beta\to\gamma}=\\&1-\mathrm{i}/2\left(\frac{e^{-2\kappa}}{1+e^{-2\kappa}/4}\right)\mathrm{x}\left(-i\frac{e^{2ik}}{1+e^{2ik}}\right)\mp\mathrm{i}/2\left(\frac{e^{-2\kappa}}{1+e^{-2\kappa}/4}\right)\mathrm{x}\left(\frac{e^{ik}}{1+e^{2ik}}\right)\end{aligned} \tag{19}$$

ISSN:1791-4469

We introduce the following two symbols, corresponding to the forbidden and the allowed region of motion respectively:

$$\rho \equiv e^{-2\kappa}/4 \qquad \text{and} \qquad \sigma \equiv e^{2ik} \tag{20}$$

and write the overall contribution as

$$C_{sym.} = -ie^{\kappa} A^{I}_{\alpha\to\beta}\left(\frac{(1+\rho)(1+\sigma)}{(1+\rho)(1+\sigma)-2\rho\sigma-2i\rho\sigma^{1/2}} - \frac{(1+\rho)(1+\sigma)}{(1+\rho)(1+\sigma)-2\rho\sigma+2i\rho\sigma^{1/2}}\right) \tag{21}$$

Direct remarks can be expressed concerning the structure of the above formulae:

a) It is obvious that when $\rho << 1$ which is equivalent to say that we have strong barriers confining the inner classically allowed region, and so practically no interaction with the continuum appears, we get

$$C_{sym.} \cong -ie^{\kappa} A^{I}_{\alpha\to\beta}\left(\frac{(1+\sigma)}{(1+\sigma)} - \frac{(1+\sigma)}{(1+\sigma)}\right) = 0 \tag{22}$$

b) It is easy to see that the condition $1+\sigma = 0$ determines the eigenvalues of the classically allowed region of motion, since we can actually write $e^{2ik} = e^{\pm i\pi} \Rightarrow k = n\pi + \pi/2, \quad n = 0,1,2,..$ which is the Bohr Sommerfield quantization rule. For example if the area of confinement is a harmonic potential of the form $U(x) = \frac{1}{2}\omega^2 x^2$ with turning points at α and –α, meaning $\mathrm{E} = \frac{1}{2}\omega^2\alpha^2$, k is equal to the quantity $\sqrt{2}\int_{-\alpha}^{\alpha} dx\sqrt{\mathrm{E} - \frac{1}{2}\omega^2 x^2} = \sqrt{2}\sqrt{\frac{2}{\omega^2}}\frac{\mathrm{E}}{2}\sin^{-1}\frac{x}{\alpha}\Big|_{-\alpha}^{\alpha} = \frac{E\pi}{\omega}$, and then the quantization rule gives $\frac{\pi E_n}{\omega} = n\pi + \frac{\pi}{2} \Rightarrow E_n = (n+1/2)\omega$, which is the harmonic oscillator spectrum.

We may now investigate the effect of the previously revealed eigenvalues on the structure of the semiclassical Greens function.

$$C_{sym.}(E_n) = -ie^{\kappa} A^{I}_{\alpha\to\beta}\left(\frac{(1+\rho)(0)}{(1+\rho)(0)+2\rho-2i\rho(-1)^n i} - \frac{(1+\rho)(0)}{(1+\rho)(0)+2\rho+2i\rho(-1)^n i}\right) \Rightarrow$$

$$C_{sym.}(E_n) = \begin{cases} -ie^{\kappa} A^{I}_{\alpha\to\beta}\left(\dfrac{(1+\rho)}{(1+\rho)-2\rho+\rho} - \dfrac{0}{4\rho}\right) \cong -i(1+\rho) & \text{for n odd} \\ -ie^{\kappa} A^{I}_{\alpha\to\beta}\left(\dfrac{0}{4\rho} - \dfrac{(1+\rho)}{(1+\rho)-2\rho+\rho}\right) \cong i(1+\rho) & \text{for n even} \end{cases} \tag{23}$$

c) Next we can now investigate the case of maximum contribution which occurs by minimizing the denominators of $C_{sym.}$. We seek for a solution of the equation $(1+\rho)(1+\sigma)-2\rho\sigma\mp 2i\rho\sigma^{1/2}=0\Rightarrow(1-\rho)\sigma\mp 2i\rho\sigma^{1/2}+(1+\rho)=0$ and we easily find $(\sigma_{1\sim 3,2\sim 4})^{1/2}=\pm\, i(1+\varepsilon),\ \mp\ i$, where we have defined $\varepsilon=\dfrac{2\rho}{1-\rho}>0$. In this way we can actually write

$$C_{sym.}=-ie^{\kappa}A^{I}_{\alpha\to\beta}\left(\frac{(1+\rho)\left(\sigma^{1/2}-i\right)\left(\sigma^{1/2}+i\right)}{(1-\rho)\left(\sigma^{1/2}-\sigma_1^{1/2}\right)\left(\sigma^{1/2}-\sigma_2^{1/2}\right)}-\frac{(1+\rho)\left(\sigma^{1/2}-i\right)\left(\sigma^{1/2}+i\right)}{(1-\rho)\left(\sigma^{1/2}-\sigma_3^{1/2}\right)\left(\sigma^{1/2}-\sigma_4^{1/2}\right)}\right)\Rightarrow$$
$$C_{sym.}=-ie^{\kappa}A^{I}_{\alpha\to\beta}(\varepsilon+1)\left(\frac{\left(\sigma^{1/2}-i\right)}{\left(\sigma^{1/2}-i-i\varepsilon\right)}-\frac{\left(\sigma^{1/2}+i\right)}{\left(\sigma^{1/2}+i+i\varepsilon\right)}\right) \tag{24}$$

It is obvious that each pole term corresponds to odd and even eigenvalue contribution. It is therefore desirable to develop the above expression around the eigenvalues of the unperturbed problem and call the expansion $C^{odd}_{sym.}$ and $C^{even}_{sym.}$ respectively. We can then write

$$C^{odd}_{sym.}\cong -ie^{\kappa}A^{I}_{\alpha\to\beta}\frac{\varepsilon+1}{\varepsilon/2+1}+ie^{\kappa}A^{I}_{\alpha\to\beta}(\varepsilon+1)\left(\frac{-i\varepsilon}{i\varepsilon+\dfrac{d\sigma^{1/2}}{dE}(E-E_n)+i\dfrac{d\varepsilon}{dE}(E-E_n)}\right)$$
$$=-i+\frac{1}{1-\rho}\left(\frac{\left(\dfrac{1}{\varepsilon}\dfrac{dk}{dE}+i\dfrac{d\ln\varepsilon}{dE}\right)^{-1}}{E-\left(E_n-\dfrac{\dfrac{d\ln\varepsilon}{dE}}{\dfrac{1}{\varepsilon^2}\left(\dfrac{dk}{dE}\right)^2+\left(\dfrac{d\ln\varepsilon}{dE}\right)^2}-i\dfrac{\dfrac{1}{\varepsilon}\dfrac{dk}{dE}}{\dfrac{1}{\varepsilon^2}\left(\dfrac{dk}{dE}\right)^2+\left(\dfrac{d\ln\varepsilon}{dE}\right)^2}\right)}\right) \tag{25}$$

Repeating the same procedure for the even eigenfunctions we similarly find

ISSN:1791-4469

$$C_{sym.}^{even} \cong ie^{\kappa} A_{\alpha\to\beta}^{I} \frac{\varepsilon+1}{\varepsilon/2+1} - ie^{\kappa} A_{\alpha\to\beta}^{I} (\varepsilon+1) \left(\frac{i\varepsilon}{-i\varepsilon + \frac{d\sigma^{1/2}}{dE}(E-E_n) - i\frac{d\varepsilon}{dE}(E-E_n)} \right)$$

$$= i - \frac{1}{1-\rho} \left(\frac{\left(\frac{1}{\varepsilon}\frac{dk}{dE} + i\frac{d\ln\varepsilon}{dE} \right)^{-1}}{E - \left(E_n - \frac{\frac{d\ln\varepsilon}{dE}}{\frac{1}{\varepsilon^2}\left(\frac{dk}{dE}\right)^2 + \left(\frac{d\ln\varepsilon}{dE}\right)^2} - i\frac{\frac{1}{\varepsilon}\frac{dk}{dE}}{\frac{1}{\varepsilon^2}\left(\frac{dk}{dE}\right)^2 + \left(\frac{d\ln\varepsilon}{dE}\right)^2} \right)} \right) \tag{26}$$

It is obvious that the form of the semiclassical Greens function is the same for both odd and even eigenfunctions. In addition the above expressions show that the Green's function is constituted by two terms, a non pole term that gives the general background of propagation and the pole term, which is the second term of the equation. The pole term reveals the complex energy eigenvalues, since now the initially prepared state of continuum 1, (particle in allowed region O), decays in the continuum 2, (particle in allowed region IV). The imaginary part of the energy pole expresses the rate with which continuum 1 decays into continuum 2 and for this it is interesting to notice that it is clearly negative as it should. This is true since *k* is an increasing function of energy and therefore the derivative *dk/dE* is a positive quantity. In addition there is a real energy shift also coming from the above interaction. Since the major contribution of the pole term can be written as a sum , we can actually write the Green's function in the following form

$$G_{sc}(E) \equiv G_{sc}^{n.p.}(E) + G_{sc}^{p.}(E) = i\frac{\exp\left[i\int_{z_1}^{\alpha} k^*(z)dz + i\int_{\delta}^{z_2} k^*(z)dz\right]}{\overline{p(z_1,z_2)}} - \sum_{n=0}^{n=n_{\max}} \frac{A_n}{E-Z_n} \tag{27}$$

where

$$A_n = \frac{\exp\left[i\int_{z_1}^{\alpha} k^*{}_n(z)dz + i\int_{\delta}^{z_2} k^*{}_n(z)dz\right]}{\overline{p_n(z_1,z_2)}\left\{(\rho+1)\left(\frac{1}{\varepsilon}\frac{dk}{dE} + i\frac{d\ln\varepsilon}{dE}\right)\right\}_{E\,=\,E_n}} \tag{28}$$

and

$$Z_n = E_n - \frac{\frac{d\ln\varepsilon}{dE}}{\frac{1}{\varepsilon^2}\left(\frac{dk}{dE}\right)^2 + \left(\frac{d\ln\varepsilon}{dE}\right)^2} - i\frac{\frac{1}{\varepsilon}\frac{dk}{dE}}{\frac{1}{\varepsilon^2}\left(\frac{dk}{dE}\right)^2 + \left(\frac{d\ln\varepsilon}{dE}\right)^2} \tag{29}$$

and where $\overline{p_n(z_1,z_2)}$ and $k^*{}_n(z)$ are defined by $\overline{p(z_1,z_2)}$ and $k^*(z)$ respectively for $E=E_n$. We should note that the n_{max} corresponds to the maximum quantum number that the barriers can

support. However a more accurate result should be produced by developing the $C_{sym.}$ in second order around the unperturbed eigenvalues, meaning

$$C_{sym.} \cong \left(\frac{(\rho+1)^{-1}}{i+\left(i\frac{d\ln\varepsilon}{dE}+\frac{1}{\varepsilon}\frac{dk}{dE}\right)(E-E_n)+i\left(\frac{-i}{2\varepsilon}\frac{d^2k}{dE^2}+\frac{1}{2\varepsilon}\left(\frac{dk}{dE}\right)^2+\frac{1}{2}\frac{d^2\ln\varepsilon}{dE^2}+\frac{1}{2}\left(\frac{d\ln\varepsilon}{dE}\right)^2\right)(E-E_n)^2} - i \right) \tag{30}$$

Proceeding the algebra by first defining the following quantities,

$$\begin{aligned} &\delta=\frac{d\ln\varepsilon}{dE},\gamma=\frac{1}{\varepsilon}\frac{dk}{dE},\beta=\frac{\varepsilon}{2}\gamma^2+\frac{1}{2}\frac{d\delta}{dE}+\frac{1}{2}\delta^2,\alpha=\frac{1}{2\varepsilon}\frac{d(\gamma\varepsilon)}{dE} \\ &\rho=\sqrt{\left(\gamma^2-\delta^2+4\beta\right)^2+\left(2\gamma\delta-4\alpha\right)^2},\varphi=\tan^{-1}\frac{2\gamma\delta-4\alpha}{\gamma^2-\delta^2+4\beta} \end{aligned} \tag{31}$$

we find for the perturbed eigenvalues, the following result

$$Z_n^* \cong E_n+\frac{-\alpha\gamma-\beta\delta}{2\left(\alpha^2+\beta^2\right)}\pm\sqrt{\rho}\frac{\alpha\cos\varphi/2+\beta\sin\varphi/2}{2\left(\alpha^2+\beta^2\right)}-\frac{i}{2}\left(\frac{\alpha\delta-\beta\gamma}{\left(\alpha^2+\beta^2\right)}\pm\sqrt{\rho}\frac{\beta\cos\varphi/2-\alpha\sin\varphi/2}{2\left(\alpha^2+\beta^2\right)}\right) \tag{32}$$

Not symmetric barriers case:

In the case of not symmetric potential barriers neither the $A^{III}_{\gamma\to\gamma}=A^{I}_{\beta\to\beta}$ nor the $A^{I}_{\alpha\to\beta}=A^{III}_{\gamma\to\delta}$ relation holds. The total path contribution from α to δ can be simply written as

$$C_{not\ sym}=\frac{A^{I}_{\alpha\to\beta}A^{II}_{\beta\to\gamma}A^{III}_{\gamma\to\delta}}{\left(1-A^{III}_{\gamma\to\gamma}A^{II}_{\beta\to\beta}\right)\left(1-A^{I}_{\beta\to\beta}A^{II}_{\beta\to\beta}\right)-A^{III}_{\gamma\to\gamma}\left(A^{II}_{\beta\to\gamma}\right)^2A^{I}_{\beta\to\beta}} \tag{33}$$

The numerator of the above fraction expresses direct propagation from α to δ without interchanging the regions of motion while the denominator equals unity when the barriers strongly confine the inner quantum well. Looking the above relation in a more accurate way and since it is true that $A^{II}_{\beta\to\gamma}=ie^{-ik_o}A^{II}_{\beta\to\beta}$, the denominator becomes a second order polynomial of the amplitude $A^{II}_{\beta\to\beta}$. Since the amplitude $A^{II}_{\beta\to\beta}=A^{II}_{\gamma\to\gamma}$ is defined according to (18c) as $A^{II}_{\beta\to\beta}=-i\frac{\sigma}{\sigma+1}$ and $A^{II}_{\beta\to\gamma}$ according to (18b) as $A^{II}_{\beta\to\gamma}=\frac{\sqrt{\sigma}}{\sigma+1}$ we can actually write equation (30) in the following form

$$C_{not\ sym}=\frac{A^{I}_{\alpha\to\beta}A^{III}_{\gamma\to\delta}\sqrt{\sigma}}{-\sigma A^{III}_{\gamma\to\gamma}A^{I}_{\beta\to\beta}+i\sigma\left(A^{I}_{\beta\to\beta}+A^{III}_{\gamma\to\gamma}\right)+\sigma+1}\equiv\frac{A^{I}_{\alpha\to\beta}A^{III}_{\gamma\to\delta}\sqrt{\sigma}}{\sigma(1-\rho^{*})+1} \tag{34}$$

where we have defined the real barrier factor $\rho^{*}>0$ as

$$\rho^{*}=-4\frac{e^{-(\kappa_1+\kappa_2)}}{\left(4+e^{-2\kappa_1}\right)\left(4+e^{-2\kappa_2}\right)}+2\left(\frac{e^{-2\kappa_1}}{4+e^{-2\kappa_1}}+\frac{e^{-2\kappa_2}}{4+e^{-2\kappa_2}}\right) \tag{35}$$

As in the symmetric case direct remarks can be expressed concerning the structure of the above formula:

a) It is obvious that when $A^{I}_{\alpha\to\beta},A^{III}_{\gamma\to\delta},\rho^{*}\ll 1$ which is the case of strong barrier confinement of the inner classically allowed region, practically no interaction with the continuum appears, we get

$$C_{not\ sym}\cong\frac{0\cdot\sqrt{\sigma}}{\sigma+1}=0 \tag{36}$$

b) Again as was shown in remark b) of the symmetric barriers case the condition $1+\sigma=0$ determines the eigenvalues of the classically allowed region of motion.

c) We can now investigate the case of maximum contribution which occurs when we minimize the denominator of the $C_{not\ sym}$. This is equivalent to finding the solution of the equation $\sigma(1-\rho^{*})+1=0$ which is $\sigma_o=-\frac{1}{1-\rho^{*}}$. Since ρ^{*} is a small real quanti-ty σ becomes also a real quantity close but not equal to -1 which contradicts to its form of $\sigma\equiv e^{2ik}$. For this we develop the denominator around $\sigma=-1$, or equivalently around the eigenvalues of the inner quantum well, and substitute to the numerator the value $\sigma_o=-\frac{1}{1-\rho^{*}}$.

$$\begin{aligned}C_{not\ sym}&\cong\frac{A^{I}_{\alpha\to\beta}A^{III}_{\gamma\to\delta}\sqrt{-\frac{1}{1-\rho^{*}}}}{\rho^{*}-\frac{d\sigma}{dE}(\rho^{*}-1)\left(E-E_n\right)-\sigma\frac{d\rho^{*}}{dE}\left(E-E_n\right)}\\&=\frac{A^{I}_{\alpha\to\beta}A^{III}_{\gamma\to\delta}\sqrt{-\frac{1}{1-\rho^{*}}}\left[\left(\frac{d\rho^{*}}{dE}+2i(\rho^{*}-1)\frac{dk}{dE}\right)_{E=E_n}\right]^{-1}}{\left\{E-E_n+\frac{\rho^{*}\frac{d\rho^{*}}{dE}}{\left(\frac{d\rho^{*}}{dE}\right)^2+\left(2(\rho^{*}-1)\frac{dk}{dE}\right)^2}\Bigg|_{E=E_n}-i\frac{2\rho^{*}(\rho^{*}-1)\frac{dk}{dE}}{\left(\frac{d\rho^{*}}{dE}\right)^2+\left(2(\rho^{*}-1)\frac{dk}{dE}\right)^2}\Bigg|_{E=E_n}\right\}}\end{aligned} \tag{37}$$

