REPORT No. 142 S
(S 2/171)

August 1970

NEDERLANDS SCHEEPSSTUDIECENTRUM TNO
NETHERLANDS SHIP RESEARCH CENTRE TNO
SHIPBUILDING DEPARTMENT LEEGHWATERSTRAAT 5, DELFT

★

OPTIMAL METEOROLOGICAL SHIP ROUTEING

(OPTIMAAL METEOROLOGISCH ROUTEREN VAN SCHEPEN)

by

DR. C. DE WIT
Department of General Sciences, Delft University of Technology

Issued by the Council

VOORWOORD

Omdat de snelheid van bepaalde scheepstypen in de laatste jaren aanzienlijk is toegenomen, wordt thans een verbeterd inzicht gevraagd in die factoren die van belang zijn voor het behouden van de snelheid.

Indien immers voor een zekere vorm van zeetransport een gewenste snelheid is vastgesteld, dan is het, juist indien deze hoog is, van essentieel belang dat deze „geïnstalleerde" snelheid zo goed mogelijk wordt benut. Hiermede wordt dus bereikt dat ook de reisduur optimaal kan zijn.

De hoofdoorzaak van verstoringen die resulteren in een toename van de reisduur is het gedrag van het schip onder invloed van de externe omstandigheden zee en wind.

Indien voor een type schip de vaartkarakteristieken bekend zijn, d.w.z. het responsiegedrag qua snelheid en bewegingen in gegeven zeetoestanden, is voor een bepaalde reis met optimaal routeren na te gaan, welk traject kan leiden tot de kortste reisduur.

Toen dan ook naar aanleiding van het proefschrift van Dr. C. de Wit: „Mathematical treatment of optimal ocean ship routeing" goede informatie beschikbaar kwam om nader op van belang zijnde problemen in te gaan, is hem gevraagd de resultaten van het proefschrift voor een rapport te bewerken.

In principe gaat de gegeven methode uit enerzijds van de reeds genoemde vaartkarakteristieken inclusief de begrenzingen door extreme verschijnselen en/of wensen van de gezagvoerder, anderzijds van golf-, stroom- en windverwachtingen van meteorologische stations voor het te bevaren gebied met begin- en aankomsthaven. Met een computerprogramma worden op basis van deze gegevens voor halve dagen tijdfronten berekend, waarmede een plot van het kortstdurende traject mogelijk wordt. In verband met de relatieve betrouwbaarheid van golf- en windvoorspellingen voor langere duur zullen de berekeningen van het kortstdurende traject regelmatig, bijvoorbeeld elke dag, met de bereikte positie als uitgangspunt herhaald moeten worden.

Het spreekt vanzelf dat waar de gezagvoerder uiteindelijk de verantwoordelijkheid voor het schip heeft, een dergelijke routeerbegeleiding slechts als een advies kan dienen.

Vooral voor de hogere snelheden mag verwacht worden dat een routeermethode in tijdwinst zal resulteren met daarnaast enige bijkomende voordelen zoals minder stormschade en meer zekerheid betreffende de te verwachten tijd van aankomst.

HET NEDERLANDS SCHEEPSSTUDIECENTRUM TNO

PREFACE

Because of the important increase in speeds for some types of ships, a better insight will be required in those factors that have an influence on the sustained speed.

If for a certain form of seatransportation the required speed has been decided and especially if this is high, it will be of the utmost importance that the "installed" speed can be made as good as possible in order to obtain optimal voyage time.

The main cause of disturbances resulting in an increase in voyage time is the behaviour of the ship as influenced by the external conditions waves and wind.

Where for a certain type of ship the sailing characteristics are known i.e. the response behaviour for speed and motions in given seastates it will be possible to construct the least time track with the aid of optimal routeing.

When with Dr. de Wit's thesis "Mathematical treatment of optimal ocean ship routeing" valuable information came available to consider important problems involved, the author has been requested to prepare the subject report.

In principle the method starts firstly with the sailing characteristics mentioned including limitations caused by extreme dynamic phenomena and/or the master's requirements, and secondly with the availability of reliable wave, current and wind forecasts for the sailing region including starting and destination point. With the aid of a computer program the timefronts for 12 hour sailing can be computed of which a plot for the least time track can be constructed. Because of the relative reliability of waves and wind forecasts for longer periods, the computations for the least time track should be regularly repeated, for instance each day, with the position reached as a new starting point.

It will be self-evident that because of the master's responsibility for the ship, the routeing should only serve as an advice.

Especially for the higher ship speeds it may be expected that routeing may result in a time gain, besides some other gains such as less storm damage and a better predictability of the expected time of arrival.

THE NETHERLANDS SHIP RESEARCH CENTRE TNO

CONTENTS

LIST OF SYMBOLS

A	Starting point (fig. 1a)
A_g	Sea wave amplitude
A_s	Amplitude of ship motion
B	Point of destination (fig. 1a)
EQ	Equator (fig. 2a)
F	Wind fetch (fig. 4.1a)
G	Navigating region (fig. 1a) (also: ship's centre of gravity)
H	Significant wave height (fig. 4.1a)
$H^+(t)$	Region of all points that can be reached at time t, starting from a certain point (fig. 3.2a)
L	Geographical longitude of P (λ transformed to sphere) (fig. 2a)
N_0	East-West curvature radius in the centre of G
O	Origin of system of axes
P	Arbitrary point of G (ship's position) (fig. 2a)
P_n	North Pole (fig. 2a)
R	Radius of sphere
S	Intersection of old and new equator (fig. 2a)
$S^+(t)$	Boundary of $H^+(t)$, "time front" (fig. 3.2a)
SK	Skew equator
T	Significant wave period (fig. 4.1a)
U	Wind velocity (fig. 4.1a)
X, Y	System of axes
a	Timefront gradient direction (fig. 5.2a)
c	Sea current velocity
d	Distance (fig. 5.1b)
$\mathrm{d}s_a$	Length of a line element on earth
$\mathrm{d}s_k$	Length of the chart mapping of a line element on earth
f	Ship's velocity
k	Scale factor
r	Distance from O (fig. 3.5a)
s	Ship's velocity with respect to the water
t	Time (in fig. 4.1a wind duration)
t_a	Starting time
t_b	Time of arrival
v	Velocity
v_0	Velocity in O (fig. 3.5a)
v_r	Velocity at distance r from O (fig. 3.5a)
x, y	Co-ordinates on XY-system
Γ	Extremal (fig. 3.2b)
α	Ship's course (fig. 1a)
γ	Skew longitude of P (fig. 2a)
δ	Skew latitude of P (fig. 2a)
ϑ	Wind direction or mean wave direction
λ	Geographical longitude of P (on ellipsoid)
ϱ_0	North-South curvature radius in the centre of G
τ	Time
φ	Geographical latitude of P (on ellipsoid)
φ_0	Initial direction (fig. 3.4a)
ψ	Geographical latitude of P (transformed to sphere) (fig. 2a)
$\underline{\psi}$	Gradient of $S^+(t)$ in P (fig. 3.3a)
note 1	Symbols $A, B, C, D, M, N, Q, R, S, T$ are also used to indicate points in various figures
note 2	(t) indicates a function of time e.g. $\alpha(t)$
note 3	Vectors are underlined e.g. $\underline{c}$, $\underline{f}$

OPTIMAL METEOROLOGICAL SHIP ROUTEING

by

Dr. C. DE WIT

Summary

The author of this report was assigned to investigate the mathematical aspects of meteorological ship routeing by Prof. Dr. R. Timman of the Delft University of Technology. The assignment was accomplished in 1968, resulting in a doctor's thesis, entitled "Mathematical Treatment of Optimal Ocean Ship Routeing" [1]. This report describes the outlines of this dissertation.
The first section contains the problem statement as well as some notational conventions.
In order to avoid metric difficulties, the navigated region G is to be mapped conformally onto a plane, keeping the scale alteration as small as possible. Section 2 gives a curse treatment of the most important mappings that fit this purpose.
In the 3rd section the minimal time problem is discussed as an application of Pontryagin's optimal control theory. The concept of a timefront is introduced and some attention is given to possible structure complications of these fronts. Also modifications of the general theory, in the event that part of a trajectory should coincide with the boundary of the region G, are briefly mentioned.
Section 4 cursely treats wave prediction methods and the determination of a ship's behaviour in sea waves.
In the 5th section the computer program to evaluate the least time track is described and the results of a practical application are exhibited. The consequences of data incertainties and the practical use of the treated method are briefly discussed.
The two appendices contain mathematical treatments of the general minimal time problem and of the minimal time problem with co-ordinate restrictions.

1 Introduction

The following data are assumed to be known: (see figure 1a)

1. A region G, part of the Atlantic or Pacific Ocean, restricted by the need of a depth of at least 40 fathoms. In this region, the starting point A and the destination B are given. as well as the starting time t_a.
2. The maximum ship's speed, taking normal safety considerations into account, in all points of G and at all times $t \geqslant t_a$. This assumption implies, that one has full quantitative knowledge of
   a. the direction and rate of the sea current, assumed stationary for a time stretch, exceeding the normal trip's duration by some 30%;
   b. the sea conditions, specified by the significant wave height and direction at all times $t \geqslant t_a$ and
   c. the ship's performance in sea waves – wind waves or swell – also possibly determined by wishes of the ship's master.

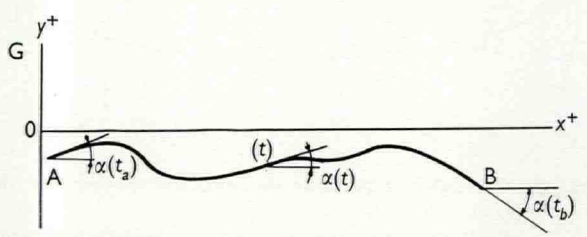


Fig. 1a. G: navigating region; A: starting point; B: destination; $\alpha(t)$: course α as a function of the time t; t_a: starting time; t_b: arrival time.

With these available data, one can attempt to determine the following quantities:

1. The ship's course α as a function of time – $\alpha = \alpha(t)$ – so that the ship, starting from A at a time t_a, arrives at B at a time t_b. This arrival time t_b is obviously later than t_a. Furthermore it can be stated, that t_b depends on the choice of the navigating function $\alpha(t)$. In view of the fact, that ships have practically always succeeded to reach B after a finite time, it is sensible to assume, that there is a collection of navigating functions $\{\alpha(t)\}$, meeting the needs that were mentioned before.
2. The optimal navigating function $\alpha^0(t)$, i.e. the navigating function that generates a minimal arrival time t_b^0.

Although this problem statement may be mathematically correct, it looks rather unflexible from a practical seaman's viewpoint. The evaluation of the least time track, also indicated as the optimal track, can only be executed with the aid of a large memory computer. Such a machine would be quite unfit to be used efficiently on board a merchant ship. Thus the optimal track would have to be evaluated at some shore computing centre.

The trouble now is, that the ship's master is primarily responsible for the ship's safety, so that he can merely be advised to take a certain route. Such a routeing advice should not be given without a rather extensive explanation, based on weather and wave predictions.

Working that way and revising the routeing advices day by day, the possibility of a discrepancy between

expected and actual ship's performance is not just academic. This means, that ship routeing has to be accompanied by an extensive mutual exchange of information between ship and shore station. Only when collaborating with good mutual respect and understanding, a significant result can be expected from ship routeing.

As for the mathematical notations in this report, the following conventions were adopted.

Whenever a scalar, like the time t or the angle φ, or a vector, like the position vector $\underline{x}$, needs to be marked by means of some index or letter, this mark is always put right below, like t_a, φ_0, $\underline{x}_b$. Vectors are denoted by underlined letters, like $\underline{x}, \underline{f}, \underline{\psi}$.

As for vector components, the indices will be placed in the right top, if the vector is a "column vector" in matrix calculus or "contravariant" in tensor calculus: $\underline{x} = (x^1, x^2)$.

If we are dealing with a row vector (in matrix calculus) or a covariant vector (in tensor calculus), the component indices are placed below: $\underline{\psi} = (\psi_1, \psi_2)$.

The notation $p := q$ stands for the assignment to give the variable p the value of the expression q. For example, after the sequence $a := 0$; $a := a+1$; $a := a+\cos(a)$; the variable a has the value $1+\cos(1) = 1.55\ldots$.