In this way the perturbed eigenvalues, for the not symmetric case, take the following form

$$Z_n \cong E_n - \frac{\rho^* \frac{d\rho^*}{dE}}{\left(\frac{d\rho^*}{dE}\right)^2 + \left(2(\rho^* - 1)\frac{dk}{dE}\right)^2}\Bigg|_{E=E_n} - i\frac{2\rho^*(1-\rho^*)\frac{dk}{dE}}{\left(\frac{d\rho^*}{dE}\right)^2 + \left(2(\rho^* - 1)\frac{dk}{dE}\right)^2}\Bigg|_{E=E_n} \tag{38}$$

Again the imaginary part is clearly negative since quantity ρ^* is normally much less than unity and the derivative *dk/dE* is a positive quantity. Again we can write the Green's function as a sum of poles meaning

$$G_{sc}(E) \equiv \sum_{n=0}^{n=n_{\max}} \frac{B_n}{E - Z_n} \tag{39}$$

where

$$B_n = \left\{ A^I_{\alpha\to\beta} A^{III}_{\gamma\to\delta} \sqrt{-\frac{1}{1-\rho^*}} \left[\left(\frac{d\rho^*}{dE} + 2i(\rho^* - 1)\frac{dk}{dE} \right) \right]^{-1} \right\}_{E=E_n} \tag{40}$$

In each case of symmetry, equation (29,32) or (38), the complex energy poles take the general form, $Z_n = E_n + \Delta_n - i\frac{\Gamma_n}{2}$ and the width of the decaying state is analytically given as follows:

$$\Gamma_{n,sym.} = \frac{\frac{2}{\varepsilon}\frac{dk}{dE}}{\frac{1}{\varepsilon^2}\left(\frac{dk}{dE}\right)^2 + \left(\frac{d\ln\varepsilon}{dE}\right)^2}\Bigg|_{E=E_n} \tag{41a}$$

and

$$\Gamma_{n,not\ sym.} = \frac{4\rho^*(1-\rho^*)\frac{dk}{dE}}{\left(\frac{d\rho^*}{dE}\right)^2 + \left(2(\rho^* - 1)\frac{dk}{dE}\right)^2}\Bigg|_{E=E_n} \tag{41b}$$

The fundamental time associated with the tunneling process if the exponential law is assumed for its evolution , is often taken as the lifetime of the quasi – bound state and is related to the width via the $\tau = \frac{\hbar}{\Gamma}$. In this way we can write for the transport time, for the symmetric and the not symmetric case respectively, the following relations

$$\tau_{n,sym.} = \hbar \frac{\frac{1}{\varepsilon^2}\left(\frac{dk}{dE}\right)^2 + \left(\frac{d\ln\varepsilon}{dE}\right)^2}{\frac{2}{\varepsilon}\frac{dk}{dE}}\Bigg|_{E=E_n} \tag{42a}$$

and

$$\tau_{n,not\ sym.} = \hbar \frac{\left(\frac{d\rho^*}{dE}\right)^2 + \left(2(\rho^* - 1)\frac{dk}{dE}\right)^2}{4\rho^*(1-\rho^*)\frac{dk}{dE}}\Bigg|_{E = E_n} \tag{42b}$$

For once again we should emphasize that the above expressions describe the transport time as long as the other time constants are significantly minimized.

5. CONCLUDING REMARKS

In this work we studied quantum transmission in a double barrier structure via the semi-classical path integral method. This kind of potential covers a large area of interest and applications from different branches of physical sciences, from quantum chemistry up to nanoelectronics. As such we demonstrated the model of a quantum tunneling diode. We produced analytic relations for the transmission amplitude, which is the Green's function for single charge transport between the two metals of the diode, due to resonant tunneling. The Green's function, as was expected, appears to have a pole structure which reveals the complex energy spectrum of the structure. The later is described via analytic relations. The imaginary part of each complex pole is related to the time needed for transport, under certain circumstances where the barrier penetration strongly dominates all the other mechanisms generating an intrinsic time constant. The above study was done for both the symmetric and the not symmetric barrier case, concerning the absence or the application respectively of an external electric field, on the diode's band profile.

REFERENCES

1. R.H. Fowler and L. Nordheim, Proc. Roy. Soc. (London) 119,173-181 (1928).
2. C. Zener, Proc. Roy. Soc. (London) A145, 523-529 (1934).
3. L. Esaki Phys. Rev. 109,603-604(L) (1958).
4. R. Tsu and L. Esaki, Appl. Phys. Lett. 22, 562-564 (1973).
5. L.L. Chang, L. Esaki and R. Tsu, Appl. Phys. Lett. 24, 593-595 (1974).
6. J. Peng, H. Chen and S. Zhou, J. Phys. : Condens. Matter 1, 5451-5462, (1989).
7. H. Cruz, A. Hernandez-Cabrera and P. Aceituno, J. Phys. : Condens. Matter 2,8953-8959, (1991).
8. Y. Zebda and A.M. Kan'an, J. Appl. Phys. 72, 559-563, (1992).
9. B. Mendez and F. Dominguez-Adame, Am. J. Phys. 62, 143-147, (1994).
10. S. Vatannia and G. Gildenblat, IEEE J. Quant. Electr. , 32, 1093-1105, (1996).
11. I. Yanetka, Acta Physica Polonica A, 116,1059-1066, (2009).
12. W.H. Miller, Adv. Chem. Phys. 25,69-177, (1974).
13. M.C. Gutzwiller, J. Math. Phys. 12, 343-358, (1971).
14. B.R. Holstein and A.R. Swift, Am. J. Phys. 50, 833-839, (1982).
15. B.R. Holstein, Am. J. Phys. 51, 897-901, (1983).
16. B.R. Holstein, Topics in Advanced Quantum Mechanics, Redwood California: Addison-Wesley,1992, pp. 209-224.
17. H. Kleinert, Path Integrals in Quantum Mechanics, Statistics and Polymer Physics, Singapore: World Scientific, 1995, 2nd edition, pp. 215-222.
18. M. Razavy, Quantum Theory of Tunneling, Singapore: World Scientific, 2003, pp. 219-240.

Water treatment with emphasis to magnetic separation methods

Spyros Golfinopoulos[a] and Konstantinos Papageorgiou[a]

[a]*University of the Aegean, Department of Financial and Management Engineering, 41 Kountouriotou Str., 82100, Chios, Greece*

Abstract. One of the most important natural resources is the water. Surface and ground waters are the main sources and there are several differences between them. We need not only water in adequate quantities to cover the human necessities but also we have to develop systems that deliver a good quality of water. The majority of human uses require fresh water. However, in water sources there are several undesirable contaminants. A conventional water treatment includes the processes, flocculation, coagulation, sedimentation, filtration and disinfection. In some cases, a serious issue that we have to deal with is the hardness of water. In this paper we study the effect of magnetic fields on the removal of calcium carbonate which is responsible for the water hardness.

INTRODUCTION

The water is a vital element for life. Man can survive without food for more than a month but it is impossible to live without water for more than one week. He needs water to survive. Every cell of the human body requires water as nourishment and to remove wastes. The lack of water reduces the amount of blood, causes many problems in health.

The role of water in human culture and civilization is well documented all over the world. It is very significant all people to have access to this resource, as water is used for drinking, preparing food, cleaning, bathing, irrigating landscapes etc.

Water is considered as a renewable resource. In our days the world's supply of groundwater is steadily decreasing due to the expanding human population and its competition for the water. Even the climate change is a significant factor that affects the available quantities of water. The strong connection between the climate and hydrologic cycle is studied intensively.

The agriculture, industrial, household, recreational and environmental activities are the main uses of water. All these uses require fresh water. All over the world, a large percent of about 70 % of water use is estimated that is for irrigation in agriculture.

The major drinking water sources include ground water, surface water, atmospheric water generation, rain water harvesting, fog collection and sea water.

Surface water and groundwater are both important sources. Approximately 98 percent of liquid fresh water exists as groundwater, much of it occurs very deep in the Earth. This makes pumping very expensive, preventing the full development and use of all groundwater resources.

The accessibility to safe drinking water are of major concern throughout the world. The drinking water must be free of chemical substances and organisms, when it is used for human

ISSN:1791-4469

consumption. The increase use of chemicals in most countries has deteriorated the water quality.

Each source of water has a unique set of contaminants. The quality of ground and surface water is not the same In most cases the ground water in contrast with surface water is microbially safe and chemically stable in the absence of direct contamination. For example groundwater contains pesticide chemicals and nitrate while surface water stores most bacteria and other microorganisms. These contaminants may be shared between the two sources due to their interconnectedness. Their contamination could be varied depending on the source of pollutant. There is no water source free from undesirable contaminants.

First of all, the water supplier has the responsibility to set the first barriers in protection of the water resources in order to provide drinking water of high quality. With this action that leads in decreasing of the contamination the degree of required treatment is reduced and the operation costs of a water treatment plant are lower. The factors that influence the quality of raw water are both natural and human. The first category includes wildlife, topography, geology, vegetation and the second refers mainly to wastewater discharges, urban and agricultural runoff, as agrochemicals and livestock or recreational use [1].

The next step is the water treatment process including coagulation, flocculation, sedimentation, filtration and disinfection. These processes are used for the removal of particles and microorganisms. The microorganisms may be pathogenic and capable to cause disease or mortality to human population. Chlorination is the most common disinfection process while chloramination, chlorine dioxide application ozonation and ultraviolet irradiation are also used [1].

DRINKING WATER

Drinking water is the water that is safe enough to be consumed by humans or used with low risk of immediate or long term harm. Another definition of water is potable. This word comes from the latin potabilis, meaning drinkable.

The occurrence of undesirable contaminants in raw water affect the human health and influence the water quality. The removal of them is necessary and obligatory in most of the cases and especially crucial when the water source is surface.

The quality parameters set for drinking water described by drinking water quality standards. In most developed countries, the water supplied to households, commerce and industry meets these standards.

All developed countries specify quality standards of drinking water to ensure the consumption of safe water In European Union drew up the Council Directive 98/83/EC on the quality of water intended for human consumption, adopted by the Council on 3 November 1998. This Directive provides a sound basis for both the consumers throughout the EU and the suppliers of drinking water [2].

In USA, the Environmental Protection Agency (EPA) sets standards that, when combined with protecting ground water and surface water, are critical to ensuring safe drinking water. EPA works with its regional offices, states, tribes and its many partners to protect public health through implementing the Safe Drinking Water Act [3].

World Health Organization (WHO) produces international norms on water quality and human health in the form of guidelines that are used as the basis for regulation and standard setting, in developing and developed countries world-wide [4]. These guidelines are the international reference point for standards setting and drinking-water safety.

WATER POLLUTION

There is evidence of widespread contamination of water resources. It isn't known exactly the health effects from the long term exposure to the contaminants that enter in the water supply system. It is noticed that all contaminants do not pose a threat to human health. When the pollutants enter water sources their concentrations dilute and are reduced by biological degradation, filtration, and adsorption to soil. Some chemicals are very stable in the environment, such as the chlorinated hydrocarbons. Some of these can accumulate in living organisms and are not readily metabolized and excreted. The most effective way of reducing contaminants in drinking water is by controlling it at the source.

Iron, manganese, barium, fluoride, hydrogen sulfide, and salt may be present in water resources at undesirable levels. Bacteria from sewage, and animal wastes are a common problem. Another issue is the high level contamination form nitrate-nitrogen levels in agriculture areas.

In order to have unpolluted water man treated it from the ancient years. The first methods of water treatment were recorded as early as 4000 B.C. In ancient Greek and India (Sanskrit) the filtering, the exposing to sunlight, the boiling and the straining were applied to improve the taste and odor of drinking water. The Egyptians after 1500 BC, using the chemical alum, discovered the coagulation method that caused the settling out of suspended particles. Another significant process for removing particles from water was filtration. This method was established during the 1700s. In Europe the filtration with sand was used by the early 1800s.This method was established as a successful way of removing particles from water and widely adopted in Europe during the nineteenth century. During the middle of nineteenth century was discovered a cholera epidemic spread through water and then the chlorination method was applied to purify the water. In this period worries about the quality of drinking water focused on pathogens in public water supply systems. The chlorination played a significant role to reduce the waterborne disease outbreaks in the early 1900s, as it was clear that filtration did not remove all the bacteria from public-use water. Today, the most effective treatment techniques for protecting water supplies from harmful microbes are filtration and chlorination [5,6].

However during chlorination, chlorine reacts with natural organic materials and forms a large category of organic compounds, named haloforms that may pose carcinogenic or mutagenic properties. A characteristic example is trihalomethane (THM) formation [7].

WATER TREATMENT PROCESSES

Drinking water treatment technologies and purification techniques have been used and continuously developed over the ages, to protect public health from pathogens and chemicals [6].

Coagulation, flocculation, sedimentation, filtration and disinfection are the most commonly applied water treatment processes. They used worldwide in the water treatment industry, from the end of nineteenth century, before the distribution of drinking water to consumers [8].

Each process is used for different purpose [1,9,10,]:

Coagulation promotes the interaction of small particles to form larger particles. The process refers to coagulant addition (i.e. addition of a substance that will form the hydrolysis products that cause coagulation), particle destabilization and interparticle collisions. Chemical coagulation has indirect impacts on the efficiency of disinfection process.

- Flocculation is the physical process of producing interparticle contacts that lead to the formation of large particles.

ISSN:1791-4469

The goal of coagulation and flocculation is to remove fine suspended particles. These particles can attract and hold bacteria and viruses to their surface. The percentage of removal for bacteria and viruses is 99.9 and 99, respectively. They can also remove some of the organic matter that gathers as water travels across the land, from raindrop to surface water. However, they cannot remove full taste and odor.

- Sedimentation is a solid–liquid separation process, in which particles settle under the force of gravity.
- Filtration can act as a consistent and effective barrier for microbial pathogens and in some cases may be the only treatment barrier. It occurs as the water passes through a substance that helps remove even smaller particles. Filters can be made of layers of sand, gravel and charcoal.

However in many cases this treatment is not sufficient. For example, adsorption and oxidation may be required if undesirable impurities contained in the water. These are two additional treatment processes, where adsorption is a type of chemical filtration while oxidation can offer water disinfection and can destroy taste, odor, algal toxins, pesticides and other soluble contaminants. Oxidation can also remove iron and manganese [9].

These processes are very common all around the world and their purpose is to remove turbidity and contaminants from water and to improve and protect water quality. The characteristics of the water, the types of water quality problems, the cost of different treatments processes and some other factors define the alternative treatments that would be chosen. A great challenge involving technological development is the need to develop technology that is suitable, applicable, and sustainable [6].

THE ROLE OF MAGNETIC MATERIALS TO WATER PURIFICATION

Water purification using physical methods such as magnetic separation, have drawn the attention of the scientific community in the past few decades. The formation of scale deposits by natural waters can significally limit or completely block the water flow in pipes or boilers and heat exchangers, resulting in heat transfer efficiency reduction and leading progressively in equipment damage. Huge amounts of energy are wasted therefore causing severe technical and economic problems in the industry.

Experimental results of a magnetic field on the precipitation process of calcium carbonate scale from a hard water have been reported [11]. Because a magnetic field is able to disturb the double ionic layer surrounding the colloidal particles and their zero potential, would also tend to reduce the nucleation rate and to accelerate the crystal growth. Experimentally the nucleation time was identified from the variations of the pH and the Ca^{++} concentration. It was shown that the magnetic treatment increases the total amount of precipitate and the effect depends on the solution pH, the flow rate and the duration of the treatment.

In order to better understand and explain how a magnetic field influences nucleation, another study [12] using X-ray diffraction and electron microscopy techniques for the characterization of the carbonates formed by heating water has shown that drawing water through a static magnetic field of approximately 0.1 T increases the aragonite/calcite ratio in the deposit.

A valuable review paper [13] presents and explains several aspects on water purification techniques using magnetic assistance. According to this paper based on the difference in adoption of physical process, magnetically assisted water purification can be classified in several categories, as follows. In the direct purification method, there is no carrier magnetic component added. The magnetic field helps in inducing crystallization and then a magnetic filter is used to remove salt ions and to prevent them to enter the pipelines. Seeding method following

by a high gradient magnetic separation device is also one of the older methods used for water purification. A coagulant cation under proper chemical conditions yields an insoluble precipitate with positive susceptibility under applied magnetic field. This property is used to trap the coagulant with contaminant on a magnetic filtration assembly. Magnetic ion exchange resins method were introduced for the removal of organic matter from ambient raw water, because the ion exchange resin beads contain a magnetized component within their structure which allows the beads to act as individual magnets and form agglomerates. Finally in the combined application method, as the name indicates, magnetic separation techniques are additionally employed in order to further improve the separation efficiency after the initial treatment of the water with electrolytic or catalytic reactions.

Magnetism helps in water purification by influencing the physical properties of the water contaminants. The method exploits the difference in behavior of particles in magnetic fields. This property is characterized by a dimensionless variable called magnetic susceptibility which is denoted as χ. The specific magnetic susceptibility is then defined as $\chi_m = \chi / \rho$, where ρ is the density of the material. Materials which repelled from the magnetic field have negative values of magnetic susceptibility and are called diamagnetics, while particles attracted towards greater intensity of the magnetics field are called paramagnetics.

Magnetics separation is based on the principal that the force acting on a particle is given by the equation $F = \mu_0 \chi_m m H \nabla H$

Therefore the component of the force acting on a particle in magnetic field in the x-axis is $F_x = \mu_0 \chi_m m H \frac{\partial H}{\partial x}$, where $\mu_0 = 4\pi 10^{-7} Wb / A.m$ denotes the magnetic permeability of vacuum, H is the magnetic field intensity, m the mass of the particle and $\frac{\partial H}{\partial x}$ is the field gradient.

From this equation it is obvious that the collection of particles is strongly depends on the creation of these large magnetic field gradients, as well as on the particle size and its magnetic properties. This can be achieved by placing magnetically susceptible wires inside an electromagnet. These wires dehomogenize the magnetic field producing large gradients around the wires that attract magnetic particles to their surfaces and tram them there. Consequently with the adoption of this technique, the formation of calcium carbonate particles expected in the bulk of the scaling water preventing them to precipitate on the walls of the distribution pipes.