A frequently used notation in this report is of the type

$$x := a(b)c$$

This means, that x is consecutively given the values a, $a+b$, $a+2b$ etc. until either x exceeds c if b – the step width – is positive, or until x becomes less than c in the event that b is negative.

2 Co-ordinate system

All evaluations regarding positions, directions and distances to be made good could be carried out on the geoid itself, sufficiently approximated by the Hayford ellipsoid. However, since the curvature of the earth is dependent of the position as well as the direction, while the meridian convergence depends on the latitude, a direct use of the ellipsoid co-ordinates φ (geographical latitude) and λ (longitude) would imply needlessly elaborate computational techniques.

A first simplification can be obtained by making a conformal mapping of the ellipsoid onto a sphere with radius $R = \sqrt{(\varrho_0 N_0)}$, where ϱ_0 and N_0 are the North-South and East-West curvature radii in the centre of G respectively.

This way the spherical mapping no longer has the disadvantage of a variable curvature, while the distance scale is practically equal to unity over a wide latitude range. The co-ordinates (φ, λ) on the ellipsoid are thus transformed to co-ordinates (ψ, L) on the sphere [2, 3].

The disadvantage of meridian convergence can only be removed by mapping a part of the sphere conformally onto a flat plane.

The Royal Dutch Weather Bureau, to be indicated from here on by the abbreviation K.N.M.I. (Koninklijk Nederlands Meteorologisch Instituut), adopted a polar stereographic projection of the earth's surface for all kinds of weather maps, so that it was quite natural to use this projection for ship routeing as well. The length of a line element on earth

$$ds_a = \sqrt{(R^2 d\psi^2 + r^2 dL^2)}$$

now has a position-dependent proportion to the length of the mapping of this element

$$ds_k = \sqrt{(dx^2 + dy^2)},$$

that is $ds_k = k \times ds_a$, where k is the scale in a point with co-ordinates (x, y) In a polar stereographic mapping this scale k depends on the spherical latitude by $k = k_0 \sec^2(\pi/4 - \psi/2)$ [3].

Between 30° and 50° Northern latitudes this scale alters about 17%. As a result of this rather considerable scale alteration the mapping of a spherical great circle, which can be taken to be the spherical mapping of a geodesic line on the ellipsoid surface, becomes a curved line with a curvature, that can not be neglected for practical purposes. This means that a polar stereographic projection does not give a quite "natural" picture of the situation.

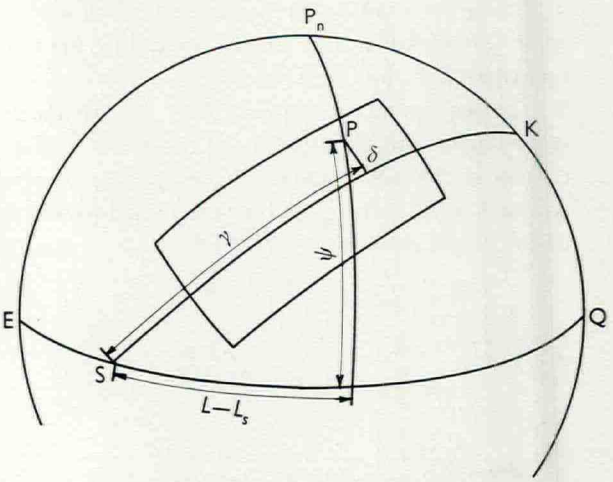


Fig. 2a. The sphere represents the spherical mapping of the earth ellipsoid.
P_n: North Pole; EQ: equator; ψ, L: latitude and longitude of P with respect to the equator and the Greenwich meridian; SK: skew equator; γ, δ: skew longtiude (relative to S) and skew latitude (relative to SK) of P.

A projection, that meets the requirements of being conformal as well as having a very slight scale alteration is the Kahn projection.

As a result of the practically constant scale a straight line segment between two points in this chart can be taken with good accuracy to be the mapping of the metrically shortest distance between these points.

This projection is obtained by enclosing the region G – which has a length/width ratio of about 4:1 - by four mutually perpendicular great circles (see figure 2a) [2].

Taking the central great circle about parallel to the longest frame lines, one can think of this great circle as an oblique equator. The intersection S of the old and the new equator can be adopted as reference point for the oblique longitude γ, while the great circle distance from a point P to the skew equator, measured along a skew meridian, is the new latitude δ. The new equator and meridians are now conformally mapped by means of the well known Mercator projection.

Taking the central great circle mapping as the X-axis and the mapping of an arbitrary skew meridian as the Y-axis, the mapping equations are

$$x = k_0 R(\gamma - \gamma_0),$$

$$y = k_0 R \ln(\operatorname{tg}\delta + \sec\delta)$$

and the scale factor becomes

$$k = k_0 \cosh(y) = k_0 \sec\delta.$$

This factor k has to be taken into account when making distance computations.

3 Mathematical theory of optimal ship routeing

3.1 *Indicatrix*

From an arbitrary point $P(x,y)$ of the region G and for an arbitrary time $t \geqslant t_a$ one can plot the ship's velocity vector for all possible – and admissible – values of the ship's course α. (See figure 3.1a). For this polar velocity curve the name "original velocity indicatrix" is adopted [1].

As a preliminary remark it should be stated that the ship's velocity, denoted by $\mathrm{d}\underline{x}/\mathrm{d}t$, is a vector valued function of the position co-ordinates x^1 and x^2, the course α and the time t:

$$\frac{\mathrm{d}\underline{x}}{\mathrm{d}t} = \underline{f}(x^1, x^2, \alpha, t),$$

or written componentwise

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = f^i(x^1, x^2, \alpha, t) \quad \text{for} \quad i = 1,2$$

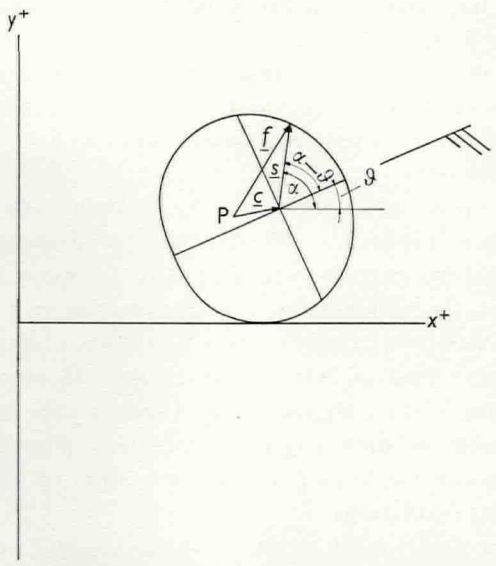


Fig. 3.1a. The original ship's velocity indicatrix for comparitively low waves, corresponding to a wind velocity 5 (Beaufort scale).
P: ship's position; $\underline{c}$: sea current; $\underline{s}$: ship's speed relative to water; α: ship's course; ϑ: wind direction; $\underline{f}$: ship's speed made good.

The vector $\underline{f}$ is composed of two vectors: $\underline{f} = \underline{c} + \underline{s}$.

Here $\underline{c}$ denotes the sea current, with a relatively small scalar value and approximately constant with respect to time. So $\underline{c} = (c^1, c^2)$ with c^i only dependent of the position co-ordinates x^1 and x^2.

The ship's propagation speed with respect to water $\underline{s}$ can be seen as a vector with a length s, depending on the significant wave height H and also varying with the ship's course α relative to the mean wave direction ϑ. As the quantities H and ϑ vary with position and time, s is implicitely dependent of the position, the time and the course α.

Mathematically stated: $\underline{s} = (s^1, s^2)$ with

$$s^1 = s(x^1, x^2, \alpha, t)\cos\alpha \quad \text{and} \quad s^2 = s(x^1, x^2, \alpha, t)\sin\alpha.$$

The components of $\underline{f}$ can now be written as

$$f^i = s(x^1, x^2, \alpha, t)\sin(i\pi/2 - \alpha) + c^i(x^1, x^2) \quad \text{for} \quad i = 1,2$$

We shall now discuss the various possibilities for the original velocity indicatrix.

In comparatively low waves the ship can proceed in all courses practically with the same speed, at least with maximal propeller speed.

The ship's speed s will be minimal for $\alpha = \vartheta$, i.e. for waves coming in from ahead. When the waves come in from one of the beams, so $\alpha = \vartheta \pm \pi/2$, s will be somewhat greater, while s is maximal for $\alpha = \vartheta + \pi$.

For other values of α the value of s can be taken so

that the velocity indicatrix becomes an ellipse (see figure 3.1a).

In this case it can be remarked that for the majority of merchant ships s is considerably greater than c. This implies that it is possible to make headway from P into all directions.

When the waves are higher, the problems arise. Both practical experience and shipbuilding laboratory investigations have revealed, that there is a sector of forbidden, i.e. non feasible courses, corresponding with sea waves coming in from about two to to six points (the compass card of 360° is divided into 32 points, so 1 point = $11\frac{1}{4}$°) aft the beam. These courses are undesirable, as they may cause unwanted phenomenae like too heavy rolling and an inclination to unstable yawing oscillations.

For these waveheights, a ship's original velocity indicatrix may become like shown in figure 3.1b.

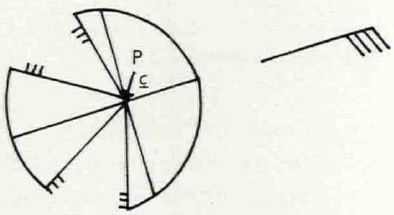


Fig. 3.1b. The original velocity indicatrix of an ocean ship in higher waves, wind speed 8 Beaufort.

Mathematically, this open or non-convex indicatrix means, that the classical calculus of variations, applicable in the case of a closed and convex indicatrix, can no longer be used. Fortunately, this variation calculus was extended essentially by Pontryagin. His theory concerns variation calculus with steering parameter restrictions and it can be applied to this particular case.

To make the application plausible, the concept of "δt-position-indicatrix" has to be introduced.

Take a time interval $\delta t > 0$ and small enough to be able to assume the ship's speed f to be a function of the central position x_0 and of the starting time t_0 within a circle with radius $2s_0\delta t$ around that centre and for times between t_0 and $t_0+2\delta t$ (s_0 is the unrestricted ship's speed). This has to express, that in a small enough subregion of the event space $\{x, t\}$ f can be assumed to be only dependent of the course α.

So $dx^i/dt = f^i(x_0^1, x_0^2, \alpha, t_0)$ in this small area.

The δt-position-indicatrix can now be constructed by plotting the vectors $f\delta t$ from P for all admissible values of α. We thus obtain an elliptic curve with interruptions.

Let M and N be the bordering points of such a restricted sector, corresponding with ship's speeds s_1 and s_2 (see figure 3.1c).

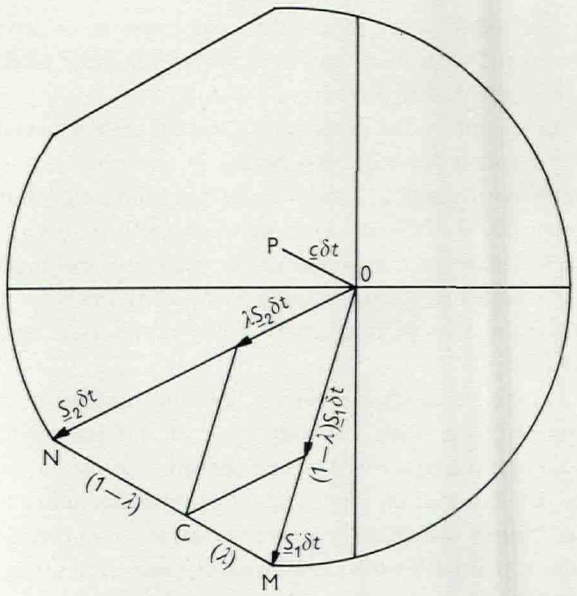


Fig. 3.1c. The effective δt-position indicatrix.
P: ship's position at time t_0; PO: current shift from time t_0 to time $t_0+\delta t$; $s_1\delta t$, $s_2\delta t$: boundary vectors of the forbidden course sector MON; C: point of the straight line segment MN; $CM/CN = \lambda/(1-\lambda)$.

It is then possible to reach an arbitrary point C of the straight line segment MN by tacking between the courses to reach M and N.

In figure 3.1c, let $MC/CN = \lambda/(1-\lambda)$. One can now reach C from P by taking the ship's speed s_2 from time t_0 to $t_0+\lambda\delta t$ and then tacking from s_2 to s_1.