Another crucial aspect is the strength of the magnetic field that we are going to employ in a magnetic separator. In cases where the pollutant are solids, separators based on permanent magnets are the most appropriate choice, while in cases where the amount of pollutants is low, high intensity magnetic field is required and the superconducting magnets should be used. A superconducting magnet is an electromagnet with coils made out of superconducting wires. During its operation is has to cooled down to cryogenic temperatures. The major advantage of using superconducting magnets, is basically their reduced operation costs, which is due to the fact that basically there are no losses of power to ohmic resistances. Finally an important application of superconducting magnets is the removal of the radio toxic hazard of plutonium which is form compounds with paramagnetic properties.

CONCLUSIONS

The clean drinking water is a crucial matter. Population growth, changes in habits and our daily needs have created a tremendous demand for potable water that meets the quality

ISSN:1791-4469

standards. The conventional processes including coagulation, flocculation, sedimentation, filtration and disinfection reduce the concentration of particulate matter and provide water that does not contain pathogenic contaminants.

One of the major problems is the water hardness. Waters with high hardness treated to remove the calcium and magnesium salts. The principle of water purification using magnetic separation methods has been analyzed and applied in the case of calcium carbonate particles removal. The efficiency of the process can be evaluated by measuring the ionic calcium concentrations before and immediately after the water treatment. Despite the fact that this is a very promising and low cost operation method, several other parameters like the geometry of the device separator, the velocity of the water flow, the temperature of the water which can affect the solubility and the precipitation of the calcium carbonate and the pH of the water, have to be studied more thoroughly in order to maximize the efficiency of the water treatment process.

REFERENCES

1. WHO (2004). Guidelines for Drinking-water Quality. 3rd Edition, Volume 1, Recommendations, Geneva
2. European Commission (1998) URL http://eur-lex.europa.eu/browse/summaries. html
3. EPA (2013) URL http://water.epa.gov/drink/standardsriskmanagement.cfm
4. WHO (2014) URLhttp://www.who.int/water_sanitation_health/dwq/guidelines/en/
5. EPA (2000). The History of Drinking Water Treatment
6. C. Ray and R. Jain (2011). Chapter 2. Drinking Water Treatment Technology -Comparative Analysis, C. Ray, R. Jain (eds.), Drinking Water Treatment, Strategies for Sustainability, Springer Science + Business Media B.V
7. S.K. Golfinopoulos (2000). The occurrence of trihalomethanes in the drinking water in Greece. Chemosphere, 41, 1761-1767
8. A. Ndabigengesere, K. S. Narasiah (1998). Quality of water treated by coagulation using Moringa oleifera seeds. Water Research, 23, 781-791
9. Cooperative Research Centre (CRC) for Water Quality and Treatment (2008). Drinking Water Treatment, Issue 2
10. M.W. LeChevallier, K-K Au (2004). Water Treatment and Pathogen Control Process, Process Efficiency in Achieving Safe Drinking Water, World Health Organization (WHO)
11. Alimi Fathi et al., Effect of magnetic water treatment on homogeneous and heterogeneous precipitation of calcium carbonate, Water Research 40 (2006) 1941-1950
12. J.M.D. Coey and Stephen Cass, Magnetic water treatment, Journal of Magnetism and Magnetic Materials 209 (2000) 71-74
13. R. D. Ambashta, M. Sillanpaa (2010). Water purification using magnetic assistance: A review. Journal of Hazardous Materials, 180, 38–49

Fuel-Efficient Ship Routing

Kyriakos Avgouleas[a] and Paul D. Sclavounos[b]

[a]*Lieutenant Commander HN, Submarine Shipbuilding Program, Hellenic Navy Detachment of Skaramangkas, Greece*
[b]*Professor, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge MA, USA*

Abstract. Optimal weather routing of ships seeks to determine the voyage route, in rough seas, which minimizes a certain metric or criterion (e.g. transit time, consumed bunker fuel, pollutant emissions etc.). In this paper the problem of minimum-fuel navigation in a sea state is studied. Advanced methods of hydrodynamic simulation are employed to compute the added resistance in waves and the responses of an example-ship in random incident waves. A fictitious trans-Atlantic route is introduced. Given the wave forecast of the geographical region of interest, the problem of determining the route that minimizes fuel consumption is numerically solved using an improved Dynamic Programming algorithm. It is concluded that significant savings in bunker fuel may result from the implementation of a decision support system based on state-of-the-art hydrodynamics, wave forecasting and optimization methods.

Keywords: Routing; optimization; hydrodynamics; resistance; fuel consumption.
PACS: 02.30.Yy, 47.35.Lf

INTRODUCTION

The shipping industry consumes about 5% of the world oil production in bunker fuel for maritime transportation. That corresponds to approximately 4 million barrels of oil a day. This fact warrants an effort to optimize navigation in a way that will bring the fuel cost down to the extent possible. A new approach to the solution of the optimal routing problem is presented in this paper. A combination of frequency domain methods for the computation of the ship hydrodynamics, state-of-the-art weather forecasts and advanced dynamic programming algorithms is utilized to generate trajectories (ship routes in the ocean) which minimize the required fuel while meeting certain safety restrictions associated with the severity of the sea.

The optimization problem associated with the selection of the fuel-minimizing route of a ship in a sea state is challenging. In calm weather the optimal route within the constraints imposed by sea lanes is usually the loxodrome (or the orthodrome for long routes), and the fuel consumption is known with a high degree of certainty given the sailing distance and the calm water resistance and propulsion characteristics of the vessel. In rough weather a number of complexities arise. The severity of the sea states to be encountered during the trip is not known a priori with certainty and must be estimated from weather forecasts supplied by a routing service. When sailing in rough seas, safety criteria must be met that protect the integrity of the hull and the cargo. They affect the vessel speed and heading in a sea state. Reliable

ISSN:1791-4469

seakeeping methods must be available in order to predict the vessel responses and the increase of her resistance in waves – the added resistance – given the attributes of the wave system. Finally, a robust and efficient dynamic optimization algorithm must be available that may be executed in real time to determine the optimal speed and heading of the vessel in rough weather in order to minimize the fuel consumption, subject to the safety constraints. These challenges are addressed in the present paper and an optimal ship routing algorithm is developed that leads to a notable reduction of the vessel fuel consumption in rough weather.

Literature Review

The problem of ship weather routing has been a subject of extensive research for many decades now. A comprehensive, albeit not exhaustive, review of the research work in the field starting from the 1950's can be found in Avgouleas (2008). Some of the developments in the last 10 years will be briefly presented here. Rathje and Beiersdorf (2005) developed a shipboard routing assistance (SRA) software to prevent containerships from encountering dangerous conditions with respect to seakeeping behavior (parametric rolling, slamming, exceedance of bending moment and shear force threshold values). Montes (2005) proposed a method for the automation of the Optimum Track Ship Routing (OTSR) system used by the US Navy. In Abramowski et al. (2006) a formal solution of the minimal time ship routing problem is presented on the basis of Pontryagin's maximum principle. Tsujimoto and Tanizawa (2006) solve the constrained optimization problem of minimum fuel routing, using the augmented Lagrange multiplier method. Böttner (2007) describes a decision support system able to provide optimal alternatives in case the ship finds herself in degraded condition (hull damage, rudder/propulsion failure etc.). The work of Szlapczyńska (2007) and extensions thereof (Szlapczyńska and Smierzchalski 2007-2009, Krata and Szlapczyńska 2011, Szlapczyńska 2013) focuses on deterministic weather routing of a sail assisted ship using Multi-Objective Evolutionary Algorithm (MOEA). Panigrahi et al. (2008) carry out a simulation of wave climate in the Indian Ocean and use it to minimize voyage time for a cargo ship route in the region. The optimization utilizes the Dijkstra algorithm and speed loss in waves is calculated from empirical curves. In Panigrahi et al. (2012) the hydrodynamics of ship motions and added resistance in waves are treated differently. Empirical speed loss curves are abandoned and standard frequency domain seakeeping computation is adopted. Sen and Padhy (2009) conducted a similar study using a Dijkstra algorithm as well, but with linear strip theory for hydrodynamic simulation. Wisniewski et al. (2009) utilize evolutionary algorithm to determine the minimum time route which safely avoids a tropical cyclone. Mezaoui et al. (2009) solve the unconstrained constant-speed minimum fuel problem in coastal navigation using high resolution forecasts and a Dijkstra optimization scheme. Dolinskaya et al. (2009) deduce shortest-time optimal paths inside convex regions of areas that a ship can sail. Their analysis is restricted to short-range optimal routing. Marie and Courteille (2009) deal with the deterministic dual-objective optimization of minimum time/minimum fuel route using Multi-Objective Genetic Algorithm (MOGA) and identifying Pareto optimal solutions. Hinnenthal and Clauss (2010) solve the same problem using similar principles but they quantify the robustness of the final solution exploiting ensemble forecasts. Delitala et al. (2010) explore the results of climatological simulations in combination with weather routing. Bruns et al. (2011) present a software tool for fuel efficient speed profile selection using potential flow CFD for calm water resistance prediction, strip theory for added resistance in waves and RANS modeling for the propulsor (both in design and off-design conditions). Gershanik (2011) addresses the challenges and benefits of weather routing optimization. The author advocates the use of the classical discrete dynamic programming algorithm (backward recurrence) for the minimum time or the minimum fuel optimal routing problem. Pipchenko (2011) solves the deterministic minimum work ship routing using genetic algorithm or a Nelder-

Mead heuristic method. Maki et al. (2011) propose a route optimization method that considers not only fuel efficiency but also ship stability (the risk of parametric rolling in particular). A real-coded genetic algorithm (GA) is used for the multi-objective optimization problem. Kobayashi et al. (2011) optimize the route of a containership in a simulated trans-oceanic passage, accounting for ship maneuvering dynamics. Dolinskaya (2012) introduces the notion of restricted turning radius in optimal path finding of a vessel in an inhomogeneous, non-stationary environment. Kosmas and Vlachos (2012) apply the simulated annealing method for the minimization of a cost function defined by the weighted sum of voyage time and a comfort parameter related to safety. Ilus and Heikkinen (2012) present a speed profile optimization approach which relies on historical and operational data collected for a specific route or leg of a voyage. The aim is to optimize energy usage while meeting prescribed constraints. Shao et al. (2012) solve the deterministic minimum fuel routing problem using a discrete 3D dynamic programming algorithm. Skoglund et al. (2012) present a method to obtain optimal routes in a directed graph using the concept of Pareto optimality. Both deterministic and ensemble forecasts are used. Tsou and Cheng (2013) combine an ant colony algorithm with a genetic algorithm for the solution of the minimal fuel and pollutant emission weather routing problem. In their recent work, Marie and Courteille (2014) construct a fuzzy logic model with data collected from ship's sensors to minimize fuel consumption of a sail-assisted motor vessel.

Model Overview

An example-ship is used from the Series 60 hullform, with the characteristics shown in table 1. A standard MARIN B-Series propeller is selected for this ship and a MAK 9M25C medium-speed, four-stroke diesel engine is matched with the propeller via a mechanical gearbox.

Resistance in calm water is calculated from standard statistical methods (Holtrop 1984). Added resistance in waves is computed using SWAN1, a frequency domain CFD code which uses 3D Rankine Panel Methods to simulate steady and unsteady potential flows around vessels with realistic hull shapes. Unlike most conventional strip and slender-body theories, SWAN1 exploits robust computational hydrodynamics methods to produce reliable results where approximate theories fail. The case of quartering and following waves can be mentioned as an example. At certain speeds they correspond to the sub-critical reduced frequency regime ($\tau<¼$) and represent challenging hydrodynamic problems intractable by strip theory. The same is true for the computation of added resistance in waves. Added resistance is a second order effect of central importance to weather routing, as it influences both the calculation of fuel consumption and the prediction of engine overload in a sea state. SWAN1 computes added resistance using direct integration and its results have been calibrated against numerous experiments carried out for America's Cup yachts (Sclavounos and Nakos 1993). In addition to resistance, all ship responses in waves are computed by SWAN1. In particular, pitch and heave motions and velocities near the ends of the hull are required to calculate probabilities of slamming and deck wetness. Limiting values imposed on these probabilities define safety constraints in the formulation of the minimum fuel consumption problem. Spectral analysis is used to estimate the mean added resistance and RMS values of the ship responses in a sea state with a given direction, significant wave height and modal period. The vessel hydrodynamic properties are computed a priori in the frequency domain across a wide range of encounter frequencies, ship speeds and wave headings and stored in a database for use in the solution of the fuel minimizing ship routing problem. Therefore, the time-consuming solution of the ship hydrodynamic problem in the time domain is avoided.

ISSN:1791-4469

TABLE 1. Principal Dimensions of Example-Ship.

Description	Symbol	Unit	Value
Length at DWL	L	m	101.7
Beam at DWL	B	m	14.28
Draft at DWL	T	m	5.7
Freeboard at DWL (bow)	H	m	2.85
Displacement	Δ	$tonne$	5700
Wetted Surface Area	S	m^2	1950
Block Coefficient	C_B	-	0.676
Prismatic Coefficient	C_P	-	0.688
Midship Section Coefficient	C_M	-	0.982

TABLE 2. Propeller Characteristics of Example-Ship.

Description	Symbol	Value
Diameter (m)	D	3.2
Pitch Ratio	P/D	1.2
Expanded Area Ratio	EAR	0.75
Number of Blades	z	5

TABLE 3. Engine Characteristics of Example-Ship.

Description	Value
Type	MAK 9M25C
Rated Power (MCR) (kW)	3000
Rated Speed (rev/min)	750
Gear Ratio	4.45:1

A third generation model, WAM cycle 4, developed at the European Center for Medium-Range Weather Forecasts (ECMWF) and adapted by the University of Athens (UoA), Greece, provides the weather forecast for the region of interest. The forecast is uploaded daily and covers a total of 168 hours. A fictitious cross-Atlantic route starting from Cape St.Vincent, Portugal, and ending at Norfolk Virginia, USA, defines the nominal voyage track considered in the present study. It is assumed that the nominal (calm water) speed for this trip is 13 knots. The nominal sailing distance is 3138.6 nautical miles following the shortest route, which is the great circle route (orthodrome). At that speed the trip would last 10.06 days. Mercator projection is used throughout. For the part of the route not covered by the wave forecast, calm water is assumed.

Keeping the start and end points fixed (“hard” constraints) and the “expiration time” of the chartering contract also fixed (i.e. the time at destination), a special Dynamic Programming (DP) algorithm is implemented in MATLAB® which calculates the optimal route with backward

recurrence. To safeguard against the shortcomings of the standard DP algorithm, namely the large requirements for memory storage and processing capacity, Iterative Dynamic Programming (IDP) is employed. IDP is an advanced extension to the standard DP algorithm (Luus 2000). It retains all the benefits of DP without suffering the infamous drawback known as the “curse of dimensionality”.

HYDRODYNAMICS

Calm Water Resistance

Calm water resistance mainly consists of three components: friction, form and wave drag. The friction drag coefficient is calculated in this paper by the ITTC 1957 model-ship correlation line

$$C_f = \frac{0.075}{(\log_{10} \mathrm{Re} - 2)^2} \tag{1}$$

The form drag coefficient is calculated from the Holtrop regression model (Holtrop 1984). It is given in the form of an enhancement factor *k* to the friction drag coefficient, based on the hull geometry and accounting for the presence of bilge keels (if present) and appendages. In this example, a bare Series 60 hull model is entered in the resistance calculations and the Holtrop resistance curve is shown in figure 1.

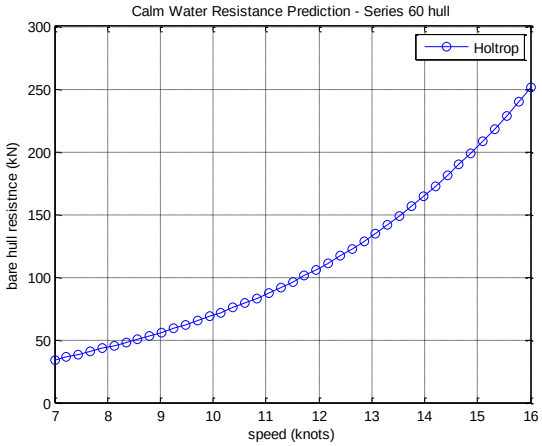


FIGURE 1. Steady Resistance from Holtrop Method.

The total calm water resistance coefficient is expressed as:

$$C_{total} = (1+k)\cdot C_f + C_w + C_a \tag{2}$$

where C_a is the model-ship correlation allowance and C_w the wave-making resistance coefficient, both obtained from the Holtrop statistical method. For a ship with total wetted surface *S* moving with constant forward speed *U* the total resistance is given by:

$$R_{calm} = C_{total} \cdot \frac{1}{2}\rho \cdot S \cdot U^2 \tag{3}$$

ISSN:1791-4469

Propulsion

The thrust T and torque Q of a propeller with diameter D rotating at a speed n in a fluid with density ρ can be expressed in non-dimensional form as thrust and torque coefficients, K_T and K_Q respectively:

$$K_T = \frac{T}{\rho n^2 D^4} \tag{4}$$

$$K_Q = \frac{Q}{\rho n^2 D^5} \tag{5}$$

The speed at which the propeller is moving in the wake is the speed of advance U_A:

$$U_A = (1-w)\cdot U \tag{6}$$

The parameter w is an average measure of the wake effect and is termed *wake fraction*. The speed of advance can be non-dimensionalized to a parameter known as the advance ratio:

$$J = \frac{U_A}{nD} \tag{7}$$

The rate at which fuel mass is consumed in the engine is:

$$q = sfc(P_B)\cdot P_B \tag{8}$$

The engine break power P_B is related to resistance through the following expression:

$$P_B = \frac{R_{total} U}{\eta_D \eta_{TRM}} \tag{9}$$

The quasi-propulsive efficiency ηD can be expressed as the product:

$$\eta_D = \eta_R \eta_o \eta_H \tag{10}$$

The relative rotative efficiency η_R and hull efficiency η_H are parameters dependent mainly on the geometry of the hull. Although these parameters depend on speed too, the dependence on hull geometry dominates and they are often assumed constant in naval architecture. For the particular example-hull used here their values are 1.035 and 1.097 respectively, interpolated from the propulsion factors of the Series 60 parent models (Todd 1963). The open water propeller efficiency can be expressed in terms of the propeller parameters:

$$\eta_o = \frac{K_T J}{2\pi K_Q} \tag{11}$$

The term η_{TRM} in equation (9) stands for the transmission efficiency, which can be decomposed further into a product of shaft and gearbox efficiency:

$$\eta_{TRM} = \eta_S \eta_{GB} \tag{12}$$

In the example considered here the gearbox and shaft efficiencies are assigned typical values of 0.96 and 0.98 respectively.

Ship Dynamics in Waves

SWAN1 solves the linear forward speed seakeeping problem of the three-dimensional flow around vessels of arbitrary geometry. Linearity allows for the solution of the problem using frequency domain methods. The ambient waves are harmonic and so are the ship motions. Real sea states are modeled by spectral analysis using the frequency domain results from SWAN1. The computational domain for this study is shown in figure 2. The rectangular grid extends 50m from the bow to the upstream free surface boundary, 100m from the stern to the downstream boundary and 140m from the ship's centerline to the lateral boundary.

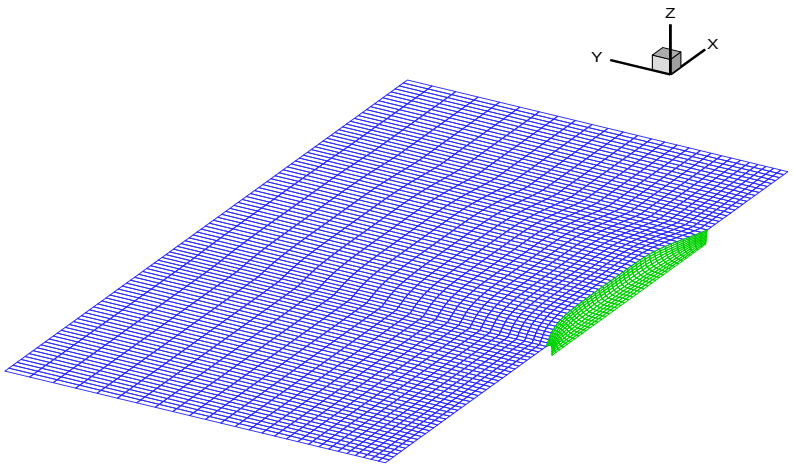


FIGURE 2. SWAN1 Computational Grid.

A ship-fixed reference frame Oxyz with the z=0 plane coinciding with the calm water surface and the positive x-axis coinciding with the positive ship axis on the waterline is related to the earth-fixed OXYZ frame through the Galilean transformation:

$$X = x + Ut \tag{13}$$

$$Y = y \tag{14}$$

$$Z = z \tag{15}$$

The total velocity potential $\Phi(\mathbf{x},t)$ can be decomposed into a sum of two terms:

$$\Phi(\mathbf{x},t) = \overline{\Phi}(\mathbf{x},t) + \varphi(\mathbf{x},t) \tag{16}$$

The first term on the right hand side represents the steady flow potential, which can be further decomposed into a basis flow and a disturbance flow:

$$\overline{\Phi}(\mathbf{x},t) = \bar{\phi}(\mathbf{x},t) + \psi(\mathbf{x},t) \tag{17}$$

The basis flow potential $\bar{\phi}(\mathbf{x},t)$ accounts for the thickness effect of the ship's hull as it

ISSN:1791-4469

encounters the uniform ambient stream of velocity -*U* (the moving ship is equivalently treated as fixed against an oncoming uniform flow). The disturbance potential $\psi(\mathbf{x},t)$ accounts for the presence of the steady wave disturbance that forms the Kelvin wave pattern. The sum of these quantities satisfies the zero normal flux condition on the mean position of the hull. In SWAN1 the linearization of the steady flow is carried out around the double-body basis flow. The free surface is replaced by a rigid wall at z=0, on which:

$$\frac{\partial \bar{\phi}}{\partial z} = 0 \tag{18}$$

$\bar{\phi}(\mathbf{x},t)$ represents the potential of the flow around the hull and its image above the plane z=0. By this representation, the hydrodynamic end effects are modeled more accurately than in typical strip and slender-body theories in which the basis flow is just the ambient free stream. The field equation, the free surface and radiation conditions and the aforementioned boundary conditions formulate a boundary value problem (BVP), which is solved numerically in SWAN1. The solution $\bar{\Phi}$ corresponds to a steady outgoing wave pattern known as the Kelvin wake.

The second term on the right hand side of (16) is the unsteady velocity potential associated with the ambient (harmonic) wave, its interaction with the hull and the resulting ship motions. It can be expressed as:

$$\varphi(\mathbf{x},t) = \Re e\left\{\tilde{\varphi}(\mathbf{x})e^{i\omega t}\right\} \tag{19}$$

In the above expression the harmonic term oscillates with the encounter frequency ω:

$$\omega = \left|\omega_o - kU\cos\beta\right| \tag{20}$$

where ω_o is the absolute ambient wave frequency defined relative to the earth fixed coordinate system and *k* the corresponding wavenumber, given by the dispersion relation in deep water:

$$k = \frac{{\omega_o}^2}{g} \tag{21}$$

Following the standard convention for relative wave direction, the angle β is measured from the stern (β=180^o means head wave).

The complex potential $\tilde{\varphi}$ in (19) is a superposition of the incident wave potential $\tilde{\varphi}_I$, the diffraction potential $\tilde{\varphi}_D$ and the radiation potentials $\tilde{\varphi}_j$, j=1,…,6 for all modes of motion:

$$\tilde{\varphi} = \tilde{\varphi}_I + \tilde{\varphi}_D + \sum_{j=1}^{6}\tilde{\varphi}_j \tag{22}$$

SWAN1 solves the relevant BVP for the diffraction potential $\tilde{\varphi}_D$ and calculates the (complex) excitation forces from the linearized Bernoulli equation extended to account for the modeling of end effects by the double-body basis flow:

$$\tilde{P} = -\rho\left(i\omega - U\frac{\partial}{\partial x}\right)\cdot\left(\tilde{\varphi}_I + \tilde{\varphi}_D\right) \tag{23}$$

$$\tilde{X}_i = \iint_{\bar{S}_B} \tilde{P} n_i dS \tag{24}$$

Furthermore, upon solving the BVP for the radiation potentials $\tilde{\varphi}_j$ SWAN1 provides the complex amplitudes of motion for all six degrees of freedom in the form of Response Amplitude Operators (RAO), normalized by the ambient wave amplitude:

$$RAO_j(\omega) = \frac{\Xi_j(\omega)}{A}, j=1,\ldots,6 \tag{25}$$

Examples of RAOs in heave and pitch in oblique (bow and stern quartering) waves for a Froude number of 0.21 (13 knots) are shown in figures 3 to 6. The RAO's are plotted against the encounter period of the wave.

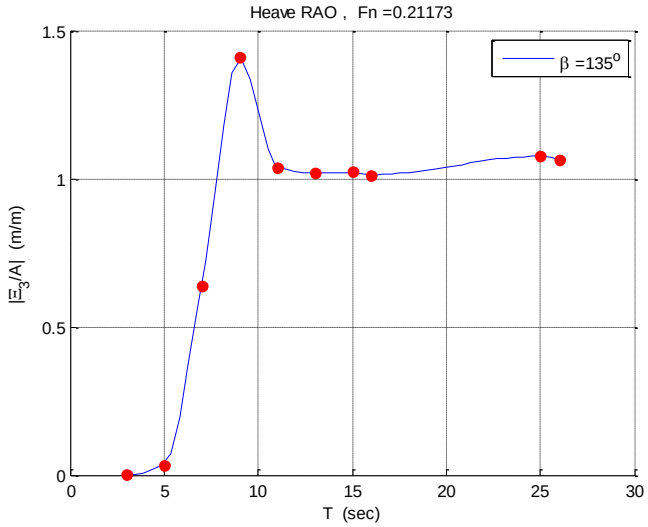


FIGURE 3. Heave RAO in Bow Waves at 13 Knots.

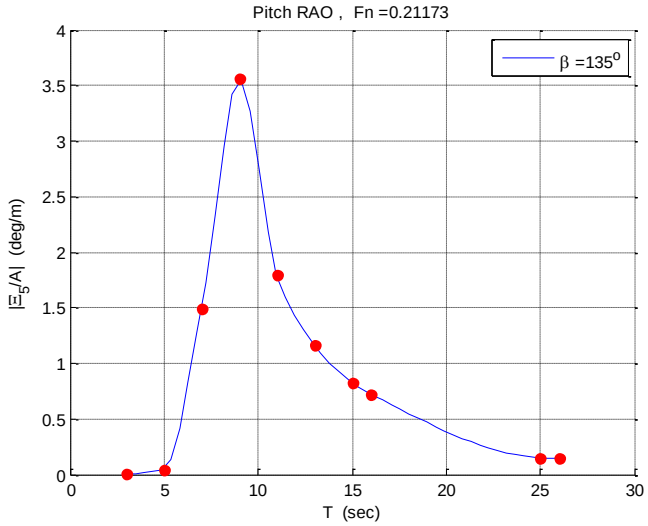


FIGURE 4. Pitch RAO in Bow Waves at 13 Knots.

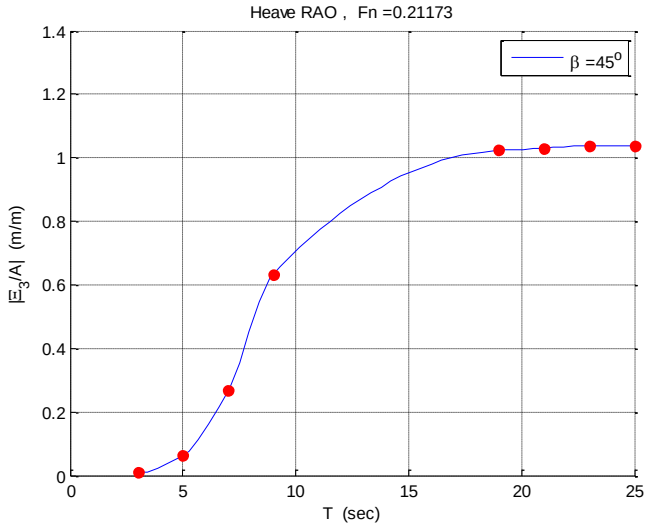


FIGURE 5. Heave RAO in Quartering Waves at 13 Knots.

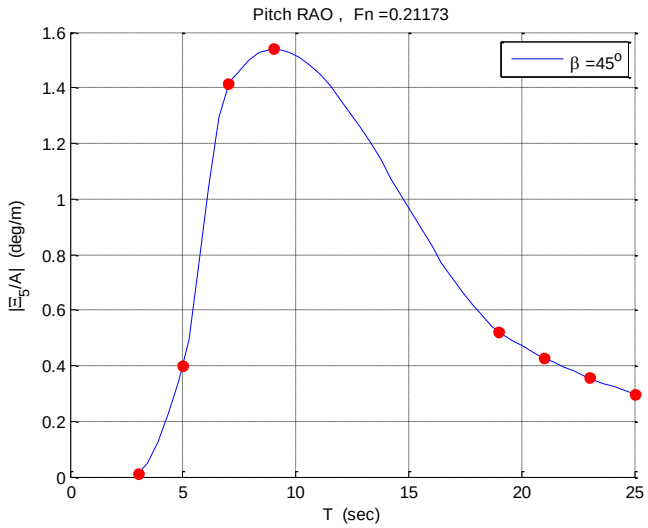


FIGURE 6. Pitch RAO in Quartering Waves at 13 Knots.

Added Resistance in Waves

Added resistance is a hydrodynamic force of paramount significance to the ship routing problem. The effect of waves manifests itself in the fuel consumption through the existence of this additive resistance force. Standard methods exist for the calculation of added resistance, but with relatively narrow range of applicability in terms of wave direction (Gerritsma and Beukelman 1972) or Froude numbers (Faltinsen et al. 1980).

SWAN1 implements the direct pressure integration method and accurately models the hydrodynamic effects near the ends of the ship which are essential for the robust prediction of the added resistance. The total velocity potential is given by (16). The total resistance is the x-component of the hydrodynamic force:

$$R_T = -\iint_{S_B} P n_1 dS_B \tag{26}$$

The above force is obtained by integrating the total hydrodynamic pressure:

$$P = -\rho\left(\frac{\partial \Phi}{\partial t} + \frac{1}{2}\nabla\Phi\cdot\nabla\Phi + gz\right) \tag{27}$$

around the instantaneous submerged hull surface S_B. Replacing Φ by the steady flow potential $\overline{\Phi}$ in (27) and integrating the resulting $\overline{P}$ over the mean surface of the submerged hull, the calm water resistance follows:

$$R_{calm} = -\iint_{\overline{S}_B} \overline{P}\overline{n}_1 dS_B \tag{28}$$

The difference between the total and calm water resistance is oscillatory, time dependent and can be written as the sum of three components:

$$R_{total} - R_{calm} = R_1 + R_2 + R_3 \tag{29}$$

$$R_1 = -\iint_{\overline{S}_B} \delta P \overline{n}_1 dS_B \tag{30}$$

$$R_2 = -\iint_{\overline{S}_B} \overline{P}\delta n_1 dS_B \tag{31}$$

$$R_3 = -\iint_{\delta S_B} \overline{P}\overline{n}_1 dS_B \tag{32}$$

In the above relations, δS_B represents the fluctuation of the mean wetted surface $\overline{S}_B$ that accounts for the ship motions, δP is the correction to the steady pressure P at some rigid point on the instantaneous position of the hull in waves and δn_1 is the difference between the x-component of the unit normal vector at the instantaneous and mean positions of the hull. By expanding (29) in a Taylor series around the mean position of the hull, the quantities δS_B, δP, δn_1 appear explicitly in terms of $\overline{\Phi}$, φ and their gradients evaluated at the mean positions of the hull and the waterline. Since the latter quantities are available from the solution of the linear seakeeping problem, the added resistance can be calculated. For a monochromatic wave of amplitude *A* and frequency *ω*, at an angle of incidence *β* the added resistance is defined as the mean of (29), namely:

$$R_w(\omega,\beta) = \overline{R_1 + R_2 + R_3} \tag{33}$$

In SWAN1 the output is expressed as an Added Resistance Operator (ARO) normalized by the square of the ambient wave amplitude:

ISSN:1791-4469

$$ARO(\omega,\beta)=\frac{R_w(\omega,\beta)}{A^2} \tag{34}$$

Figures 7, 8 plot ARO vs. encounter period for the same wave headings as in figures 3 to 6.

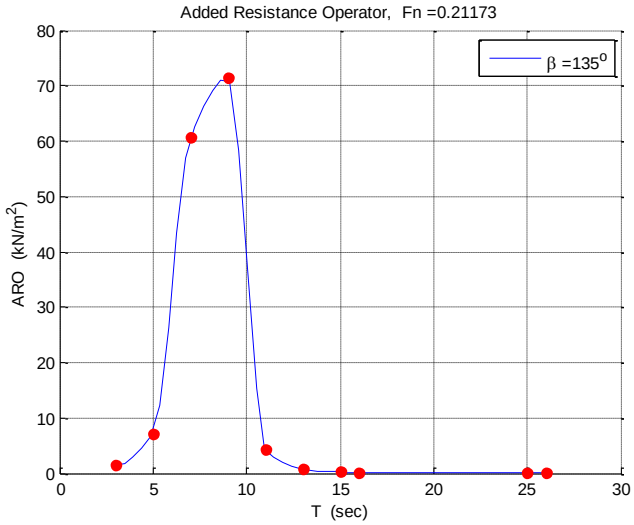


FIGURE 7. Added Resistance Operator in Bow Waves at 13 Knots.

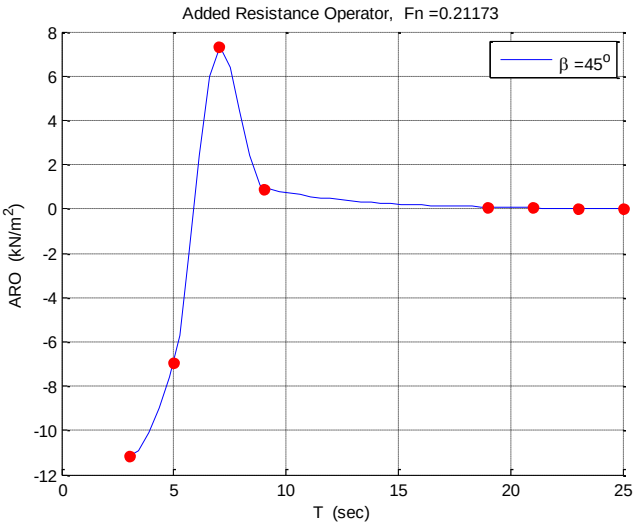


FIGURE 8. Added Resistance Operator in Quartering Waves at 13 Knots.

For irregular seas characterized by a spectrum $S(\omega_0,\theta)$ the mean added resistance can be found by integrating the ARO and the spectral density over all constituent absolute wave frequencies and headings:

$$\bar{R}_w=2\int_0^{\infty}d\omega_0\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}d\theta\,S(\omega_0,\theta)\cdot ARO[\omega(\omega_0),\theta] \tag{35}$$

As explained below, wave and swell are treated as locally unidirectional, propagating along the mean directions computed by the forecasting model. In this setting integration over θ in (35) need not be carried out. In the evaluation of the mean added resistance by (35) the dependence of the ARO on the encounter frequency defined by (20) has been used to transform the integration in (35) over the absolute wave frequency.