The original δt-position-indicatrix can now be replaced by its convex closure, for which the name "effective δt-position-indicatrix" seems suitable. This effective indicatrix generally consists of two elliptically curved arcs, connected by two straight line segments. In this case it can as well be stated, that the stream vector c is always and everywhere small enough for P to be located inside its own effective indicatrix.

Thus we can base further considerations on the fact, that always and everywhere in G it is possible to make headway into all directions.

3.2 *Set of reachable points, timefronts, extremals*

For a given starting point A and starting time t_a, let us think of the set of all points, that can be reached at a given time $t > t_a$. This set of reachable points will be denoted by $H^+(t)$ [5]. It may be remarked, that every point of this set can generally be reached in more than one way, depending on the choice of $\alpha(\tau)$, with $t_a \leqslant \tau \leqslant t$.

The boundary of such a reachable region is defined as the timefront $S^+(t)$. This boundary curve can be

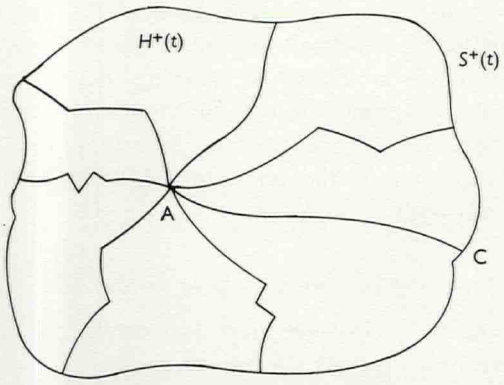


Fig. 3.2a. $H^+(t)$: region of reachable points at time t, starting from A at time t_a; $S^+(t)$: boundary of $H^+(t)$, called "timefront".

seen as the collection of extreme points, that can be ultimately reached at a time t (see figure 3.2a).

After these introductions, a few important implications have to be mentioned.

Lemma 3.2α

Let C be a point of timefront $S^+(t)$. Then the time t, at which C is reached from A, is minimal.

To prove this lemma, suppose there is a time $t' < t$, at which C is reachable from A. In that case C would be a point on the boundary or inside the set of reachable points $H^+(t')$.

Referring to the considerations in 3.1 about the effective position indicatrix it now follows, that $H^+(t)$ is everywhere wider than $H^+(t')$, because t is later than t'. So points of $H^+(t')$ become undoubtedly inner points of $H^+(t)$. Thus it is impossible for a point C of $H^+(t')$ to lie on the boundary of a larger region $H^+(t)$. The assumption, that a boundary point of $H^+(t)$ could be reached from A at an earlier time t' is thus leading to a contradiction and it cannot be true. This proves lemma 3.2α.

A trajectory, starting in A at a time t_a, that leads the ship to a boundary point of $S^+(t)$, will be called an extremal. A fundamental property of extremals is stated by the next lemma.

Lemma 3.2β

If the curve Γ_{AB} is an extremal from A to B, with starting time t_a and time of arrical t_b, then every subarc of Γ_{AB} is an extremal as well.

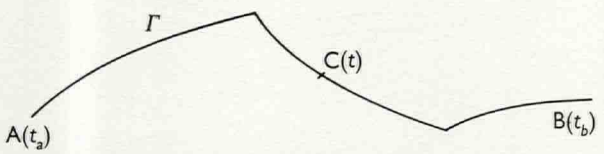


Fig. 3.2b. Γ: extremal from A to B; t_a: starting time; t_b: arrival time; $C(t)$: position on Γ at time t, with $t_a < t < t_b$.

For an explanation of this lemma, consider figure 3.2b.

The lemma says, that for an arbitrary point C of an extremal Γ_{AB}, passed at a time t, this point cannot be reached from A at an earlier time than t. It also states, that, if the ship starts at C at a time t, the subarc Γ_{CB} of Γ_{AB} is an extremal.

To illustrate the proof of these two statements, let us start with the last one. If Γ_{CB} is not an extremal, then B could be reached from C at an earlier arrival time t_b'. This contradicts the optimality of t_b as an arrival time from A at time t_a.

To prove the first part – Γ_{AC} is an extremal – one can introduce the concept of "set of initial points" $H^-(t)$, corresponding to the common arrival point B at the given arrival time t_b. The boundary of such a set $H^-(t)$, denoted by $S^-(t)$, has a property, which is analogous to lemma 3.2α:

The starting time t from a point of $S^-(t)$, the boundary of $H^-(t)$, is maximal to arrive in B at time t_b.

Working with this concept and assuming Γ_{AC} to be no extremal then leads once more to a contradiction regarding the extremality of Γ_{AB}.

In principle it is now possible to give a constructive solution of the optimal track problem (see figure 3.2c).

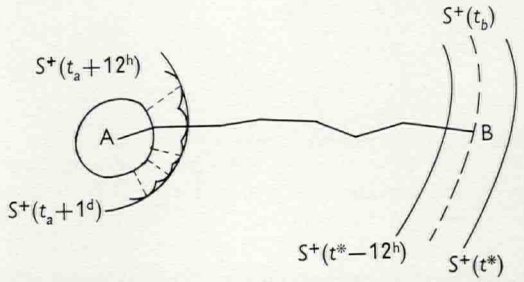


Fig. 3.2c. $S^+(t_a+12^h)$: Timefront for a time, 12 hours after the starting time t_a; $S^+(t_a+1^d)$: timefront for one day after the start; $t^* = t_a + k \times 12^h$, where k is a positive integer.

Taking 12^h time steps, one can construct the 12^h-position-indicatrix from A, which can be seen as a first guess of the timefront $S^+(t_a+12^h)$. The conditions at the boundary may be somewhat different from those in A on account of position as well as time differences, so one could apply corrections to these boundary points in accordance with these differences.

Accepting this corrected curve as the timefront $S^+(t_a+12^h)$, one could construct 12^h-position-indicatrices from the points of this first timefront and correct these on account of changes in time and position.

Taking these corrected indicatrices for granted, one can construct their envelope, which is the timefront $S^+(t)$ for $t = t_a + 1^d$. This procedure could be repeated, until the destination B becomes an inner or

possibly a bordering point of $H^+(t^*)$, while B lies outside $H^+(t^*-12^h)$. The optimal arrival time t_b could then be estimated by linear interpolation. The optimal track could be constructed by executing the same procedure, except for the fact, that we now have to "start" in B and "proceed" with decreasing time until A is reached.

The disadvantage of this procedure is, that it is rather elaborate and hardly fit to be programmed.

The enveloping procedure can be avoided by making a more effective use of lemma 3.2β, i.e. "an extremal goes from timefront to timefront".

3.3 *Pontryagin's Maximum Principle*

Starting in P on timefront $S^+(t)$, the lane between $S^+(t)$ and $S^+(t+\delta t)$, with δt positive but rather small, has to be covered in the fastest possible way. See figure 3.3a. Let $\underline{\psi}(t)$ be the gradient vector of $S^+(t)$ at P, then the velocity $\underline{f}$ in P has to be selected so that its projection on $\underline{\psi}(t)$ has a maximal length.

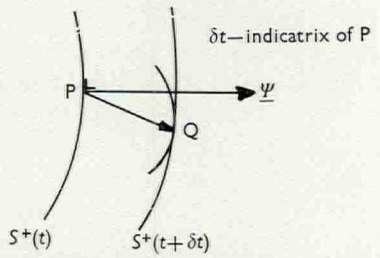


Fig. 3.3a. $S^+(t)$, $S^+(t+\delta t)$: timefronts for times t and $t+\delta t$, with $0^h < \delta t < 1^h$; $\overline{PQ} = \underline{f}t$: total ship's optimal displacement in the time interval $(t, t+\delta t)$; $\underline{\psi}$: gradient vector of $S^+(t)$ in P.

Therefore the optimal course $\alpha^0(t)$ has to be selected so that the innerproduct $(\underline{\psi}, \underline{f})$ is maximal [4].

This is Pontryagin's Maximum principle. The enveloping procedure can now be replaced by a much simpler construction, exhibited in figure 3.3b.

Given P on $S^+(t)$, draw the normal in P on $S^+(t)$ – i.e. the line perpendicular to the tangent to $S^+(t)$ in P – and construct the indicatrix of P. Take a line parallel to the timefront tangent in P and shift it until it "touches" the indicatrix, meaning that it has only one common point with the indicatrix, all other points lying on the "P-side" of that touching line. This touching point is the point of $S^+(t+\delta t)$ that corresponds with P on $S^+(t)$.

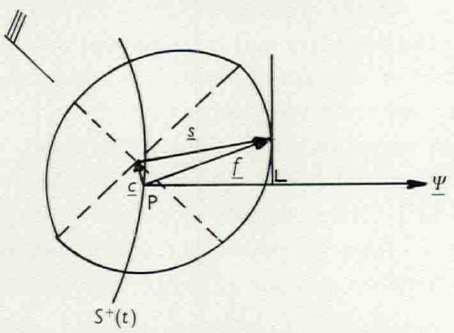


Fig. 3.3b. $S^+(t)$: timefront for time t; $\underline{\psi}$: timefront's gradient in P; $\underline{c}$: sea current vector; $\underline{s}$: ship's speed vector relative to water; $\underline{f}$: total ship's speed.
The elliptic curve is the ship's effective velocity indicatrix.

3.4 *Principles to construct the optimal track*

Revising the procedure, mentioned at the end of 3.2, we can now describe the construction of a timefront $S^+(t+\delta t)$ from a (supposed to be) given timefront $S^+(t)$:

> Take a number of points of $S^+(t)$ with sufficiently small subsequent distances. Determine the timefront gradient directions in these points and take optimal δt-steps, applying the maximum principle. Correct these steps on account of changes in time and position.
> Then one has a collection of points, through which a smooth curve can be laid, that is to represent the new timefront $S^+(t+\delta t)$.

This procedure has the disadvantage of being rather elaborate. It so seems sensible to investigate the possibility of constructing just a few extremals not timefrontwise, but one by one, with which one could attempt to hit the destination B.

In order to construct the subsequent points of one single extremal, it is obvious that the value of the timefront gradient, to be indicated by the vector $\underline{\psi}$, is needed to be able to select the proper course α by means of the maximum principle. Fortunately it can be asserted that there are mathematical means to evaluate the change of this vector $\underline{\psi}$ with time, when following an extremal.

One can now construct one single extremal from A by taking a fair estimation for the inital value of $\underline{\psi}$. (See figure 3.4a). Assuming this initial $\underline{\psi}$ to have a unit length, this starting gradient becomes $\underline{\psi}(t_a) = (\cos\varphi_0, \sin\varphi_0)$, where φ_0 is the angle of $\underline{\psi}(t_a)$ with the X^+-

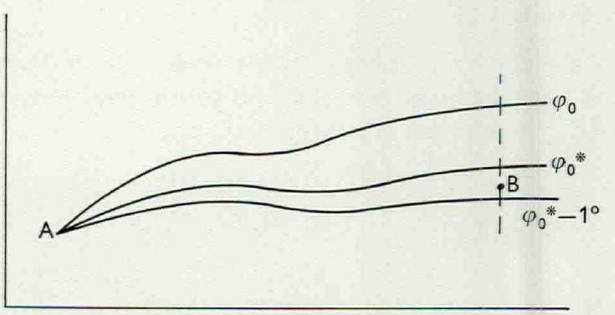


Fig. 3.4a. For various initial direction φ_0 of $\underline{\psi}$, extremals are constructed, until B is approximately hit by a specimen.

direction. The problem of finding the proper AB-extremal now seems to be to select φ_0 properly. This could be done by means of a well known and frequently used procedure called "trial & error", as shown in figure 3.4a.

Starting with a first guess for φ_0, this may produce an extremal, that passes North of B. Then φ_0 can be replaced by $\varphi_0 - 1°$ and the corresponding extremal could be constructed. This scanning could continue, until B is enclosed by two tracks with intitial gradient directions φ_0^* and $\varphi_0^* - 1°$. The proper φ_0 could then be fairly estimated by means of linear interpolation.

However this method is inadequate for solving the present problem. The cause of this is, that there may be more than one "extremal", that can be traced by means of this trial & error procedure. In the next paragraph this possibility of the occurrence of more than one solution will be shown by means of an example.