SEA STATE MODELING AND FORECASTING

By means of an advanced CFD tool like SWAN1, the ship hydrodynamics in calm water and in waves is properly modeled, as outlined in the preceding section. In addition to resistance and powering, predicting the variability of the ocean environment is a necessity in optimal ship navigation. In this paper the representation of the sea state is derived from the results of the ECMWF global version of the WAM numerical model. The model solves the wave transport equation (i.e. the governing law of wind-wave energy balance) without any assumption on the spectral shape. The solution is the spectral density given at 28 different frequencies and 24 wave directions for every grid point on the global map (Emmanouil et al. 2007). After processing the spectral information, the output contains significant wave height, maximum expected wave height, mean and peak frequencies and mean wave direction. For swell waves the output provides swell height, mean direction and mean frequency. These results constitute the wave and swell forecasts which are promulgated daily in 1-hour time intervals at each grid point. For reasons of computational efficiency the study presented here makes use of the forecast data in 3-hour time steps. The spatial resolution of 0.5×0.5 degrees corresponds to an area of roughly 55×55 km.

Although WAM4 calculates the spectrum, the enormous size of the data files makes it impossible to process the direct information of spectral density $S(\phi,\lambda,\omega,\theta,t)$ as a function of position (i.e. longitude λ -latitude ϕ), frequency ω, direction θ and time t. Practical considerations, therefore, dictate the need to back-fit the output parameters of the model (namely mean frequency and wave height) to a standard spectrum which is anticipated to represent the actual sea state reasonably well. A bi-modal Bretschneider spectrum is selected for this purpose. By virtue of linearity, two separate single-peaked Bretschneider spectra can be superimposed, one describing the local storm (i.e. wind-generated waves) the other describing the swell. The expression for each of them is:

$$S(\omega) = 0.278\frac{\bar{\omega}^4}{\omega^5}\bar{H}^2 e^{-0.437(\bar{\omega}/\omega)^4} \tag{36}$$

where $\bar{\omega}$ and $\bar{H}$ stand for the mean frequency and height of the wave (or swell) respectively. $\bar{H}$ is related to the significant wave height H_s via:

$$\frac{H_s}{\bar{H}} = 1.6 \tag{37}$$

The forecast time interval of 3 hours corresponds to a time window for which the assumption of stationarity of the sea state holds well. Stationarity, in turn, permits the utilization of frequency domain methods and spectral analysis in hydrodynamics which together with linearity allow the calculation of the mean added resistance from expression (35) and the estimation of the ship response statistics from analogous expressions. For example, the variance of heave motion (mode 3 in SWAN1) can be easily calculated, given the wave

ISSN:1791-4469

spectrum and the RAO in heave:

$$\sigma_3^2 = \int_0^\infty S(\omega_o)\left|RAO_3(\omega_o)\right|^2 d\omega_o \tag{38}$$

The variances of other derived responses may be evaluated in a similar manner and used in the equations of the constraints in the optimal control problem in question.

All hydrodynamic attributes of a particular vessel of interest in the minimum fuel routing problem are contained in the ship response RAOs and the added resistance ARO. These quantities are computed once, in advance, as functions over all absolute wave frequencies, ship speeds and headings in unidirectional waves and stored in a hydrodynamic database customized for each vessel. They are subsequently used in expressions (35) and (38) together with the wave spectra supplied by the weather forecasts. As will be seen below the mean added resistance and RMS values of the ship responses are the quantities that enter the minimum fuel routing problem and their evaluation is very efficient given the ARO and RAOs which are retrieved from the hydrodynamic database.

Treating the spectrum as locally unidirectional in the vicinity of each position on the map (grid point **x**)**,** the variance in equation (38) and the added resistance in (35) are both functions of the local prevailing direction of propagation θ_0 of the sea state (i.e. the mean direction). Technically, the local directionality of the sea state is modeled by multiplying the spectral density by a "spreading function", which in this setting is a delta function located at θ_0:

$$S(\omega_0,\theta) = \delta(\theta-\theta_0)\cdot S(\omega_0) \tag{39}$$

Integration of an expression like (35) over all angles with the spectrum cast in the form of (39) eliminates all but one direction θ_0. If the assumption of long-crested waves propagating in one direction is lifted, some directional spreading should be included in the model at the higher computational cost incurred by the double integration in (35). The use of different spreading functions (i.e. delta, $\cos^2()$, $\cosh^{-2}()$, Hasselmann formula, Mitsuyasu formula etc.) and their effect on the optimal solution and computational time will be addressed in future work.

MINIMUM FUEL ROUTING

Powering and Fuel Consumption in Rough Seas

The equations of motion of the ship sailing in the ocean can be written in spherical coordinates as:

$$\dot{\phi}(t) = \frac{1}{\bar{R}_{earth}}\cdot U(t)\cdot\cos\left(p(t)\right) \tag{40}$$

$$\dot{\lambda}(t) = \frac{1}{\bar{R}_{earth}\cdot\cos\left(\phi(t)\right)}\cdot U(t)\cdot\sin\left(p(t)\right) \tag{41}$$

The speed of the vessel is denoted by *U* and the course is denoted by *p* (measured relative to the true north). *U* refers to the speed through water (STW), which in the absence of currents coincides with the speed over ground (SOG). The latter is accurately measured by the GPS receivers on board. Ocean currents are not considered in the present study, but their inclusion in the model is straightforward. The spherical coordinates in the above equations are the latitude

$\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, measured from the equator (positive north), and the longitude $\lambda \in (-\pi, \pi]$ measured from the Greenwich meridian (increasing eastbound). $\bar{R}_{earth}$ is the mean radius of the earth. The differential distance between positions (ϕ, λ) and $(\phi + d\phi, \lambda + d\lambda)$, if measured along the rhumb line (loxodrome) connecting the two locations, is given by:

$$dS = U(t) \cdot dt = \bar{R}_{earth} \cdot \sqrt{d\phi(t)^2 + \cos^2\left(\phi(t)\right) \cdot d\lambda(t)^2} \tag{42}$$

The loxodrome is the constant heading route connecting any two locations on the sphere. On a Mercator map this route appears as a straight line. On the other hand, the shortest distance between two points on the sphere is obtained if these two points are connected with a great circle arc. This is the orthodrome route. The great circle distance between points (ϕ_2, λ_2) and (ϕ_2, λ_2) is:

$$S = \bar{R}_{earth} \cdot \cos^{-1}\left(\sin(\phi_1)\sin(\phi_2) + \cos(\phi_1)\cos(\phi_2)\cos(\lambda_2 - \lambda_1)\right) \tag{43}$$

It is customary in navigation to approximate the great circle route with a sequence of rhumb line segments. This is the approach adopted throughout this paper.

The position vector $\mathbf{x}(t) = \left(\phi(t) \quad \lambda(t)\right)^T$ is the state vector of the system, while $\mathbf{u}(t) = \left(U(t) \quad p(t)\right)^T$ is the control vector. Assuming quasi-steady conditions, the thrust delivered by the propulsor is balanced by the total resistance:

$$T(U, p, \phi, \lambda, \tilde{t}) = \frac{R_{calm}(U) + \bar{R}_w(U, p, \phi, \lambda, \tilde{t})}{1 - t} \tag{44}$$

where $\tilde{t}$ denotes time to distinguish it from the thrust deduction factor defined as $t = \frac{T - R}{T}$. Combining equations (4), (6), (7) and (44) yields the propeller load curve:

$$K_T = \frac{R_{calm}(U) + \bar{R}_w(\mathbf{u}, \mathbf{x}, \tilde{t})}{\rho(1-t)(1-w)^2 U^2 D^2} J^2 \tag{45}$$

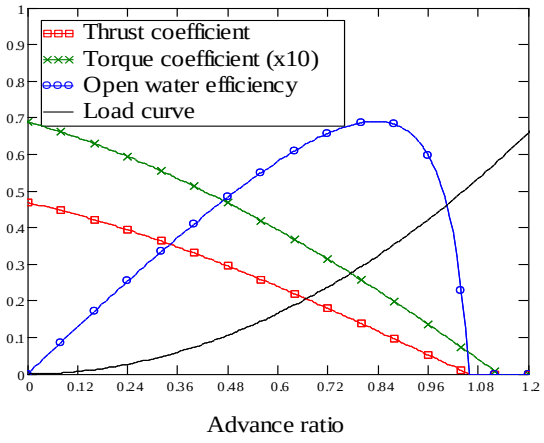


FIGURE 9. Typical B-Series Chart. Intersection of load curve with thrust coefficient curve uniquely defines all the operational parameters of the propeller.

ISSN:1791-4469

Equation (45) together with the K_T curve of the B-series chart define a nonlinear system of equations for the unknowns K_T and J. Upon solving this system, the parameters K_Q and η_o are readily calculated from (5) and (11) respectively. Figure 9 shows a graphical solution on the chart. The point of intersection of the load curve with the thrust coefficient curve uniquely determines the advance ratio J, from which the corresponding K_Q and η_o can be read off the chart. From equations (8)-(12), the fuel rate can be expressed formally as a function of position in the ocean, control setting (i.e. speed and course) and time:

$$q(\mathbf{u},\mathbf{x},\tilde{t}) = sfc \cdot \frac{R_{calm}(U) + \bar{R}_w(\mathbf{u},\mathbf{x},\tilde{t})}{\eta_R \cdot \eta_o(\mathbf{u},\mathbf{x},\tilde{t}) \cdot \eta_H \cdot \eta_S \cdot \eta_{GB}} \cdot U \tag{46}$$

The product of the last two terms in (46) is the engine load (break power). The dependence of the term *sfc* on engine load has been suppressed for brevity. Figure 10 demonstrates this dependence for the MAK 9M25C engine.

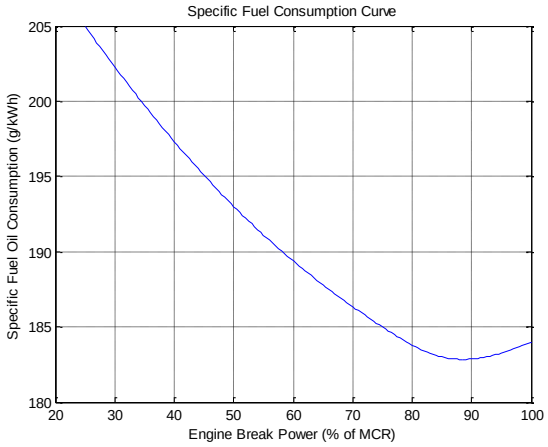


FIGURE 10. Specific Fuel Consumption as a Function of Engine Load (data source: ref [22]).

The integral of (46) over the entire duration of the trip defines a scalar performance index:

$$I = \int_0^{t_f} sfc \cdot \frac{R_{calm}(U) + \bar{R}_w(\mathbf{u},\mathbf{x},\tilde{t})}{\eta_R \cdot \eta_o(\mathbf{u},\mathbf{x},\tilde{t}) \cdot \eta_H \cdot \eta_S \cdot \eta_{GB}} \cdot U d\tilde{t} \tag{47}$$

The above index quantifies the total amount of fuel which has been consumed upon reaching the fixed final time t_f. It is the cost function to be minimized in the optimization process.

Inequality Constraints

While seeking to minimize (47) there are certain limiting factors which in a typical optimal control problem appear as inequality constraints. Violation of these constraints is prohibited at

all times and this restriction changes the structure of the solution.

State Constraints, namely limits in ϕ and λ, could represent regions of forbidden navigation, such as shallow waters, land or ice. The numerical simulation in this study was carried out in the Atlantic Ocean below the arctic circle, so the possible influence of shallow waters and ice is not investigated.

Control Constraints are bounds imposed by the propulsion plant capabilities. For example, the particular ship cannot exceed 15.45 knots in calm seas with the given engine at the particular loading condition, assuming clean hull. In addition, the engine load (an implicit function of the controls) is not permitted to exceed an envelope function $g(P_B,n_e)$ which is depicted in figure 11, otherwise the engine will be overloaded. P_B and n_e represent the engine load and speed respectively.

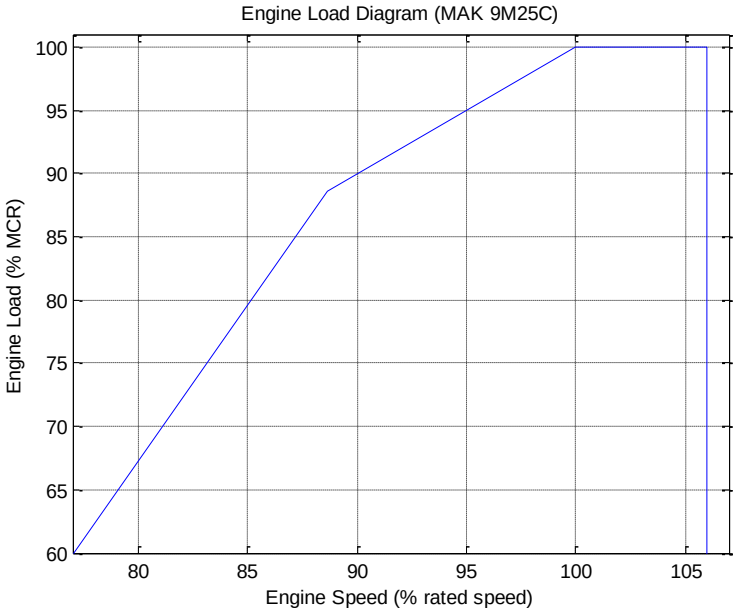


FIGURE 11. Engine Operating Envelope (data source: ref. [22]).

Therefore, the control bounds can be described as follows:

$$P_B(U,p) \le g(P_B,n_e) \tag{48}$$

$$U_{\min} \le U \le U_{\max} \tag{49}$$

In principle, the lower bound for speed should be zero. However, in the optimization example described below it was found that there is no gain in lowering the speed below 7 knots. As verified by SWAN, the responses in speeds lower than 7 are practically those of the ship at zero speed. Regardless of the initial conditions or the trial forecasts, it was observed that the program would always choose to avoid severe weather by altering the ship's course while retaining speed above 7. Furthermore, there exists a lower bound for engine speed to avoid excessive vibrations. In order to save computational effort the lower bound $U_{\min}$ was set to 7 knots. This limit corresponds (in calm water) to an engine speed of 311 revs/min, well above the

ISSN:1791-4469

minimum allowable speed of 250 revs/min for this particular engine type (ref. [22]).

Safety Constraints are enforced to ensure the secure transportation of the cargo and the physical integrity of the ship herself, while moving in rough seas. Quantifying safety is not a trivial task. The approach proposed here is to enforce “hard” inequality constraints on the probabilities of occurrence of certain extreme events associated with sea severity. Two such events are green water on deck (or deck wetness) and slamming. Other motion-related constraints could also be specified, at increased computational cost, depending on the type of vessel and nature of risks that must be taken into account when sailing in severe weather. For instance, parametric rolling is a dangerous situation that should be avoided especially in container transportation. In what follows, only restrictions on deck wetness and slamming have been considered, as they are deemed sufficient to demonstrate the concept and the benefits of constrained optimal routing. It should be noted that these two events are related to pitching and heaving, two longitudinal modes of ship motion which can be predicted by SWAN1 with remarkable accuracy. Considering a point Π_1 in the uppermost part of the bow (bulwark), deck wetness occurs when the relative motion exceeds the freeboard f at point Π_1. The probability of this event follows the Rayleigh distribution (Ochi 1998) and is given by:

$$P(\text{water on deck}) = e^{-\frac{f^2}{2\sigma_r^2}} \tag{50}$$

The variance of the relative motion, σ_r^2, is calculated by an expression similar to (38). Slamming occurs when the relative motion at a point Π_2 in the keel near the bow exceeds the draft H at that point and when the relative velocity exceeds a critical value:

$$V_{cr} = 0.093\sqrt{g \cdot L} \tag{51}$$

The critical velocity V_{cr}, gravitational acceleration g and waterline length L enter equation (51) in SI units. Assuming that the two events are statistically independent, the probability of slamming follows the Rayleigh distribution:

$$P(\text{slamming}) = e^{-\left(\frac{H^2}{2\sigma_r^2}+\frac{V_{cr}^2}{2\sigma_V^2}\right)} \tag{52}$$

where σ_r^2 and σ_V^2 are the variances of the relative motion and velocity, respectively, at point Π_2.

When the frequency of occurrence of slamming or deck wetness rises above certain limits, the motions are so severe that the ship’s safety is compromised and voluntary speed reduction is effected. The limiting probabilities assume the following values (Faltinsen 1990):

$$P(\text{slamming}) \le 0.03 \tag{53}$$

$$P(\text{water on deck}) \le 0.07 \tag{54}$$

Substituting (50) and (52) in (53) and (54), respectively, and taking the logarithms of both sides the safety constraints take the form of inequalities involving the RMS values of the relative motion and velocities at selected points on the hull. The RMS values of these and other derivative seakeeping quantities may be evaluated easily from their definition and the use of expressions analogous to (38).

Formulation of the Optimal Routing Problem

The complete formulation can be summarized in the standard format for constrained optimization problems:

minimize $$I = \int_0^{t_f} sfc \cdot \frac{R_{calm}(U) + \bar{R}_w(\mathbf{u}, \mathbf{x}, \tilde{t})}{\eta_R \cdot \eta_o(\mathbf{u}, \mathbf{x}, \tilde{t}) \cdot \eta_H \cdot \eta_S \cdot \eta_{GB}} \cdot U d\tilde{t}$$

subject to:

Dynamic constraint:

$$\dot{\phi}(t) = \frac{1}{\bar{R}_{earth}} \cdot U(t) \cdot \cos\left(p(t)\right)$$

$$\dot{\lambda}(t) = \frac{1}{\bar{R}_{earth} \cdot \cos\left(\phi(t)\right)} \cdot U(t) \cdot \sin\left(p(t)\right)$$

Control bounds:

$$P_B\left(U, p\right) \le g\left(P_B, n_e\right)$$

$$U_{\min} \le U \le U_{\max}$$

Safety constraints:

$$P(\text{slamming}) = e^{-\left(\frac{H^2}{2\sigma_r^2} + \frac{V_{cr}^2}{2\sigma_V^2}\right)} \le 0.03$$

$$P(\text{water on deck}) = e^{-\frac{f^2}{2\sigma_r^2}} \le 0.07$$

initial and terminal conditions:

$$\mathbf{x}(0) = \begin{pmatrix} \phi(0) \\ \lambda(0) \end{pmatrix} = \begin{pmatrix} \phi_0 \\ \lambda_0 \end{pmatrix}$$

$$\mathbf{x}(t_f) = \begin{pmatrix} \phi(t_f) \\ \lambda(t_f) \end{pmatrix} = \begin{pmatrix} \phi_f \\ \lambda_f \end{pmatrix}$$

and final time:

t_f prescribed

In order to solve the problem numerically the discrete counterparts of the cost function and

ISSN:1791-4469

dynamic constraint are needed. The sailing time is divided into *N* stages of length:

$$\Delta t = \frac{t_f}{N} \tag{55}$$

The equations of motion are approximated as:

$$\phi(k+1) = \phi(k) + \frac{1}{\bar{R}_{earth}} \Delta t U(k) \cos p(k) \tag{56}$$

$$\lambda(k+1) = \lambda(k) + \frac{1}{\bar{R}_{earth} \cdot \cos\left(\phi(k)\right)} \Delta t U(k) \sin p(k) \tag{57}$$

with $\phi(1) = \phi_0$, $\lambda(1) = \lambda_0$, $\phi(N+1) = \phi_f$, $\lambda(N+1) = \lambda_f$ and $k = 1, 2, ..., N$.

The integrand in (47) is the Lagrangian. In discrete form the integral is approximated by a sum and the performance index becomes:

$$I = \sum_{k=1}^{N} q\left(U(k), p(k), \phi(k), \lambda(k), k\right) \Delta t \tag{58}$$

NUMERICAL SOLUTION

Dynamic Programming

Dynamic Programming (DP) was introduced by Richard Bellman in the 50's and has proven to be an invaluable tool in optimal control, especially for problems that do not admit an analytical solution. Many such problems do not satisfy optimality or existence and uniqueness conditions. In these cases the classical calculus of variations formulation falls short. Yet, these problems may possess an optimal solution which can be found by DP. Mathematical principles of the method can be found in Bellman (1957). A standard DP algorithm is presented in Kirk (1970).

Iterative Dynamic Programming (IDP) is a recent evolution of the same idea. It applies the recurrence relation known as the Bellman equation in an iterative fashion, converging to the solution in each iteration. The IDP algorithm can be implemented with surprisingly few grid points, thus eliminating the most prominent weakness of the standard DP algorithm: as the number of dimensions grows the vast size of memory storage becomes prohibitive. In fact, the IDP version adopted in this paper uses only a single grid-point. The algorithm outlined below in 10 steps, is adapted by Luus (2000):

1. Discretize the time interval $[0, t_f]$ into N time stages of equal length Δt. Δt should be chosen to coincide with the forecast interval. If $t_f \neq N \cdot \Delta t$, an additional time stage is required of length $\Delta t' = t_f - N \cdot \Delta t$.

2. Make an initial "guess" for the entire control sequence $\mathbf{u}^*(k)$ $(k = 1, ..., N)$. In the iteration following the first, this will be the optimal control obtained in the previous iteration. Using $\mathbf{u}^*(k)$

integrate (40)-(41) to generate an initial nominal trajectory $\mathbf{x}^*(k)$, where $k=1,...,N+1$. Alternatively, one can "guess" the nominal trajectory $\mathbf{x}^*(k)$ and compute the corresponding initial control history $\mathbf{u}^*(k)$.

3. Choose the number of controls c_r $(r=1,2,...,m)$ for each element of the mx1 control vector $\mathbf{u}$. The total number of controls is $C=c_1\cdot c_2\cdot ...\cdot c_m$. Choose the control region contraction factor γ.

4. Choose an increment $\Delta\mathbf{u}$ that determines the span of control region around the central value $\mathbf{u}^*$, i.e.

$$^{j}\mathbf{u}^* \pm \Delta\mathbf{u}\cdot {}^{j}R \tag{59}$$

The left superscript j denotes iteration number. The quantity ^{j}R is a scaling factor applied to $\Delta\mathbf{u}$. It is responsible for narrowing the control region around $^{j}\mathbf{u}^*$ in every iteration. For the first iteration $^{1}R=1$.

5. Discretize each element r of $\Delta\mathbf{u}$ into $\frac{c_r}{2}$ quantized values. These values are selected randomly in the range $\Delta\mathbf{u}$.

6. Start the iterations by setting $j=1$, $^{j}R=1$.

7. Move two steps back from the terminal state to the beginning of stage $N-1$. This corresponds to time $t_f-2\Delta t$. Integrate forward along the next stage, that is from time $t_f-2\Delta t$ to time $t_f-\Delta t$, C times using C different values of $\mathbf{u}$:

$$\mathbf{u}(N-1)={}^{j}\mathbf{u}^*(N-1)\pm\Delta\mathbf{u}\cdot {}^{j}R \tag{60}$$

If any of the elements of $\mathbf{u}(N-1)$ falls outside the allowable values for control, clip them at the upper or lower bound as appropriate. The integration starts from the known state $^{j}\mathbf{x}^*(N-1)$ of the current optimal path. Calculate the reached state $\mathbf{x}^r(N)$ $(r=1,2,...,C)$ and integrate along the final stage to close in on the (fixed) destination point (i.e. the terminal state). Calculate the associated performance index I for each of the C different controls for this two-leg route. Find the r for which I is minimized and store the corresponding $\mathbf{u}^r(N-1)$ as $^{j+1}\mathbf{u}^*(N-1)$.

8. Move one step further back, to the beginning of stage $N-2$ (corresponding to time $t_f-3\Delta t$). Starting from state $^{j}\mathbf{x}^*(N-2)$ integrate one step forward, up to time $t_f-2\Delta t$, using C different values of $\mathbf{u}$:

$$\mathbf{u}(N-2)={}^{j}\mathbf{u}^*(N-2)\pm\Delta\mathbf{u}\cdot {}^{j}R \tag{61}$$

Calculate the reached state $\mathbf{x}^r(N-1)$. For every r, carry out the integration of the remaining trajectory up to the terminal state. Use the optimal control $^{j+1}\mathbf{u}^*(N-1)$ derived in the previous step to integrate. Calculate the performance index for this segment of the path (i.e. from time $t_f-3\Delta t$ to t_f). Among the C different values, store the minimum I and the corresponding control as $^{j+1}\mathbf{u}^*(N-2)$.

9. Repeat steps 7 and 8 until the initial state is reached (initial condition). Integrate forward along the complete path using all available combinations of control as before and determine the

ISSN:1791-4469

best one for the initial stage, i.e. ${}^{j+1}\mathbf{u}^*(1)$. This concludes the first iteration. A complete control sequence ${}^{j+1}\mathbf{u}^*(k)$, $(k=1,2,...,N)$ and an optimal trajectory ${}^{j+1}\mathbf{x}^*(l)$, $l=1,2,...,N+1$ are now available for the next iteration.

10. Reduce the size of the control region by setting ${}^{j+1}R=\gamma\,{}^{j}R$. Increase the iteration index j by 1 and repeat the algorithm starting from step 6.

The above algorithm assumes that the whole voyage duration t_f is covered by a wave/swell forecast. If the duration of the voyage exceeds the maximum forecast time, as is the case in the example below, the algorithm needs to be adapted to assume calm water conditions for the segment of the route beyond forecast coverage. In this case, the remaining part of the route will be the great circle arc to the destination point. Figure 12 shows the extent of forecast coverage along the nominal great circle route.

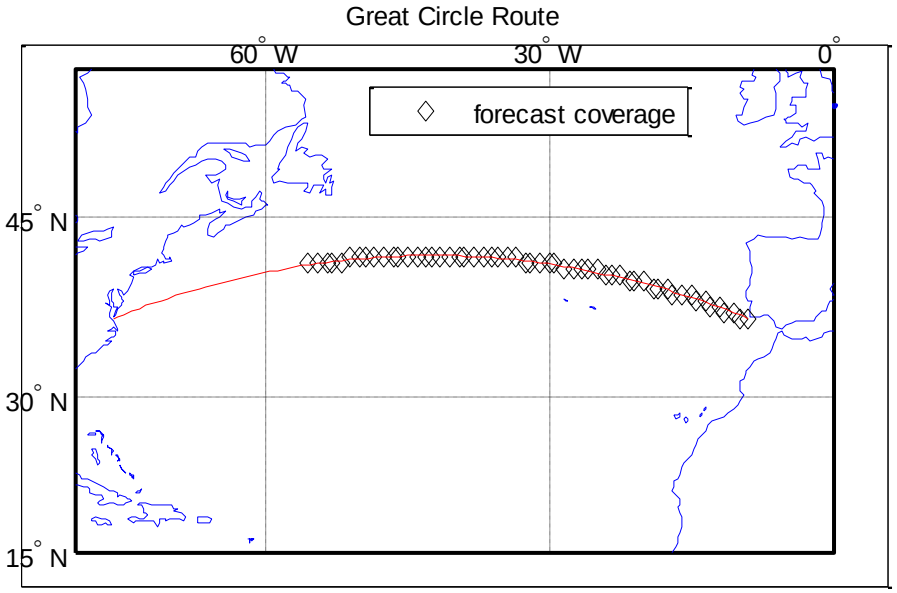


FIGURE 12. Forecast Grid Points Closest to the Great Circle Route.

Assumptions

The length of time stages, as defined in the algorithm above, is chosen to coincide with the forecast interval of 3 hours. Such a short-term description of a sea state allows much confidence in the stationarity assumption discussed earlier. The control policy (i.e. speed and course) is piecewise constant over this time step. No accelerations are taken into account. The transients associated with changing the control input (e.g. engine rpm) are very short relative to the problem's time scale and are neglected. Other components of resistance (such as appendage drag, air resistance, trim or steering resistance) are not included in the analysis. These are either steady components which can be easily incorporated in the model or relatively small in magnitude. The effect of currents is also omitted, as discussed above.

At each stage (time) and for every state (spatial location) the program calculates fuel consumption and checks for constraint violation. For that, it draws information from stored forecast data. In the particular example presented here, the global forecast uploaded on the FTP server of UoA on October 9, 2014 contains the relevant information. For the fraction of the

trip not covered by the forecast calm water is assumed. Ship response and added resistance data are obtained from SWAN1, via the constructed hydrodynamic database, for 10 different speeds (from 7 to 16 knots) and 5 different headings (from 0 to 180 degrees on either side). The actual data are stored in memory in the form of large matrices. Intermediate values are obtained by interpolation. The number of iterations is set to 20 and γ is chosen to be 0.7. The discretization of controls yields a total of C=50 combinations of speed and course (5 speed settings and 10 course settings are used in the optimization code).

A longer trip, on the order of several weeks, can be handled by the IDP algorithm as well. In this case, more than one 168-hour forecasts are needed to cover the whole voyage duration. The algorithm is executed at the starting time using the available forecast at that time. The following day, as soon as the next forecast becomes available, the algorithm is executed again with different initial conditions. These are naturally the coordinates of the ship's new position. The final time is also adjusted accordingly. Now the part of the total route not covered by a sea state prediction is shorter. The program is run daily using updated forecasts in a similar fashion until the destination is reached. This process may be executed very efficiently onboard the vessel. It can, in fact, be implemented in short duration voyages as well.

RESULTS AND DISCUSSION

A trans-oceanic passage would normally be planned along the great circle route. In calm seas this is the minimum fuel route, because it is the shortest one. Regardless, the rhumb line route is often considered by the navigator as a reasonable alternative, especially when the great circle routes pass closer to the poles and through more severe seas and/or navigational hazards (e.g. icebergs). Table 4 below presents a comparison of calculation results for the two routes in the numerical example worked out in this text. It can be seen that the loxodrome (rhumb line) is 2.3% longer than the orthodrome (great circle) and about 2.5% more fuel-expensive in calm seas. In rough seas, on the other hand, computation of fuel consumption when sailing the great circle route shows a 45.5% increase in fuel cost compared to the respective calm seas quantity. For the rhumb line route the increase of fuel cost in waves is 14%.

TABLE 4. Great Circle vs. Rhumb Line Routes.

Description	Unit	Value
Great Circle (GC) Distance	naut.miles	3138.6
Rhumb Line (RL) Distance	naut.miles	3210.5
Nominal Voyage Time	days	10.06
Calm Water Speed (GC)	knots	13.0
Calm Water Speed (RL)	knots	13.3
Calm Water Fuel Cost (GC)	tonne	66.89
Calm Water Fuel Cost (RL)	tonne	68.59
Predicted Fuel Cost (GC)	tonne	97.31
Predicted Fuel Cost (RL)	tonne	78.24

Figures 13 to 16 show the fraction of each trajectory in which some safety constraint is violated. In the middle part of the great circle route the WAM forecast predicts wind waves and swell with significant height locally in excess of 9m and 4m respectively. For the rhumb line the predicted figures are about 3.5m and 3m respectively. In both routes the ship encounters these

ISSN:1791-4469

sea states from head and bow directions. The rhumb line route violates only the deck wetness limiting probability of occurrence.

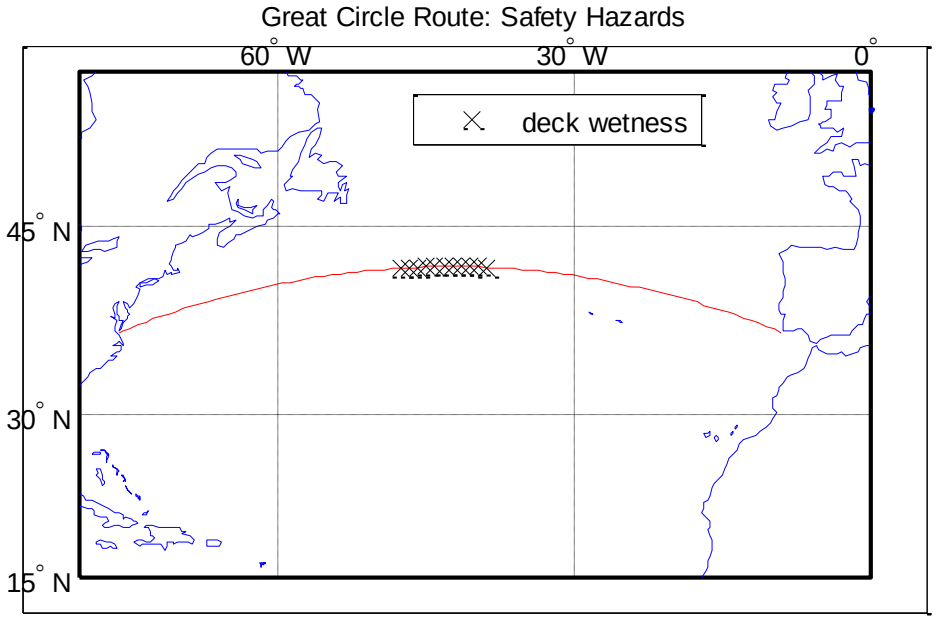


FIGURE 13. Violation of Deck Wetness Constraint along the Orthodrome.

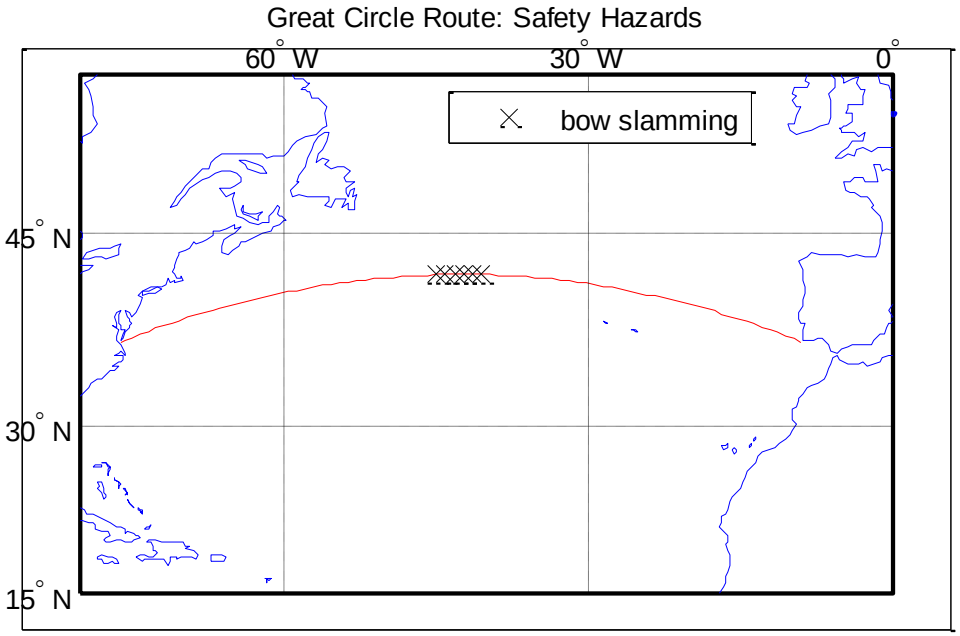


FIGURE 14. Violation of Slamming Constraint along the Orthodrome.

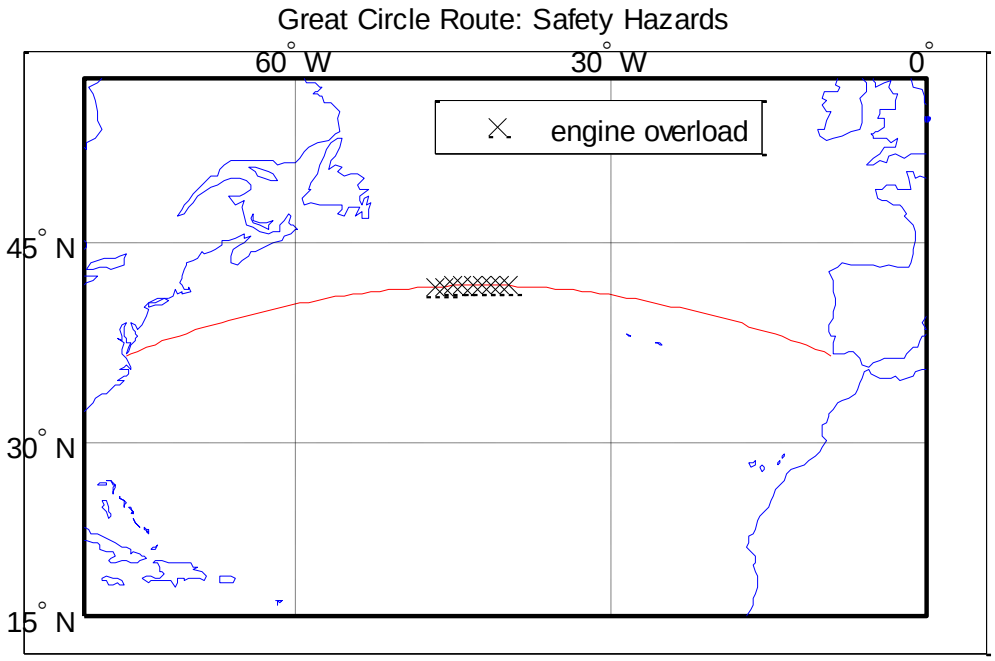


FIGURE 15. Violation of Engine Overload Constraint along the Orthodrome.