3.5 *Absolute and relative extremals*

Consider the velocity field $\underline{v} = (v\cos\alpha,\ v\sin\alpha)$, with $v = 2 - e^{-(x^2+y^2)/2}$, shown in figure 3.5a.

The velocity v has the minimal value 1 for $x = y = 0$. Furthermore, on a circle around the origin O with radius r the velocity has the constant scalar value

$$v_r = 2 - e^{-r^2/2}.$$

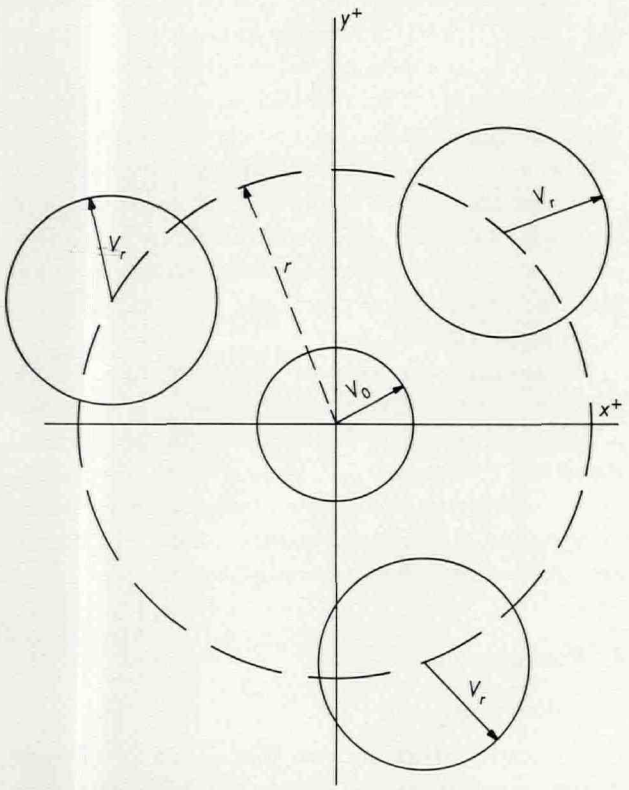


Fig. 3.5a. v_0: velocity in $O = 2 - e^0 = 1$; v_r: velocity at distance r from $O = 2 - e^{-r^2/2}$;

r	0.5	1	2	3	∞
v_r	1.12	1.39	1.86	1.99	2

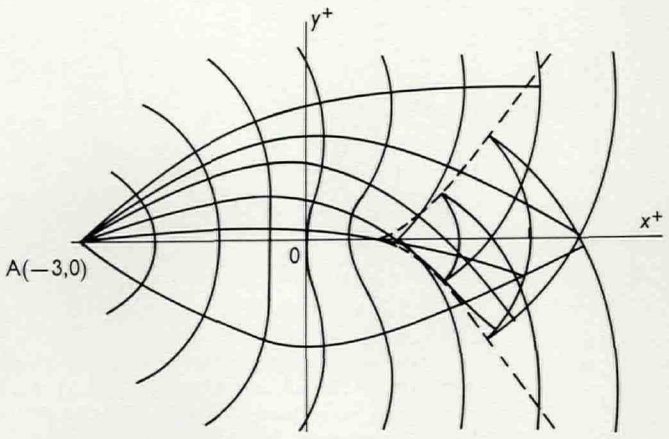


Fig. 3.5b. Complete figure of timefronts and extremals with start in A $(-3,0)$. The velocity indicatrix is a circle, the scalar velocity is a monotonously increasing function of $r = (x^2+y^2)^{\frac{1}{2}}$.

Starting at $(-3,0)$ it is now possible to construct extremals by numerical means for various initial values α_0 of the course α. This was done for $\alpha_0 = = -70°(0.5°)70°$. The total result is shown in figure 3.5b.

This figure shows that the timefronts have reverse points and a double point, after the origin is passed [6]. This phenomenon will now be briefly explained.

The velocity indicatrix for this case is a circle with radius $2 - e^{-r^2/2}$, where r is the distance to O. It will now be clear, that the timefronts, starting as almost circular convex curves, will become less curved near the X-axis as time increases. Thinking of a timefront $S^+(t+\delta t)$ to originate from $S^+(t)$ as the envelope of δt-position indicatrices, drawn from all points of

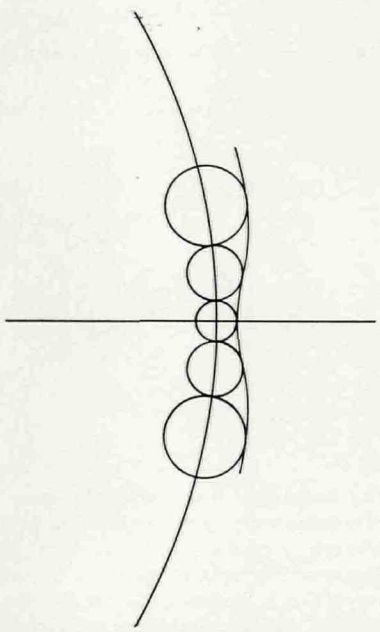


Fig. 3.5c. The origin of a concave dent in a timefront, when passing a velocity minimum.

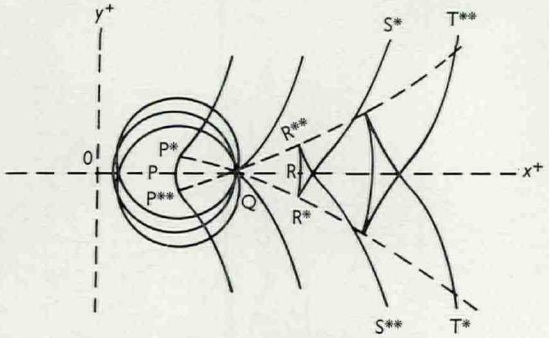


Fig. 3.5d. The formation of reverse points and double points in timefronts. The three circles thorugh Q have approximately equal radii. Their centress are P^*, P and P^{**} respectively.

$S^+(t)$, it is also clear, that the timefronts get a concave dent near the X-axis, like shown in figure 3.5c.

In figure 3.5d one can see, how the concave part P^*PP^{**} is compressed to one point Q after a time δt. After this, the timefront parts that correspond to extremals, passing North and South of O, start to overlap each other. The sector of reverse points of timefronts T^*QT^{**} envelopes the "extremals" to points of the rear parts, like R^*RR^{**}.

If B is located inside a sector like T^*QT^{**}, there are three solutions, i.e. three AB-extremals. This is shown in figure 3.5e.

In this figure, B is located on the timefronts $S^+(t_b)$ – fully drawn –, on $S^+(t_b')$ – dashed curve – and on $S^+(t_b'')$ – dotted curve –. It is clear that $t_b < t_b' < t_b''$, so that the trajectory, that takes the ship from A to B at a time t_b is the proper extremal.

In view of this phenomenon, it is obviously necessary to delete those parts of a timefront $S^+(t)$, that are not part of the boundary of $H^+(t)$. We also see, that the timefronts have to be constructed as extensively as possible.

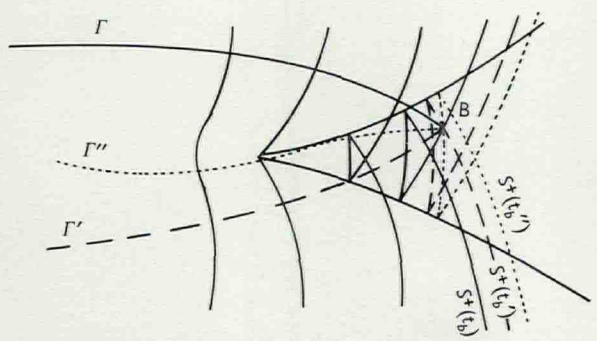


Fig. 3.5e. The destination B is reached by three "extremals", all constructed by local application of the maximum principle. Γ is the absolute extremal, with t_b as the corresponding arrival time. Γ' is a relative extremal, arrival time t_b', meaning that the arrival time t_b' is only minimal, when considering trajectories that pass South of the velocity minimum. Γ'' is no real extremal at all, not even a local one.

3.6 *Co-ordinate restrictions*

So far the extremals and timefronts were only considered in the event that they were located entirely inside G, the navigating region. Obviously, timefront arcs, that are located outside of G, are to be deleted. This deleting may however lead to the possibility, that B cannot be reached by following a common extremal. Figure 3.6a shows, that this possibility can very well occur.

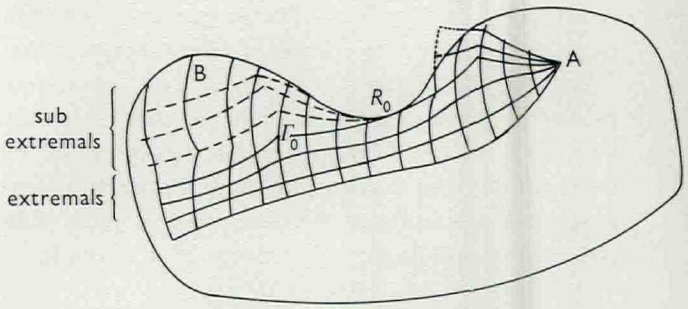


Fig. 3.6a. Free and isoperimetric extremals. Γ_0 is the bordering specimen of a family of free extremals. Γ_0 touches the boundary of G at R_0. The dashed drawn subextremals start in R_0 as boundary arcs of G. Sooner or later they leave this boundary again, proceeding from then on as extremals inside of G.

The extremal Γ_0 touches the boundary of G at R_0. Assume the corresponding time to be t_0. Then the arc of Γ_0 past R_0 forms an undesired limit for $H^+(t)$ for $t > t_0$, as it is impossible to construct extremals on the other side of Γ_0, because they would have arcs outside of G. The solution to this problem is rather simple:

In R_0 we can follow a certain stretch of the boundary curve and then let the extremal go its normal way to the inside of G again. These trajectories are called subextremals. They are to be constructed to such an extend that the corresponding timefronts reach as far as the boundary of G.

The mathematical analysis, that is the basis of this solution, is rather complicated. Besides, the indicated principle is not rigorously followed in the practical numerical construction of the extremals. Therefore, these mathematical details are omitted here. However, in Appendix II, interested readers can find the derivation of this "revised maximum principle".

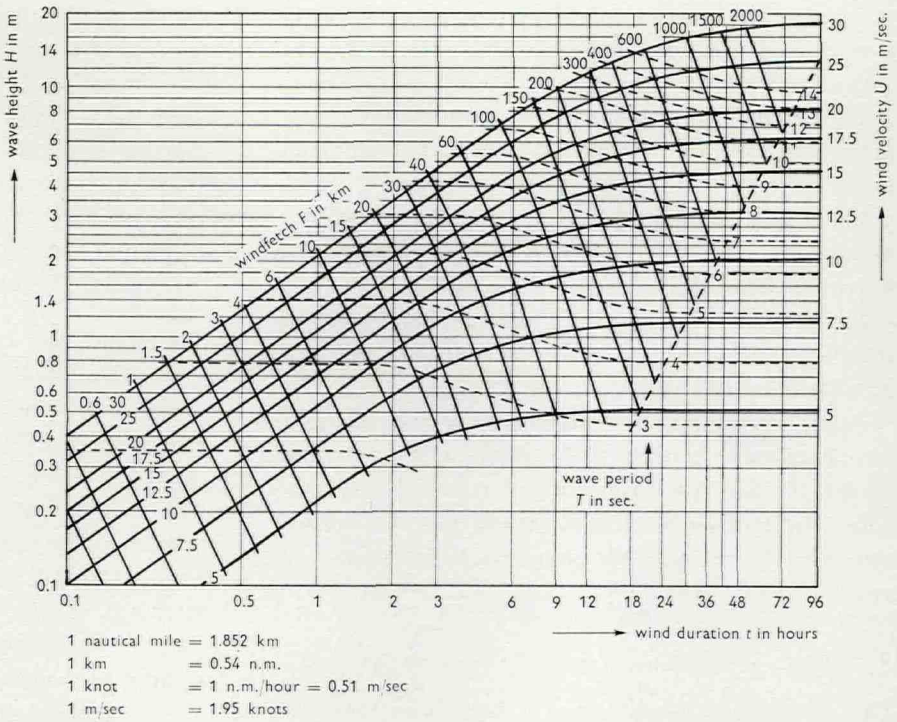
4 Practical data

4.1 *Wave prediction*

Briefly stated, there are two fundamentally different mathematical models as a basis for wave prediction techniques [7].

The Sverdrup-Munk theory adopts the concept of a "Significant wave", defined as an elementary sine-

Fig. 4.1a.
Graphs to determine significant height H and period T of wind-generated sea waves from wind velocity U and either duration t or fetch F.



wave, that is supposed to represent the entire sea wave pattern, as far as the energy density is concerned, while the phase speed equals the average of these speeds of all the sea wave components.

In spite of the fact, that this assumption meets some mathematical objections, the results of the computations, based on this theory, are fairly in accordance with practice, as far as the prediction of wind generated sea waves is concerned.

If the wind has blown in a certain area during a time t with a wind velocity U, while the wind fetch F is known, then t, U and F determine the "significant wave height", which can be defined as the average wave-height of the highest $\frac{1}{3}$ part of a large number (about 1000) of waves.

This determination of the significant wave height can be done by means of a nomogram, like shown in figure 4.1a.

This graphical aid is a reprint from [8], a very handy and practical booklet.

The investigations of ship's performances in sea waves are practically all based on a stochastical sea wave model, designed by Pierson. The sea is assumed to be perfectly chaotic, while the total sea wave pattern is seen as a linear combination of waves of various frequencies (from 0 to ∞) and directions (from $-\pi/2$ to $\pi/2$ with respect to the mean wind direction). A spectral function S, depending on the frequency σ and the direction relative to the mean wind direction ϑ_0, gives

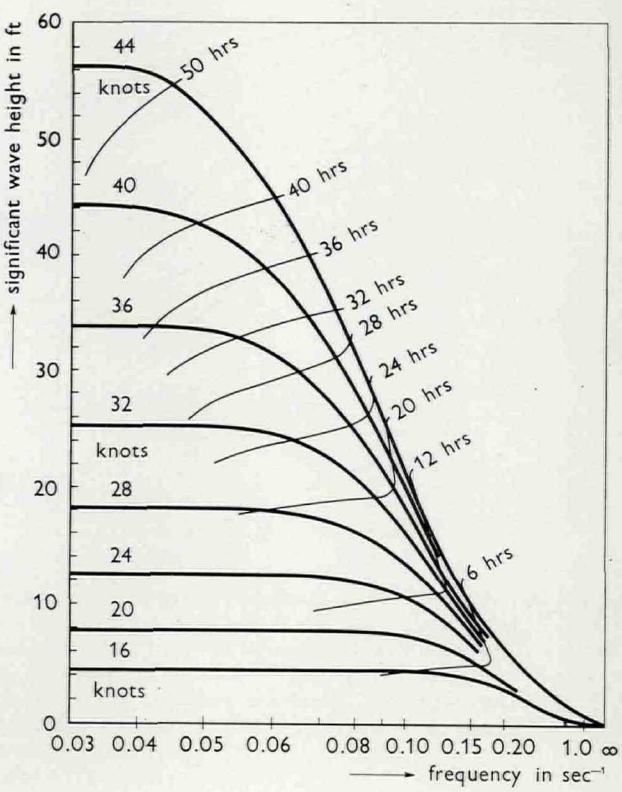


Fig. 4.1b. Nomogram, based on the Neumann spectrum, to determine the significant wave height of wind generated sea waves as a function of the wind speed and the wind duration. For instance, a wind speed of 28 knots and a wind duration of 12 hours can generate a sea wave pattern with an average height of 11 feet. To see this, project the intersection of the 28-knots-line and the 12-hours-line on the vertical scale. In a similar diagram, the wave height can be determined as a function of the wind speed and the wind fetch. Denoting these heights by H_d and H_f, the significant wave height, predicted by the Pierson-Neumann model, becomes the minimum of H_d and H_f.

the probability distribution of the energy of waves in a frequency-direction element with value (σ, ϑ) and with band widths $\Delta\sigma$ and $\Delta\vartheta$.

The designation of a suitable spectral energy distribution function $S(\sigma, \vartheta)$ was first done by G. Neumann in 1952.

Using this Neumann spectrum, one can compose a nomogram, shown in figure 4.1b. This figure was borrowed from [9].

Comparison of these wave predictions, based on the Neumann spectrum, with practical data, did not give satisfactory results.

After Neumann's first estimation, several attempts have been made to improve this wave energy spectrum. So far, the best result was booked in 1965 by Moskowitz. The Pierson-Moskowitz model has also turned out to be the best possible design for the prediction of swell.

4.2 *Ship's performance data*

The analysis and prediction of a ship's behaviour in a given sea wave pattern has turned out to be a very difficult problem.

Graphs of ship's speeds in sea waves with various heights for waves coming in from ahead, from one of the beams or from astern have been published for a variety of types of ships. Examples are shown in figure 4.2a.

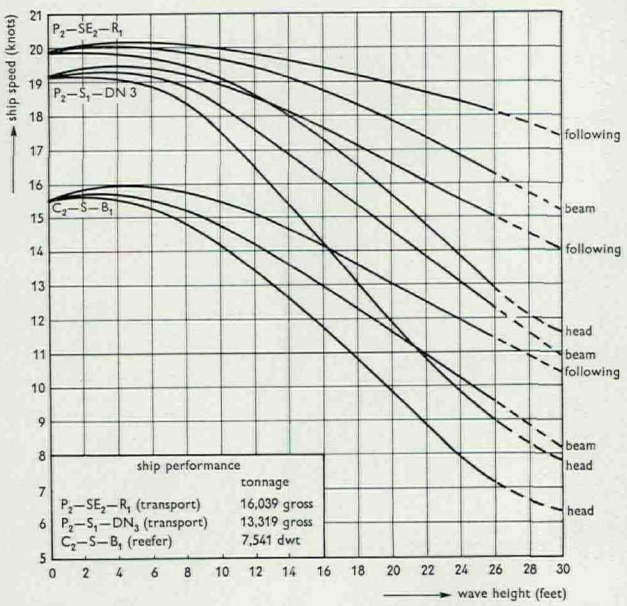


Fig. 4.2a. Maximum speed graphs for three types of ships, experimentally determined by the U.S. Naval Oceanographic Office. For example, for the C_2-type freighter with a normal speed of 15.5 knots, in waves with 20 feet significant height the quantities d_1, d_2 and d_3 – see 5.2 – become: $d_1 = 12^h \times 13$ knots $= 156$ miles
$d_2 = 12^h \times 11.5$ knots $= 138$ miles
$d_3 = 12^h \times 9.8$ knots $= 117.6$ miles

They were practically all based on experimental data, revealed by practical experience [10]. Their reliability is far from optimal. In the Netherlands this problem has been and is still being worked on rather vigourously by a group of naval architects, attached to the Shipbuilding Laboratory of the Delft University of Technology and directed by Gerritsma. A brief summary of the results of the research work will now be given [11, 12, 13, 14, 15, 16].

a. The chaotic sea wave pattern is seen as a composition of mutually independent elementary sinewaves with various frequencies, wave-heights and directions. For the energy probability distribution the Pierson-Moskowitz spectrum is adopted.
b. The various ship's movements also are linear compositions of elementary harmonic oscillations. The most important ship movements are:
   pitching: oscillations about a thwartships axis through the ship's gravity centre G,
   rolling: oscillations about a longships axis through G,
   heaving: vertical shifting of G,
   swaying: thwartships shifting of G.

For a given elementary sinewave with a given amplitude and apparent frequency σ^* it is possible to evaluate the input values of all the movements mentioned above.

There are various methods to determine the ship's responses to these inputs. A response is characterized by two numbers, the first of which is the amplitude ratio $H = A_s/A_g$, with A_s = the real amplitude of pitching, heaving etc. and A_g = the sea wave input amplitude, while the second one is the phase lag ε of the real movement with respect to the sea input.

The amplitude ratios and phase lags appear to be fairly well independent of the input amplitudes and they only depend upon the input frequencies.

The quantities $H(\sigma^*)$ can be used to determine the pitch, heave, roll and sway spectra by simply multiplying the corresponding wave input spectra $A_g(\sigma^*)$ by $H(\sigma^*)$.

(Note that the amplitude probability distribution is given as a function of the true frequency σ, so that it is necessary to transform this to a function of the apparent frequency σ^*, according to $\sigma^* = \sigma + \sigma^2 s/g$, where s is the projection of the ship's speed on the wind direction.)

The phase lag ε plays an important part in the estimation of the ship's speed decrease. It is also important, when questions like the coincidence of pitch and heave are considered.

Thus the knowledge of these frequency responses open the possibility for

a. evaluation of the ship's resistance and
b. evaluation of the probability, that some undesired phenomenon occurs, like too much variance in the accelerations in the ship's fore and aft parts, slamming or shipping green water.

It is also possible to evaluate the ship's speed, at which one of these phenomenae has an occurrence probability less than a desired maximum.

Briefly stated, the ship's performance can be determined by sea conditions as well as by a set of wishes of the ship's master. The occurrence of sectors of forbidden courses, mentioned in 3.1, can be made plausible by considering the following facts:

a. The apparent frequency of waves, that come in from aft the beam, is rather low. The ship may be sort of caught in a wave trough for a couple of minutes. On account of the small midships draught the couple to decrease a possible list of some 20° or 30° may be very small or even zero on account of the small metacentric height. As a result, the ship maintains a considerable list for quite a while. This phenomenon has been frequently experienced on merchant ships, going some 25 knots.
b. The low apparent wave frequency may cause the ship's stern to be lifted from the still water level for quite a while. This may result in bad steering conditions and heavy yawing.
c. The roll-generating torque can have a periodic component with a period that is favourable to lead the ship to a steady state of constant rolling, i.e. constant in amplitude. For practically all kinds of cargo, stowed either classically or in big load units, this rolling resonance can bring ship and cargo into an uncontrollable situation.

5 Program to evaluate the optimal track

5.1 *Oceanographic, meteorological and ship's performance data*

In the polar stereographic projection or in the Kahn projection of the concerned area an X- and Y-axis are assumed and the mapping near the origin of a stretch of 150 or 200 nautical miles is taken as a length unit. Points with integer co-ordinates are taken as data grid points. Thus the sea current and wave data are introduced as a two-dimensional array, like sxr, syr, hzr, hrx, hry, tzr [0:18, −3:5].

In the arrays sxr$[i,j]$ and syr$[i,j]$ the values of the ocean current x- and y-components are read, taking 12^h as a time unit. The arrays hzr$[i,j]$ and tzr$[i,j]$ are to contain the significant wave heights and the mean directions that these waves come from.

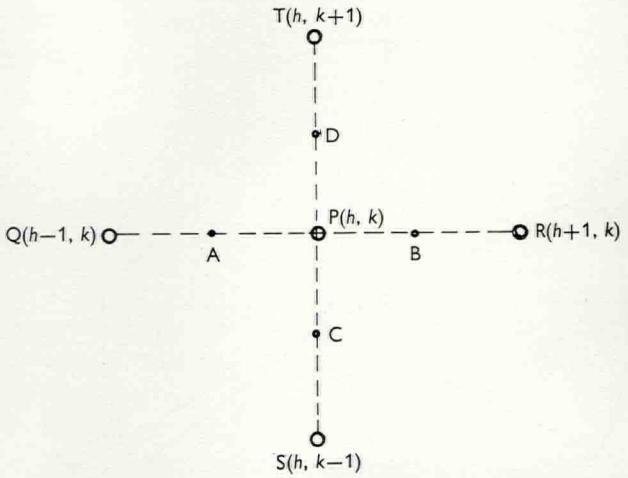


Fig. 5.1a. Grid points P, Q, R, S, T and in-between points A, B, C, D to explain the computation of the partial derivatives *hrx* and *hry* of the sea wave height *hzr* with respect to x and y.

The grid of points with integer co-ordinates is dense enough to get a fair representation of the sea current components and of the wave direction, if interpolated linearly between grid values.

The wave data can be taken from wave prediction charts, constructed by the K.N.M.I. routeing office. A program for semi-automatic copying of these data is being prepared, as well as a program to determine wave data from wind velocity, duration and fetch.

The quantities hrx$[i,j]$ and hry$[i,j]$ stand for the increases of the significant wave height in the X^+- and Y^+-directions respectively.