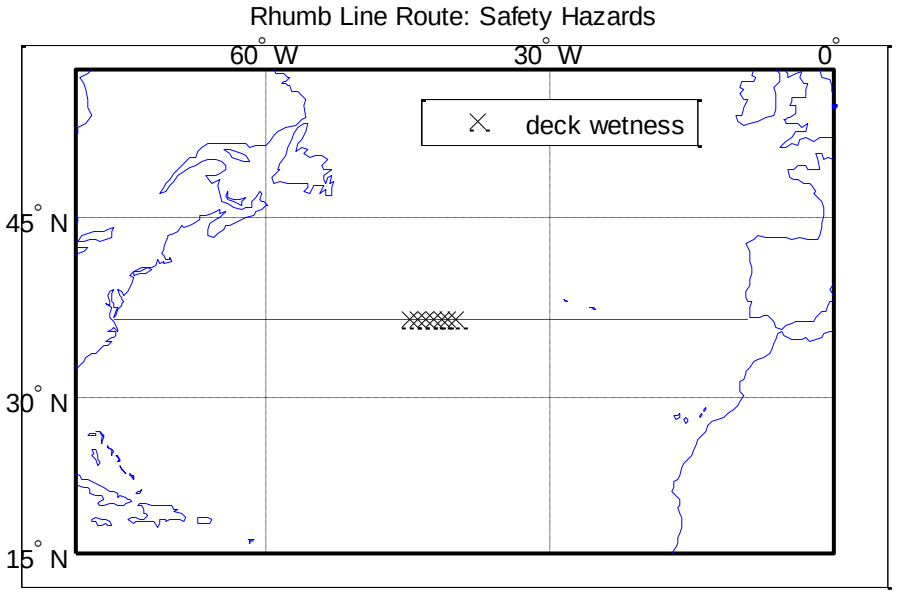


FIGURE 16. Violation of Deck Wetness Constraint along the Loxodrome.

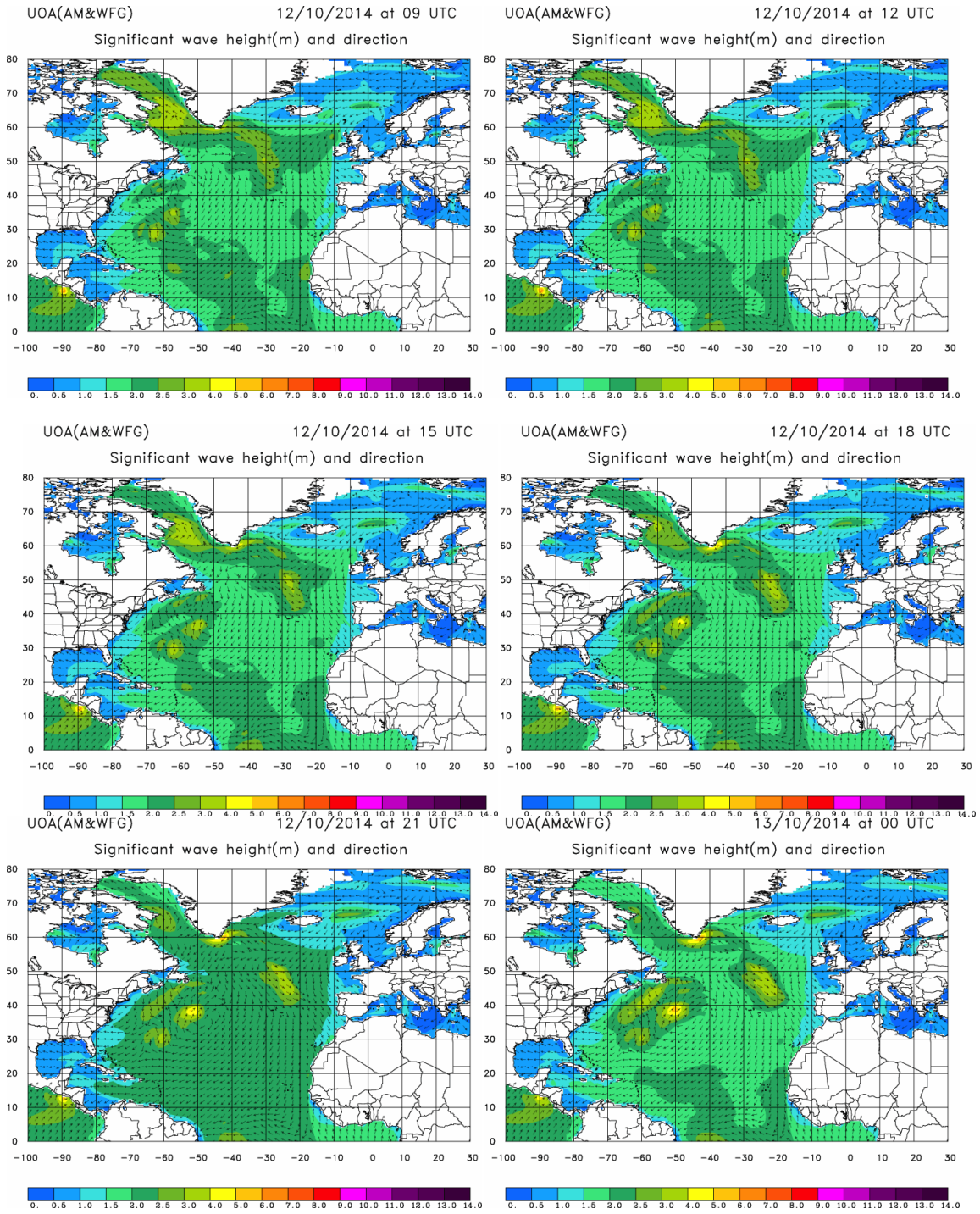


FIGURE 17. Evolution of Wind and Swell Wave System as Forecasted between 105h and 120h of Voyage Time.

The algorithm described in the previous section is coded in MATLAB® to generate the fuel minimizing route. First, the code is executed without any constraint considerations. The orthodrome enters the computations as the nominal route (initial guess). The output is depicted in figure 18. The computed optimal route, after 25 iterations, saves 19.8% more bunker fuel than the nominal great circle but only marginally less than the rhumb line (around 0.3%). However,

this route exceeds the specified threshold for deck wetness probability in the portion designated by 'x' marks. This would normally be an infeasible route in the constrained counterpart of the optimization problem, hence it would be ruled out during the code execution. To verify this, the optimization code is run again checking for constraint violation this time. The result is graphically reproduced in figure 19. The optimal control histories for ship's course and speed settings are shown in figures 20 and 21 respectively. Evidently, the optimal route turns sharply towards the north while slowing down at the same time, in order to avoid the high head/bow seas around the middle segment of the trajectory. Then it turns back towards the previous track. After the expiration of the available forecast the gap to the destination point is bridged with a great circle arc.

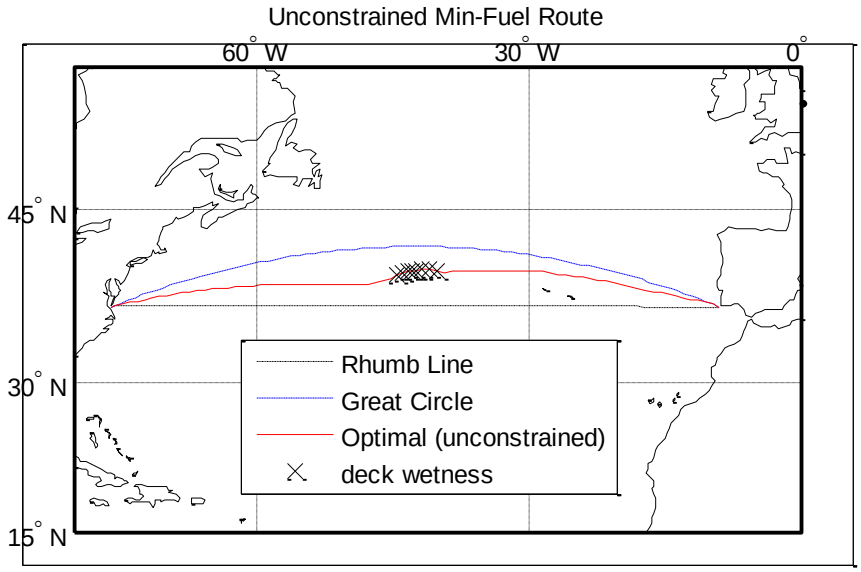


FIGURE 18. Optimal Route of the Unconstrained Minimum Fuel Problem.

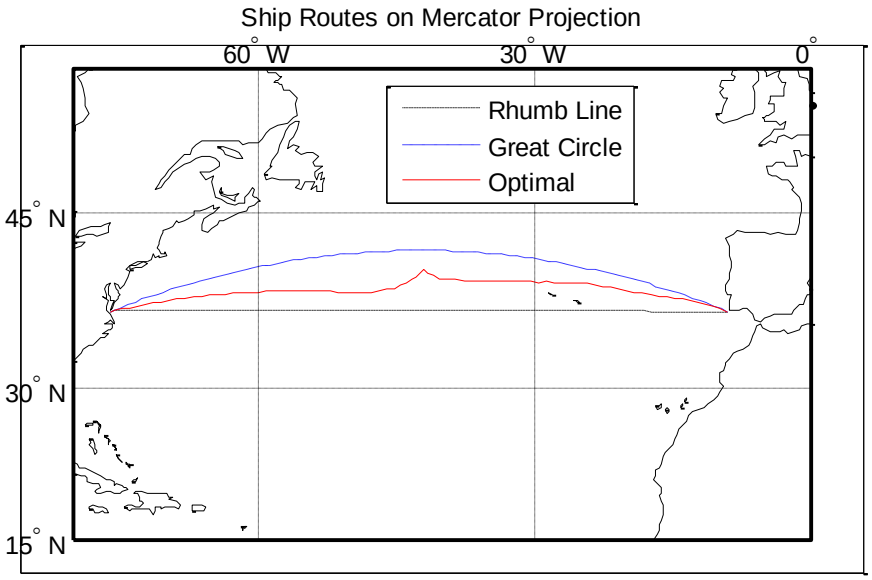


FIGURE 19. Optimal Route of the Constrained Minimum Fuel Problem.

ISSN:1791-4469

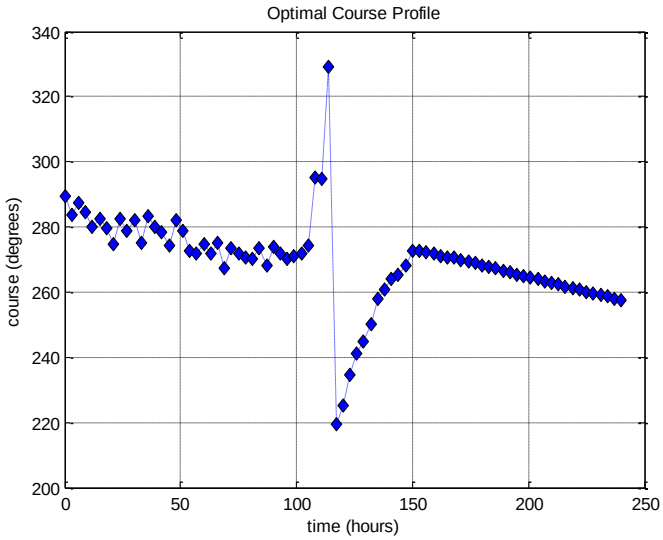


FIGURE 20. Optimal Control History for Ship's Heading.

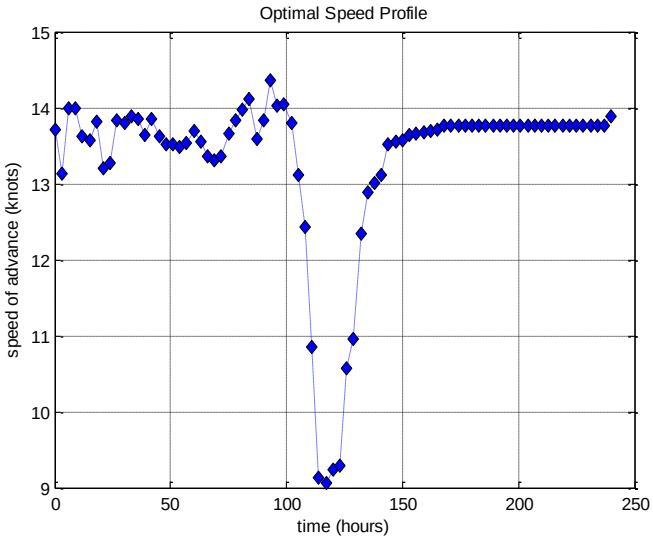


FIGURE 21. Optimal Control History for Ship's Speed.

The structure of the computed optimal solution is intuitive, if one observes the sea state evolution in the mid Atlantic during the voyage of interest (figure 17), bearing in mind that the ship's freeboard is only 2.85m. In this numerical experiment the behavior of the optimal solution is explored under the influence of longitudinal motion-related constraints only. As a result, turning the ship's bow away from the waves both reduces added resistance and alleviates heaving and pitching motions.

Convergence, Accuracy and Sensitivity

Speed of convergence and accuracy are both dominated by the control space *C* and the contraction factor *γ*, as defined above in the IDP algorithm. Convergence is illustrated in figure

22 for the constrained optimization problem. To expedite convergence, the optimal solution computed before for the unconstrained problem was used as the initial route in the algorithm. A discussion about tuning the IDP parameters (C, γ, number of iterations) can be found in Avgouleas (2008). The particular optimization scheme with the IDP settings mentioned above was found to be the best compromise between speed of execution and accuracy of the solution.

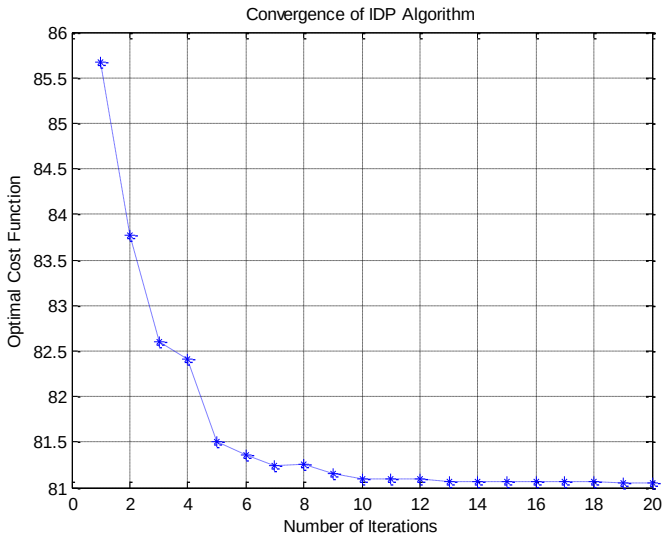


FIGURE 22. Convergence to the Optimal Solution with γ=0.7.

To assess the sensitivity of the solution to the imposed safety constraints several trials were conducted relaxing or lifting one constraint at a time and then both simultaneously. It is concluded that relaxing the slamming constraint has no effect on the optimal solution. Deck wetness, in contrast, is dominant. In all cases, violation of the slamming constraint occurred together with violation of the deck wetness constraint, but the converse was not true. Altering the limiting value for the probability of occurrence of this event (i.e. green water on deck) changes the structure of the solution significantly. This suggests that the type and limiting values of the constraints must be selected carefully and consistently with the risks that particular vessels and their cargoes encounter in severe sea states. The observed redundancy of the slamming constraint could justify enforcement of a different restriction, such as rolling angle or acceleration. Some caution should be exercised regarding the accuracy of computation of ship motions, if they are to be incorporated as constraints in the program. For example, prescribing a rolling angle (RMS) restriction implies computation of rolling RAOs. If these computations are performed on the basis of ideal flow principles the results may not be realistic, as roll is primarily a viscous phenomenon.

Fuel Savings

There is no simple way to unambiguously quantify the savings in fuel cost in this constrained optimization simulation. The lack of knowledge, in advance, of the navigation track the captain would follow to avoid rough seas makes it difficult to directly compare it to the optimal solution and evaluate its beneficial effect. An estimate may however be obtained by examining feasible routes that satisfy the constraints and comparing them to the optimal.

Any route which brings the ship from the point of departure to the destination point without violating the imposed constraints as well as the voyage final time belongs to the set of feasible routes. The one yielding the lowest value of the assigned performance index (fuel cost, in

ISSN:1791-4469

particular) is the optimal route. Given the difference in the forecasted severity of the sea between loxodrome and orthodrome, as discussed earlier in this section, existence of a plausible track was investigated to the south of the rhumb line route. This baseline-feasible, sub-optimal route (figure 23) satisfies the prescribed constraints. It was determined by trial and error. Calculation yields a fuel cost of 85.93 tonne for this route. The optimal route of figure 19 requires 5.7% less bunker fuel than the baseline feasible one. Table 5 summarizes the calculated savings. Although the calculated saving rates are heavily dependent on the predicted sea states of the particular example, the benefit of utilizing a decision support system for route selection over sheer intuition or mariner's experience is indisputable.