Considering a point P (fig. 5.1a) with integer co-ordinates (h,k) and the neighbouring points $Q(h-1,k)$, $R(h+1,k)$, $S(h,k-1)$ and $T(h,k+1)$, we can take the in-between points:

A, between P and Q, co-ordinates $(h-\frac{1}{2},k)$,
B, between P and R co-ordinates $(h+\frac{1}{2},k)$,
C, between P and S, co-ordinates $(h,k-\frac{1}{2})$ and
D, between P and T, co-ordinates $(h,k+\frac{1}{2})$.

Denoting the waveheights in A, B, C and D by h_a, h_b, h_c and h_d, the waveheight increments are now defined by

$$\text{hrx}[h,k] = h_b - h_a \quad \text{and} \quad \text{hry}[h,k] = h_d - h_c.$$

These quantities had to be introduced to prevent small areas of high waves from being neglected by linear interpolation. As for the ship's performance data, there are five quantities to determine the 12^h-position indicatrix, denoted by d_1, d_2, d_3, ψ_{g1} and ψ_{g2}. Their significance is illustrated in figure 5.1b.

d_1: ship's distance, made good in 12 hours with wind and waves coming in from astern,

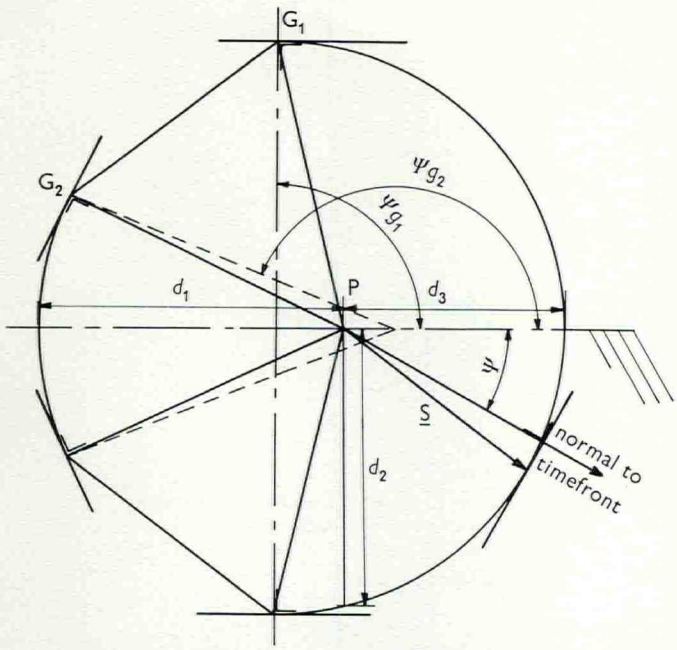


Fig. 5.1b. This figure shows an example of the 12^h effective position indicatrix. The notations are explained in the text of this paragraph.

d_2: ship's 12^h-distance with wind and waves coming in from one of the beams,

d_3: ship's 12^h-distance with wind and waves coming in from ahead.

If G_1 is the windward border of one of the sectors of course constraints, then ψ_{g1} is the angle between the line, perpendicular to the indicatrix tangent at G_1 and the wind direction.

Let G_2 be the leeward border of that sector. Draw a line in G_2, that touches the indicatrix there and then draw the line in G_2, that is perpendicular to this tangent. This last line makes an angle ψ_{g2} with the wind direction.

These indicatrix parameters are computed by means of linear interpolation between grid values, stored in the array elements d[1:3, 0:30] and cospsig[1:2, 0:30].

As for d[h,k] – h and k are integers with $1 \leqslant h \leqslant 3$ and $0 \leqslant k \leqslant 30$ – the first index (h) indicates that the wind comes in from astern (1), abeam (2) or ahead (3). The second index (k) refers to a waveheight of $k/2$ meters. The array cospsig h,k contains the values of cos ψ_{gi} with $i = 1,2$ for waveheights of 0(0.5)15 meters.

5.2 *Evaluation of the subsequent timefronts*

Let the co-ordinates of a number of points of a timefront S_j, for a time $t_j = t_a + j \times 12$h, be known. These co-ordinates are denoted by x_{i0} and y_{i0}. The number of points is about 40. If the distance between two subsequent points is more than 40 nautical miles, then a point on the middle of the corresponding segment is added. This is executed by a procedure, called "verfijn". The timefront gradient direction (a) can now be evaluated, like shown in figure 5.2a.

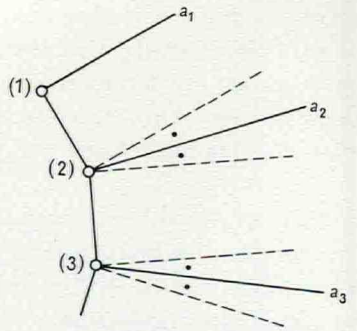


Fig. 5.2a. The timefront is known by the location of the discrete points (1), (2) etc. The gradient in a vertex, like (1), is taken to be perpendicular to the segment (1)-(2). For other points, the gradient's direction is approximated by the direction of the bissectrix of the perpendiculars to the line segments from the point in question to the two neighbouring points.

This is executed by a procedure, called "normaal".

With the co-ordinates x_{i0}, y_{i0} and the value of a the first estimation of the optimal 12^h-step can be evaluated. This happens by calling the procedure "optistep". Briefly stated, this procedure works as follows.

1. The current components sx and sy, the waveheight hz and the wave direction tz are estimated by linear interpolation between grid values.
2. The indicatrix parameters d_1, d_2, d_3, ψ_{g1} and ψ_{g2} are determined by linear interpolation, using hz as the determining variable.
3. The entire 12^h-step has the components dx and dy. The first assignments to get these components are: "dx:= sx" and "dy:= sy".

Looking back at figure 5.1b we now take the point with co-ordinates $(x_{i0}+sx, y_{i0}+sy)$ as the origin and rotate the axes, so that the wind direction tz is the new X^+-direction. The timefront's normal direction now is $\psi = a - tz$. Now an ellips is constructed through the points with local co-ordinates $(-d_1,0)$, $(0,d_2)$ and $(d_3,0)$. For a given value of ψ the optimal 12^h-displacement, which is the displacement with the longest projection on the ψ-direction, can now be computed. The displacements dx' and dy' with respect to the local axes now have to be transformed back to the above mentioned amounts dx and dy.

Remembering that the sea current was already accounted for, we have to give the assignments

$$\mathrm{d}x := \mathrm{d}x + \mathrm{d}x' \times \cos(tz) - \mathrm{d}y' \times \sin(tz);$$

$$\mathrm{d}y := \mathrm{d}y + \mathrm{d}x' \times \sin(tz) + \mathrm{d}y' \times \cos(tz);$$

This completes the optistep procedure.

The co-ordinates of the points of the first estimation of S_{j+1}, denoted by S^*_{j+1}, are denoted by x_{i1} and y_{i1}.

After reading in new sea wave data for the time $t_{j+1} = t_j + 12^h$, the procedures "normaal" and "optistep" are repeated, giving new displacements dx and dy. These displacements are computed for the positions of the points of S^*_{j+1} with co-ordinates (x_{i1}, y_{i1}). The co-ordinates of the "corrected" points of S_{j+1}, denoted by x_{i2} and y_{i2}, are obtained by giving the assignments

$$x_{i2} := (x_{i1} + x_{i0} + \mathrm{d}x)/2;$$

$$y_{i2} := (y_{i1} + y_{i0} + \mathrm{d}y)/2;$$

5.3 *Timefront revision*

The navigated region G is scanned by a collection of extremals, which start in A with a 40° wide sector. After each timefront evaluation there are four reviding procedures to be applied.

1. "rand": Points outside of G are deleted.
2. "schoonmaak" and
3. "ontlussen": With these tests, points of S_{j+1}, that are no bordering points of the set of reachable points H_{j+1}, are deleted, being irrelevant for
4. further evaluations.
   "verfijn": This procedure, already mentioned in 5.2, takes care of a sufficiently dense net of extremals, with which the area is scanned.

In order to explain the procedures "schoonmaak" and "ontlussen", consider figures 5.3a and 5.3b.

In 5.3a the points with numbers 14, 15, 16, 17 and 18 have to be deleted. This is done, by checking if a timefront segment – here the segment 13–14 – has an intersection with any following segments. In this case the timefront chord 18–19 intersects the 13–14 chord and the points with numbers greater than 13 and less than 19 are deleted as a result. Point nr. 19 gets number 14 etc.

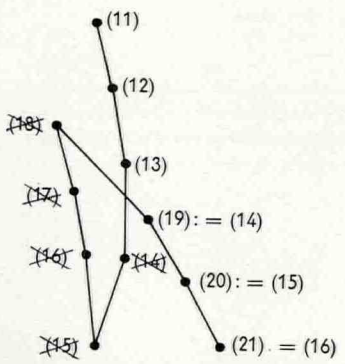


Fig. 5.3a. The occurrence of overlapping of a timefront is frequently met after passing a region with comparatively high waves. Rear echelon points are delected, being no part of S_{j+1}, the boundary of H_{j+1}.

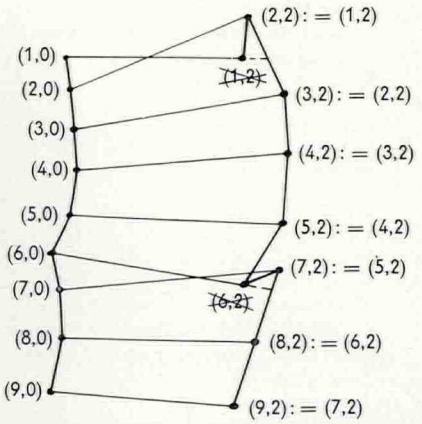


Fig. 5.3b. As timefronts are built up by straight line segments between discrete points, the occurrence of one single rear echelon point is also possible, especially near the timefront's vertices. Being not part of S_{j+1}, they also have to be deleted.

Another possible situation to occur is shown in figure 5.3b. This situation is checked for by seeing if the extension of a segment from (x_{i0}, y_{i0}) to (x_{i2}, y_{i2}) has an intersection with any of the segments between two consecutive points of S_{j+1}.

It is necessary to maintain an accurate bookkeeping of the numbers that are deleted by the procedures "ontlussen" and "schoonmaak", while the computing program must also include a recording of the new numbers, brought in by the refinement procedure "verfijn". These numbering mutations are stored in two arrays, nu$[h, j]$ and mu$[h, j]$. These arrays are used, when the destination B is overwashed by a timefront and the indices of the consecutive points of the optimal track have to be traced back.

5.4 *Determination of the optimal trajectory and the time gain*

The passage of B is checked for by a procedure called "fuik". To explain this procedure, consider the revised

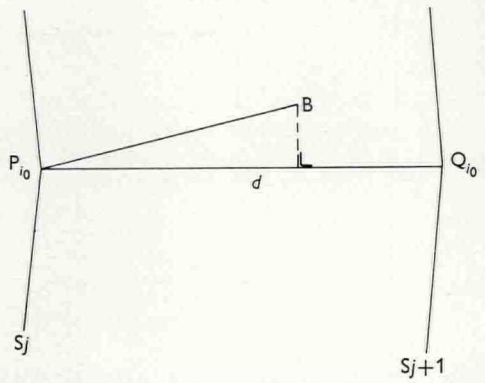


Fig. 5.4a. The destination B is found to lay outside of H_j, while being an inner point of H_{j+1}. The timefront computing procedure now stops and the optimal track from A to B can be determined.

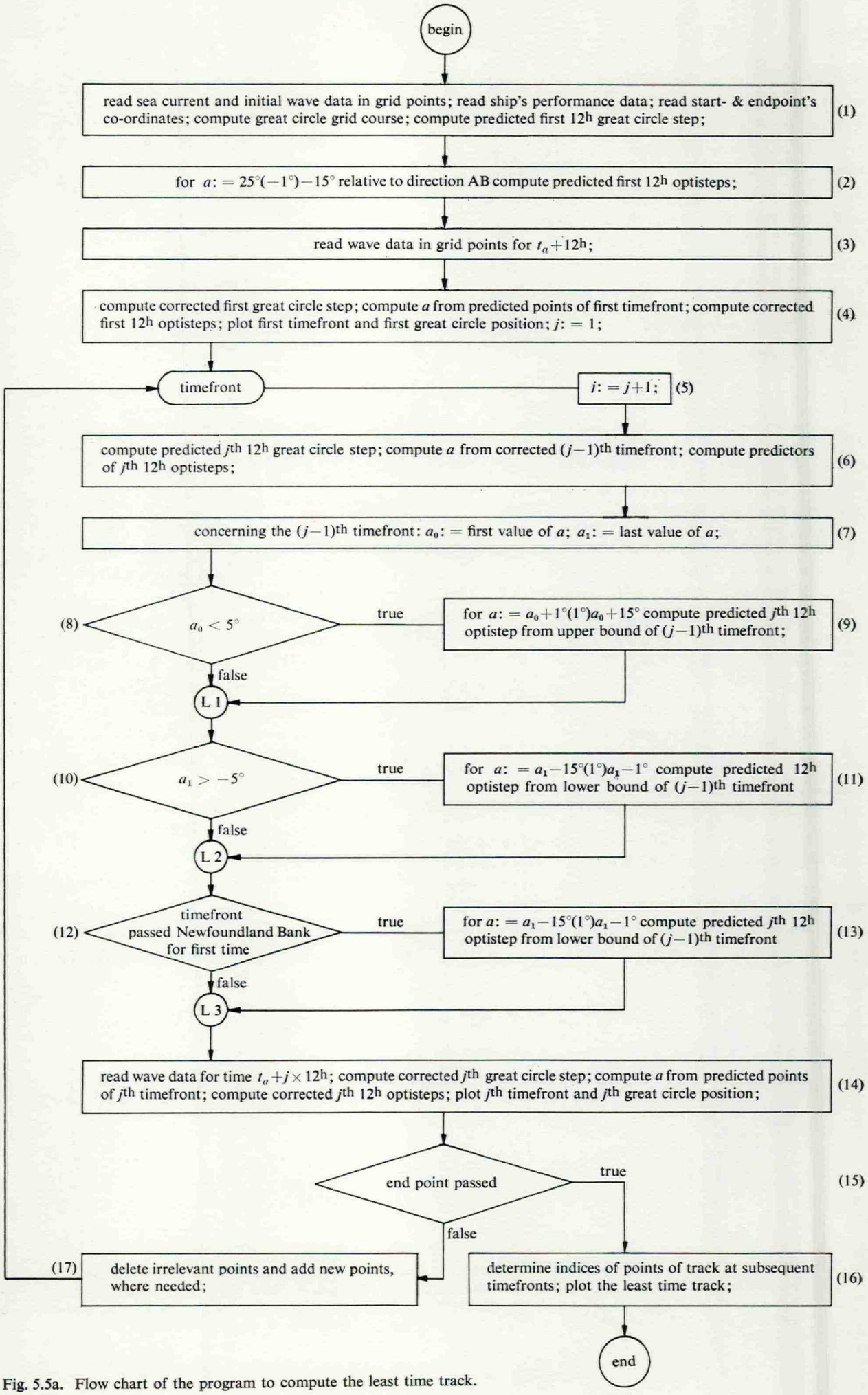


Fig. 5.5a. Flow chart of the program to compute the least time track.

timefront S_j and the unrevised timefront S_{j+1}. First the index is determined of the point of S_j, that is closest to B. This index value is assigned to the integer variable i_0. Let us denote this point by P_{i_0} and the corresponding point of S_{j+1} by Q_{i_0}. Let the distance from P_{i_0} to Q_{i0} be d. We can now say, that B is covered by H^+_{j+1}, if the projection of $P_{i_0}B$ onto $P_{i_0}Q_{i_0}$ is not greater than d.

As soon as this is found to be true, the arrival time and the points of the optimal track can be determined.

A fair idea about the time gain of the optimal track can be obtained by comparing it with the time, needed for the same ship to pass the same composition of obstacles, when steering a great circle's course.

5.5 *Flow chart of the program to compute the least time track*

See fig. 5.5a on page 20.

5.6 *An application to practice*

The computing program, corresponding to the flow chart of the preceding paragraph, was applied to a fictive merchant ship with a service speed of $15\frac{1}{2}$ knots and a set of performance graphs, taken frpm figure 4.2a. The sectors of forbidden courses were taken in accordance with data, that were gathered by the K.N.M.I. routeing office from practical experience. The ship was supposed to sail from Bishop Rock to Nantucket Shoals, starting on March 1st 1967, 0^h GMT.

The result of the computation is shown in figure 5.6a.

The timefronts in this figure are marked $S_{0.5}$, S_1, $S_{1.5}$ etc. For example : $S_{3.5}$ is the timefront for March 4th, 12^h GMT.

The arrows near the timefronts refer to the sea waves, i.e. wind generated waves when fully drawn and swell when drawn dashed. The numbers near the arrows give the waveheights in meters. The optimal track is the fully drawn curve, composed of straight line segments from timefront to timefront. Possible small course alterations and tacking manoeuvres are not indicated, being a matter in wich the ship's master is autonomous.

The dotted line is the great circle track, on which the positions after 0.5^d, 1^d, 1.5^d etc. are marked by crosses.

Regions of iceberg danger near Newfoundland and Nova Scotia were taken into account by simulating extremely high waveheights in the grid points near by.

5.7 *Considerations regarding data incertainties and practical use*

As noted before, the reliability of weather predictions can be taken inversely proportional to the time last ahead, the prediction is given for. This means, that the least time track, constructed on the basis of weather and sea estimations for a time, from 0 to 8 days ahead, can only be taken as a real extremal for the next one or two days to come. A change in the predicted data would mean, that the track has to be revised. If a ship would thus follow a day-by-day corrected extremal,

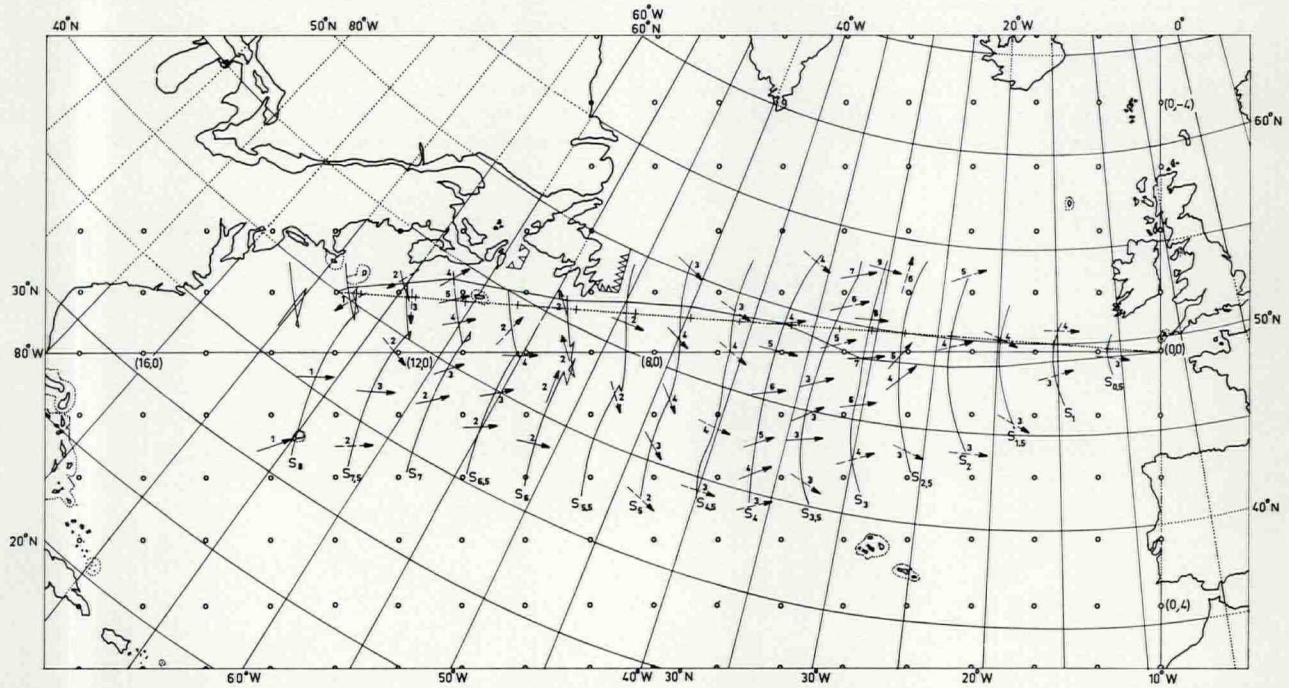


Fig. 5.6a. Result of the computation of an optimal track from the English Channel to the U.S. East coast. The least time track (fully drawn) is seen to avoid areas of extremely high waves and it also shows the advantages of making for the leeward side of Newfoundland as soon as possible.

the entire track will consist of parts, that are extremal arcs from day to day, but the composition of these extremal arcs will in general not be an extremal.

It seems rather precarious to make a numerical estimation of the error, that is made this way. A possibility to collect statistical data on this matter would be to simulate ship's crossings in the past, based on weather and sea predictions that were really made at that time, and then compare this with the real extremal, that should have been taken.

As a concluding remark on the efficiency it can be stated, that the past seven years of experimental ship routeing did not show a substantial average time gain. However, the decrease of cases of extremely high storm damage as well as the better predictability of the expected time of arrival – an essential figure for ships that carry general cargo to the North American East Coast – can be marked as important advantages of this method of shore-piloted ocean navigation.

For other methods of numerical ship routeing the references [17, 18, 19, 20] can be consulted.

References

1. Wit, C. de, Mathematical Treatment of Optimal Ocean Ship Routeing, Thesis Delft, 1968.
2. Driencourt, L. and J. Laborde, Traité des Projections des Cartes Géographiques, Hermann et Cie, Paris, 1932.
3. Roon, J. van, Leerboek der Zeevaartkunde III, C. de Boer Jr., Hilversum, 1949.
4. Pontryagin et al., L. S., Mathematical Theory of Optimal Processes, Pergamon Press, 1964.
5. Halkin, H., On the necessary conditions for optimal control of nonlinear systems, Journal d'Analyse Mathématique, Jerusalem, 1964.
6. Jong, H. M. de, Theoretical Aspects of Aeronavigation and its Application in Aviation Meteorology, Staatsdrukkerij- en Uitgeverijbedrijf, The Hague, 1956.
7. Blair Kinsman, Wind waves, their generation and propagation on the ocean surface, Prentice Hall, 1965.
8. Groen, P. and R. Dorrestein, Zeegolven, Staatsdrukkerij- en Uitgeverijbedrijf, The Hague, 1958.
9. Pierson, W. J., G. Neumann and R. W. James, Practical Methods for Observing and Forecasting Ocean Waves, U.S. Hydrographic Office Publ. No. 603, Washington, 1955.
10. Joseph, E. J. and J. M. Kipper Jr., Wave Climatology as an Aid to Ship Routing in the North Atlantic Ocean, U.S. Naval Oceanographic Office, Washington D.C., 1963.
11. Gerritsma, J., Behaviour of a ship in a sea-way, Neth. Ship Research Centre TNO Report no. 84 S, Delft, 1966. Also: Intern. Shipb. Progress, July 1966.
12. Gerritsma, J., J. J. van den Bosch and W. Beukelman, Propulsion in regular and irregular waves, Shipbuilding Laboratory, Delft Univ. of Techn., Publ. No. 17, 1961.
13. Leeuwen, G. van, The lateral damping and added mass of a horizontally oscillating shipmodel, Neth. Ship Research Centre TNO Report no. 65 S, Delft, 1964.
14. Gerritsma, J. and W. Beukelman, Analysis of the modified strip theory for the calculation of ship motions and wave bending moments, Neth. Ship Research Centre TNO, Report no. 96 S, Delft, 1967.
15. Wahab, R., Amidships forces and moments on a "Series 60" model in waves from various directions, Neth. Ship Research Centre TNO, Report no. 100 S, Delft, 1967.
16. Vugts, J. H., The hydrodynamic coefficients for swaying, heaving and rolling cylinders in a free surface, Neth. Ship Research Centre T.N.O., Report no. 112 S, Delft, 1968.
17. Faulkner, F. D., Numerical methods for determining optimum ship routes, Journal of the Institute of Navigation, Vol. 10, No. 4, U.S.A., 1964.
18. Bleick, W. E. and F. D. Faulkner, Minimal-Time Ship Routing, Journal of Applied Meteorology, U.S.A., April 1965.
19. Nagle, F. W., Ship Routing by Numerical Means, U.S. Navy Weather Research Facility, Report 32-0361-042, Norfolk, Virg., 1961.
20. Haltiner, G. J., W. E. Bleick and F. D. Faulkner, Use of Long-Range Weather Forecasts in Ship Routing, U.S. Navy Weather Research Facility, Norfolk, Virg., 1965.