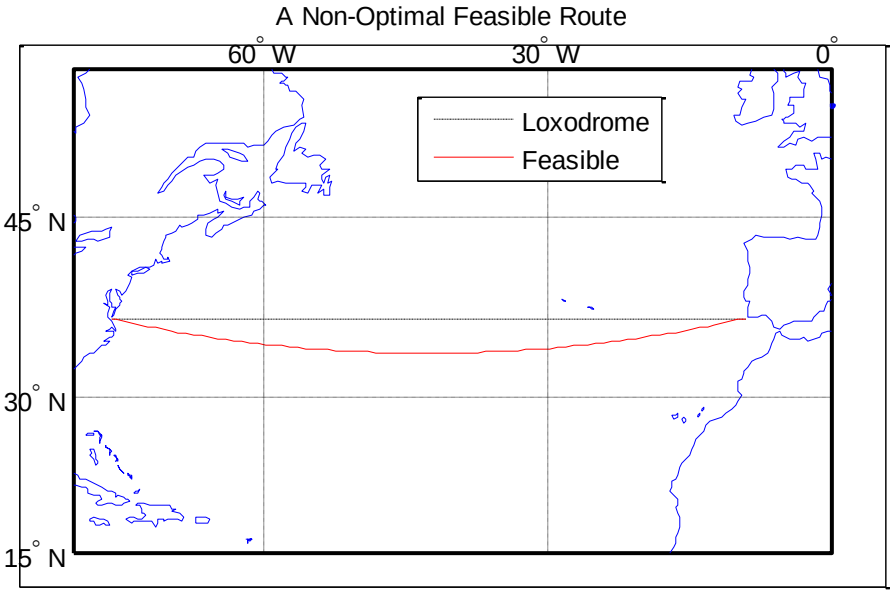


FIGURE. 23. A Baseline Feasible Route South of the Loxodrome.

TABLE 5. Summary of Fuel Savings.

Description	Value
Fuel Cost of Baseline Feasible Route (tonne)	85.93
Fuel Cost of Nominal Route (tonne)	97.31
Fuel Cost of Optimal Route (tonne)	81.06
Fuel Savings Relative to Nominal (%)	**16.7**
Fuel Savings Relative to Baseline Feasible (%)	**5.7**

CONCLUSIONS

Fuel-efficient navigation is explored and a solution to the minimum cost routing problem is proposed. An understanding of the governing physics of the problem in question has been emphasized. The dynamics of ship motions and sea state evolution is best captured by state-of-the-art tools like SWAN1 and WAM4. Particular attention should be paid to the added resistance in waves, as it is a driving factor in the vessel routing problem. A comprehensive model for the

optimal control of a ship in waves is developed and a solution has been shown to exist and has been generated using Dynamic Programming. This technique proves to be particularly appealing as it circumvents problematic questions of existence and uniqueness of solution, optimality conditions, existence of derivatives and gradients, all basic requirements of the classical calculus of variations. Furthermore, it is ideally suitable for fast, real time implementation onboard the vessel. Iterative Dynamic Programming, an intelligent alternative to the standard DP algorithm, is found to produce fast results with reasonable accuracy. The main advantage of this alternative, pointed out by Luus (2000) who introduced it, lies in its robustness in obtaining the global optimum. Although there is always a possibility of obtaining a local optimum, careful selection of the IDP parameters usually eliminates this risk and it is proven to give remarkable results in challenging problems, compared to existing well-established methods such as sequential quadratic programming and others. The structure of the optimal solution is strongly affected by the enforced constraints. The findings reveal a great potential in the application of optimal routing methods in shipping. The example-ship used in this paper is a small cargo liner. The magnitude of the derived savings projects to a notable reduction of bunker fuel expenses, especially for tankers, bulk carriers and containerships.

Accurate weather forecasts need to be the subject of a continuous research effort for the results of the present fuel efficient routing algorithm to be significant and useful. The current formulation carries out the optimization of the expected value of the cost function conditioned on a given weather forecast known deterministically. Current weather forecasting technologies allow for the accurate prediction of the expected sea states days in advance. This enables the direct implementation of the optimal ship routing method developed in the present article for the reduction of fuel consumption for trips of the order of weeks by breaking the sailing time into sub-intervals over which the weather forecast is known with a high degree of certainty.

The shipping industry consumes approximately 4 million barrels of oil daily. Assuming that cargo vessels sail into severe sea states approximately 25% of their sailing time, the implementation of the fuel efficient routing algorithm developed in the present study, even using a modest 5% fuel saving rate, would result approximately into a daily reduction of fuel consumption of approximately 50,000 barrels of oil (=0.25x0.05x4,000,000). Other than the cost savings by the shipping industry this reduction of crude oil consumption results in a reduction of emissions of CO_2 and other greenhouse gases into the atmosphere. The combined liquid fuels obtained from the refining of an average barrel of oil will produce a minimum of 317kg of CO_2 when consumed. Therefore a reduction of fuel consumption by the shipping industry of 50,000 barrels of oil daily would result in a reduction of 15,850 tonnes of CO_2 emissions. These figures may climb higher under conditions that result in greater fuel savings rates such as the 26% estimated by Tsujimoto and Tanizawa (2006), or the 22.1% estimated by Mezaoui et al. (2009) etc.

The optimal routing methods developed in the present study are readily applicable to the optimal sailing of yachts at the America's Cup level and in open ocean racing. The cost function that must be minimized when sailing a yacht is the time to destination. Weather uncertainty arises from the wind and wave forecasts and accurate forecasts are currently available a day or more in advance. The resistance of sailing yachts in calm water and in a sea state is analogous to that of ships, augmented by the induced drag caused by the wake of the keel, winglets, rudder and sails. The "propulsion" of a sailing yacht arises from the lift force exerted by the wind on the sails. The controls to be optimally selected by the dynamic programming algorithm are the settings of the rudder and trim tabs and the trimming of the sails. Given the advanced state of development of the hydrodynamic and aerodynamic performance attributes of hull forms, appendages and sails the introduction of optimal navigation strategies based on real time dynamic programming methods stands to offer a significant edge in competitive yacht sailing.

ISSN:1791-4469

A Look Ahead

The model presented in this paper belongs to the broader category of white-box models, namely models fully transparent in the incorporation of the governing physics to the best possible extent. On the other extreme, black-box models exist that are oblivious of the governing physics but rely on real time on-board measurements of relevant parameters to derive the quantities of interest using system identification techniques. The abundance of existing literature makes use of either approach, or combination of both (the so-called "grey-box" models) and claims fuel savings ranging from 2% to over 25%. The savings anticipated by deterministic optimal routing models are unlikely to materialize with the degree of confidence expected by ship owners and operators. The cornerstone of weather routing is wave forecasting. As already mentioned above, the state of the art in wave modeling and forecasting is deterministic, yet the nature of predicting the future sea state is by definition stochastic. Furthermore, variation of a range of parameters would lead to poor prediction of fuel consumption even under the hypothesis of perfect knowledge of future weather. Just to name a few, such parameters could be hull/propeller fouling, engine performance degradation, propeller performance in waves, loading condition, water density, presence of unpredicted currents, bunker fuel heating value and so on. A stochastic optimal weather routing formulation would seek to determine the optimal control strategy which remains optimal under all state variables and uncertainties associated with the problem. This task is challenging in multi-dimensional and stochastic settings and is being addressed as outlined below.

Recent advances in control theory allow the explicit solution of the multi-dimensional stochastic optimal control problem by casting the ship evolution equations in matrix state-space form. In the context of fuel-efficient ship routing the states include: the ship fuel consumption, the propeller RPM, the vessel coordinates relative to an earth-fixed frame, the vessel horizontal velocities, the vessel yaw angle, the vessel yaw rate, the rudder angle, the rudder angular velocity and the parameters discussed in the preceding paragraph. Casting the coupled resistance, maneuvering and seakeeping problems in a multi-dimensional state-space form and allowing for the weather forecasts to be known only stochastically leads to an analytical and efficient determination of the ship route with the minimum fuel consumption using the methods described in Yong and Zhou (1999). The generality and efficiency of this optimal control algorithm would allow its real-time implementation on a ship using a desktop PC. Its implementation to the fuel-efficient ship routing problem is very promising and is the subject of ongoing research.

ACKNOWLEDGMENTS

The contribution of Professor George Kallos from the University of Athens and Professor George Galanis from the Hellenic Naval Academy, Greece, is deeply acknowledged. They provided the output of the WAM4 numerical model for wave and swell forecasting.

REFERENCES

1. Abramowski, Przemyslaw, Tomasz Abramowski, and Zenon Zwierzewicz. 2006. "Formal Solution of Ship Weather Routing Problem via Pontryagin's Maximum Principle". *Síntesis Tecnológica*. 3 (1): 27-31.
2. Avgouleas, K. 2008. "Optimal Ship Routing". SM Thesis, Massachusetts Institute of Technology.
3. Bellman, R.E. 1957. *Dynamic Programming*. Princeton, N.J.: Princeton University Press.
4. Böttner, C. U. 2007. "Weather Routing for Ships in Degraded Condition". *International Symposium of Maritime Safety, Security and Environmental Protection*, Athens, Greece.

5. Bruns, A., K. Christiansen, and D. Rossow. 2011. “FSG.EcoPilot – an Onboard Tool for Fuel Efficient Speed Profile Selection”. In *10th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT'11)*, Berlin, Germany.
6. Delitala, A. M. S., Gallino, S., Villa, L., Lagouvardos, K.,Drago, A. 2010. “Weather Routing in Long-Distance Mediterranean Routes”. *Theoretical and Applied Climatology* 102 (1-2): 1-2.
7. Dolinskaya, I. S., M. Kotinis, M. G. Parsons, and R. L. Smith. 2009. “Optimal Short-Range Routing of Vessels in a Seaway”. *Journal of Ship Research*, Society of Naval Architects and Marine Engineers 53 (3): 121-9.
8. Dolinskaya, Irina S. 2012. “Optimal Path Finding in Direction, Location, and Time Dependent Environments”. *Naval Research Logistics* (NRL) 59 (5): 325-39.
9. Emmanouil, G., G. Galanis, and G. Kallos. 2007. “Assimilation of Radar Altimeter Data in Numerical Wave Models: an Impact Study in Two Different Wave Climate Regions”. *Annales Geophysicae*, 25(3): p. 581-596.
10. Faltinsen, O.M., K. J. Minsaas, N. Liapis, and S. O. Skørdal. 1980. “Prediction of Resistance and Propulsion of a Ship in a Seaway”. In *13th Symposioum on Naval Hydrodynamics*. T. Inui. Editor. The Shipbuilding Research Association of Japan: Tokyo. p. 505-530.
11. Faltinsen, O.M. 1990. *Sea Loads on Ships and Offshore Structures*. Cambridge Ocean Technology Series, 1, Cambridge; New York: Cambridge University Press.
12. Gerritsma, J., and W. Beukelman. 1972. “Analysis of the Resistance in Waves of a Fast Cargo Ship”. *International Shipbuilding Progress*, 19(217): p. 285-293.
13. Gershanik V.I. 2011. “Weather Routing Optimisation - Challenges and Rewards. Proceedings of the Institute of Marine Engineering, Science and Technology Part A: *Journal of Marine Engineering and Technology*. 10 (3): 29-40.
14. Hinnenthal, J., and G. Clauss,. 2010. “Robust Pareto-Optimum Routing of Ships Utilising Deterministic and Ensemble Weather Forecasts”. *Ships and Offshore Structures* 5 (2): 105-14.
15. Hotrop, J. 1984. “A Statistical Re-Analysis of Resistance and Propulsion Data”. *International Shipbuilding Progress*, 31(363): p. 272-276.
16. Ilus, T., and A. Heikkinen. 2012. “Challenges in Vessel Speed Optimization”. In *11th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT'12)*, Liege, Belgium.
17. Kirk, D.E. 1970. *Optimal Control Theory: an Introduction.* Englewood Cliffs, N.J.: Prentice-Hall.
18. Kobayashi, E., T. Asajima, and N. Sueyoshi. 2011. “Advanced Navigation Route Optimization for an Oceangoing Vessel”. In *Methods and Algorithms in Navigation – Marine Navigation and Safety of Sea Transportation*, eds. Adam Weintrit, Tomasz Neumann, 149-155CRC Press.
19. Kosmas, O. T., and D. S. Vlachos. 2012. “Simulated Annealing for Optimal Ship Routing”. *Computers & Operations Research* 39 (3) (3): 576-81.
20. Krata, P., and J. Szlapczynska. 2011. “Weather Hazard Avoidance in Modeling Safety of Motor-Driven Ship for Multicriteria Weather Routing”. In *Methods and Algorithms in Navigation – Marine Navigation and Safety of Sea Transportation*, eds. Adam Weintrit, Tomasz Neumann, 165-172CRC Press.
21. Luus, R. 2000. *Iterative Dynamic Programming*. Boca Raton: Chapman & Hall/CRC.
22. MAK 9M25C Project Guide –Propulsion, Issue July 2008, Caterpillar Motoren GmbH & Co. KG, (downloaded from www.stet.pt).
23. Maki, Atsuo, Youhei Akimoto, Yuichi Nagata, Shigenobu Kobayashi, Eiichi Kobayashi, Shigeaki Shiotani, Teruo Ohsawa, and Naoya Umeda. 2011. “A New Weather-Routing System that Accounts for Ship Stability Based on a Real-Coded Genetic Algorithm”. *Journal of Marine Science and Technology* 16, (3) (09): 311-322.
24. Marie, S., and E. Courteille. 2009. “Multi-Objective Optimization of Motor Vessel Route”. In *Marine Navigation and Safety of Sea Transportation*, ed. Adam Weintrit, 411-418CRC Press.
25. Marie, S., and E. Courteille. 2014. “Sail-Assisted Motor Vessels Weather Routing Using a Fuzzy Logic Model”. *Journal of Marine Science and Technology* 19: 265-279.
26. Mezaoui, B., K. Takashima, and R. Shoji. 2009. “On the Fuel Saving Operation for Coastal Merchant Ships Using Weather Routing”. In *Marine Navigation and Safety of Sea Transportation*, ed. Adam Weintrit, 431-436CRC Press.
27. Montes, A. 2005. “Network Shortest Path Application for Optimum Track Ship Routing”. MSc thesis, Naval Postgraduate School.
28. Ochi, M.K. 1998. *Ocean Waves: the Stochastic Approach*. Cambridge Ocean Technology Series, 6, Cambridge, U.K.; New York: Cambridge University Press.

ISSN:1791-4469

29. Sen D. and Padhy C.P. 2009. “Development of a Ship Weather Routing Algorithm and its Application for Indian Coastal Sea Routes”. *International Journal of Ecology and Development* 12 (W09): 59-74.
30. Panigrahi, J. K., Padhy, C. P., Sen, D., Swain, J.,Larsen, O. 2012. “Optimal Ship Tracking on a Navigation Route between Two Ports: A Hydrodynamics Approach”. *Journal of Marine Science and Technology* 17 (1): 59-67.
31. Panigrahi, J. K., Tripathy, J. K., Umesh,P.A.,. 2008. “Optimum Tracking of Ship Routes in 3g-WAM Simulated Rough Weather Using IRS-P4 (MSMR) Analysed Wind Fields”. *Journal of the Indian Society of Remote Sensing* 36 (2): 149-58.
32. Pipchenko, O. 2011. “On the Method of Ship's Transoceanic Route Planning”. In *Methods and Algorithms in Navigation – Marine Navigation and Safety of Sea Transportation.* eds. Adam Weintrit, Tomasz Neumann, 157-163CRC Press.
33. Rathe, H., and C. Beiersdorf. 2005. “Decision Support for Container Ship Operation in Heavy Seas – Shipboard Routing Assistance” In *4th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT'05)*, Hamburg, Germany.
34. Sclavounos, P.D., and D.E.Nakos. 1993. “Seakeeping and Added Resistance of I.A.C.C. Yachts by Three-Dimensional Rankine Panel Methods”. In *11th Chesapeake Sailing Yacht Symposium.* Annapolis MD.
35. Shao W., Zhou P.,Thong S.K. 2012. “Development of a Novel Forward Dynamic Programming Method for Weather Routing”. *Journal of Marine Science and Technology (Japan)* 17 (2): 239-51.
36. Skoglund, L., J. Kuttenkeuler, and A. Rosén. 2012. “A New Method for Robust Route Optimization in Ensemble Weather Forecasts”. Trita-AVE / KTH Aeronautical and Vehicle Engineering 2012:67 .
37. Szłapczyńska, J. 2007. “Multiobjective Approach to Weather Routing”. *Advances in Marine Navigation and Safety of Sea Transportation.* Proceedings of TransNav 2007, Gdynia 2007. Gdynia Maritime University.
38. Szłapczyńska, J. 2013. “Multicriteria Evolutionary Weather Routing Algorithm in Practice”. *International Journal on Marine Navigation and Safety of Sea Transportation*, TransNav 7 (1): 61-5.
39. Szlapczyńska, J., and R. Smierzchalski. 2007. “Adapted Isochrone Method Improving Ship Safety in Weather Routing with Evolutionary Approach”. *International Journal of Reliability, Quality and Safety Engineering.* 14 (06): 635-645.
40. Szlapczyńska, J., and R. Smierzchalski. 2009. “Multicriteria Optimisation in Weather Routing”. In *Marine Navigation and Safety of Sea Transportation*, ed. Adam Weintrit, 423-429CRC Press.
41. Todd, F.H. 1963. “Series 60 Methodical Experiments with Models of Single-Screw Merchant Ships”. Washington: For sale by the Supt. of Docs., U.S. Govt. Print. Off.
42. Tsou, Ming-Cheng and Hung-Chih Cheng. 2013. “An Ant Colony Algorithm for Efficient Ship Routing”. *Polish Maritime Research*. 20(3): 3-50
43. Tsujimoto, M., and K. Tanizawa. 2006. “Development of a Weather Adaptive Navigation System Considering Ship Performance”. In *Actual Seas, 25th International Conference on Offshore Mechanics and Arctic Engineering*, Proceedings of OMAE 2006, Germany.
44. Wisniewski, B., J. Chomski, and P. Medyna. 2009. “Application of the 1-2-3 Rule for Calculations of a Vessel’s Route Using Evolutionary Algorithms”. In *Marine Navigation and Safety of Sea Transportation*, ed. Adam Weintrit, 419-422CRC Press.
45. Yong, J. and X. Z. Zhou. 1999. *Stochastic Controls: Hamiltonian Systems and HJB Equations.* Applications of Mathematics Series, 43, Springer-Verlag, New York.