APPENDIX 1

Differential equations to construct an extremal

Let $\{\Gamma\}$ be the collection of all extremals from A to a point of the timefront $S^+(\tau)$ with $\tau > t_a$ (see figure A.1a). Otherwise stated, $\{\Gamma\}$ is the collection of solutions of $d\underline{x}/dt = \underline{f}(\underline{x}, \alpha, t)$ with initial point $\underline{x}(t_a) = \underline{x}_a$ and $\underline{x}(\tau)$ on $S^+(\tau)$.

A subcollection C of $\{\Gamma\}$ is formed by a field of extremals – i.e. a family of extremals with no mutual intersection points except A – with the additional property, that the timefronts $S^+(t)$ are differentiable curves, as far as they correspond to extremals of the C-field. The extremals of C, that cover a subregion G_c of $H^+(\tau)$, can be characterised by attaching a real number s to each specimen. Thus the co-ordinates of a point of G_c are functions of t and of s: $x^i = x^i(t,s)$ with $i = 1,2$.

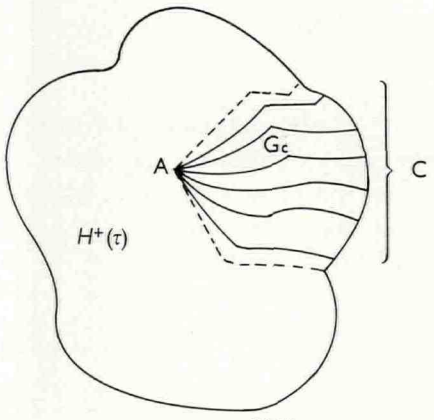


Fig. A.1a.

The timefronts that correspond to C were assumed to be differentiable curves, so $x^i(t,s)$ is partially differentiable with respect to s everywhere in G_c, except of course in the boundary points, where there is only left or right differentiability.

Let Γ_0 be an extremal of C with $s = s_0$, corresponding with the navigating function $\alpha_0(t)$. Take two time instants t_1 and t_2 so that $t_a < t_1 < t_2 \leqslant \tau$. Let Γ_0 intersect the timefronts $S^+(t_1)$ and $S^+(t_2)$ in P_1 and P_2 respectively, so

$$x^i_{P_j} = x^i(t_j, s_0) \quad \text{for} \quad i,j = 1,2.$$

Consider a neighbouring extremal Γ_ε, also a member of the C-family, generated by the navigating function $\alpha_\varepsilon(t)$. Let Q_1 and Q_2 be the intersecting points of Γ_ε with $S^+(t_1)$ and $S^+(t_2)$. Let the s-value of Γ_ε be $s_0+\varepsilon$. Then we have

$$x^i_{Q_j} = x^i(t_j, s_0+\varepsilon) \quad \text{for} \quad i,j = 1,2.$$

This situation is exposed in figure A.1b.

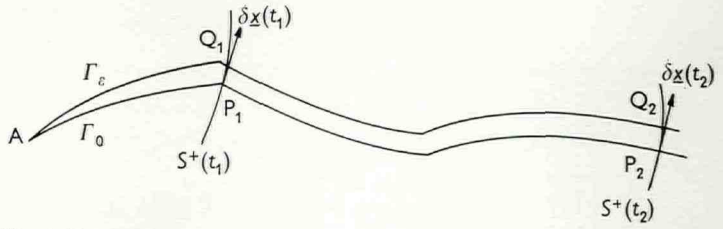


Fig. A.1b.

The tangent vector, that touches $S^+(t)$ in a point of Γ_0, is defined by

$$\delta x^i(t) = \frac{\partial x^i(t, s_0)}{\partial s}, \quad (i = 1,2) \tag{A.1.1}$$

so that

$$x^i(t, s_0+\varepsilon) = x^i(t, s_0) + \varepsilon\delta x^i(t) + o(\varepsilon) \tag{A.1.2}$$

where $o(\varepsilon)$ has the property

$$\lim_{\varepsilon\to 0} \frac{o(\varepsilon)}{\varepsilon} = 0.$$

As for the two neighbouring extremals Γ_0 and Γ_ε it can be asserted that the velocities at a time t in the points $\underline{x}(t,s_0)$ of Γ_0 and $\underline{x}(t,s_0+\varepsilon)$ of Γ_ε are also "close to each other", except possibly for the occurence of a finite number of bending points, i.e.

$$\lim_{\varepsilon\to 0} f^i(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t) = f^i(\underline{x}(t, s_0), \alpha_0(t), t) \tag{A.1.3}$$

almost everywhere on $[t_a, \tau]$.

A "one-point-exception" to this is shown in figure A.1c.

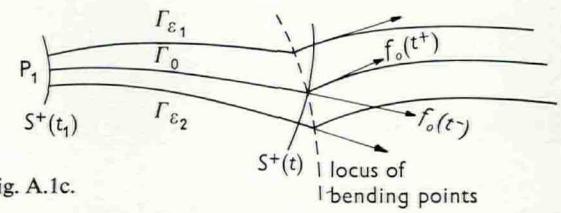


Fig. A.1c.

Suppose $\varepsilon_1 > 0$ and $\varepsilon_2 < 0$, then it is clear that the velocities $\underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t)$ have different limits for $\varepsilon \uparrow 0$ and $\varepsilon \downarrow 0$. These limits are indicated in figure A.1c as $\underline{f}_0(t^-)$ and $\underline{f}_0(t^+)$. May the attention now be concentrated on the way, the timefront tangent $\delta\underline{x}(t)$ changes with time along Γ_0. We can put

$$\underline{x}(t_2, s_0+\varepsilon) = \underline{x}(t_1, s_0+\varepsilon) + \int_{t_1}^{t_2} \underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t)\,dt$$

and

$$\underline{x}(t_2, s_0) = \underline{x}(t_1, s_0) + \int_{t_1}^{t_2} \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)\mathrm{d}t.$$

Subtracting this and using (A.1.2) gives

$$\varepsilon\delta\underline{x}(t_2) = \varepsilon\delta\underline{x}(t_1) +$$

$$+\int_{t_1}^{t_2} \{\underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t) - \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)\}\,\mathrm{d}t + \underline{o}(\varepsilon), \tag{A.1.4}$$

where $\underline{o}(\varepsilon)$ denotes a vector in R^2 with components of the $o(\varepsilon)$-type. The expression (A.1.4) can be split up in two ways:

1:

$$\varepsilon\delta\underline{x}(t_2) = \varepsilon\delta\underline{x}(t_1) + \underline{I} + \underline{J} + \underline{o}(\varepsilon) \tag{A.1.5}$$

with

$$\underline{I} = \int_{t_1}^{t_2} \{\underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t) - \underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_0(t), t)\}\,\mathrm{d}t \tag{A.1.5a}$$

and

$$\underline{J} = \int_{t_1}^{t_2} \{\underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_0(t), t) - \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)\}\,\mathrm{d}t \tag{A.1.5b}$$

2:

$$\varepsilon\delta\underline{x}(t_2) = \varepsilon\delta\underline{x}(t_1) + \underline{I}^* + \underline{J}^* + \underline{o}(\varepsilon) \tag{A.1.6}$$

with

$$\underline{I}^* = \int_{t_1}^{t_2} \{\underline{f}(\underline{x}(t, s_0), \alpha_\varepsilon(t), t) - \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)\}\,\mathrm{d}t \tag{A.1.6a}$$

and

$$\underline{J}^* = \int_{t_1}^{t_2} \{\underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t) - \underline{f}(\underline{x}(t, s_0), \alpha_\varepsilon(t), t)\}\,\mathrm{d}t \tag{A.1.6b}$$

The expressions for $\underline{J}$ and $\underline{J}^*$ can be written as

$$\underline{J} = \int_{t_1}^{t_2} \left\{ \sum_{j=1}^{2} \frac{\partial \underline{f}(\underline{x}(t, s_0(, \alpha_0(t), t)}{\partial x^j} (x^j(t, s_0+\varepsilon) - x^j(t, s_0)) + \underline{o}(\varepsilon) \right\} \mathrm{d}t,$$

$$\underline{J}^* = \int_{t_1}^{t_2} \left\{ \sum_{j=1}^{2} \frac{\partial \underline{f}(\underline{x}(t, s_0), \alpha_\varepsilon(t), t)}{\partial x^j} (x^j(t, s_0+\varepsilon) - x^j(t, s_0)) + \underline{o}(\varepsilon) \right\} \mathrm{d}t$$

In view of (A.1.3) we may now conclude that

$$\lim_{\varepsilon\to 0} \underline{J}/\varepsilon = \lim_{\varepsilon\to 0} \underline{J}^*/\varepsilon = \int_{t_1}^{t_2} \sum_{j=1}^{2} \frac{\partial \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)}{\partial x^j} \delta x^j(t)\,\mathrm{d}t \tag{A.1.7}$$

Concentrating on $\underline{I}$ and $\underline{I}^*$, we can conclude from (A.1.5 & 6) that

$$\delta\underline{x}(t_2) = \delta\underline{x}(t_1) + \lim_{\varepsilon\to 0} \underline{I}/\varepsilon + \lim_{\varepsilon\to 0} \underline{J}/\varepsilon \tag{A.1.5c}$$

$$\delta\underline{x}(t_2) = \delta\underline{x}(t_1) + \lim_{\varepsilon\to 0} \underline{I}^*/\varepsilon + \lim_{\varepsilon\to 0} \underline{J}^*/\varepsilon \tag{A.1.6c}$$

Considering (A.1.7) it can be asserted that

$$\lim_{\varepsilon\to 0} \underline{I}/\varepsilon = \lim_{\varepsilon\to 0} \underline{I}^*/\varepsilon = \underline{L}.$$

Using the $\underline{I}^*$-expression (A.1.6a), we can write

$$\underline{L} = \int_{t_1}^{t_2} \lim_{\varepsilon\to 0} \frac{\underline{f}(\underline{x}(t, s_0), \alpha_\varepsilon(t), t) - \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)}{\varepsilon}\,\mathrm{d}t$$

This expression justifies the conclusion that $\underline{L}$ is almost everywhere differentiable to the upper bound t_2. Omitting the index, we can now state that $\mathrm{d}\underline{L}/\mathrm{d}t$ exists almost everywhere on $[t_a, \tau]$.

From (A.1.5a) we see that

$$\frac{\mathrm{d}\underline{I}}{\mathrm{d}t} = \underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t) - \underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_0(t), t) \tag{A.1.8}$$

Remembering the maximum principle and the assumption that $\alpha_\varepsilon(t)$ was an optimal navigating function – because Γ_ε is an extremal – it is clear that

$$(\underline{\psi}_\varepsilon, \underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_\varepsilon(t), t)) \geqslant (\underline{\psi}_\varepsilon, \underline{f}(\underline{x}(t, s_0+\varepsilon), \alpha_0(t), t)),$$

where $\underline{\psi}_\varepsilon$ is a normal vector to $S^+(t)$ in $\underline{x}(t, s_0+\varepsilon)$.

So according to (A.1.8) we have

$$\left(\underline{\psi}_\varepsilon, \frac{\mathrm{d}\underline{I}}{\mathrm{d}t}\right) \geqslant 0 \tag{A.1.9}$$

so for

$$\varepsilon > 0: \left(\underline{\psi}_\varepsilon, \frac{\mathrm{d}\underline{I}/\varepsilon}{\mathrm{d}t}\right) \geqslant 0 \tag{A.1.10}$$

and for

$$\varepsilon < 0: \left(\underline{\psi}_\varepsilon, \frac{\mathrm{d}\underline{I}/\varepsilon}{\mathrm{d}t}\right) \leqslant 0 \tag{A.1.11}$$

Consideration of $\underline{I}^*$ in (A.1.6a) and using the maximum

principle on the extremal Γ_0, leads to the analogous conclusion that for

$$\varepsilon > 0: \left(\underline{\psi}_0, \frac{\mathrm{d}\underline{I}^*/\varepsilon}{\mathrm{d}t}\right) \leqslant 0 \qquad \text{(A.1.12)}$$

and for

$$\varepsilon < 0: \left(\underline{\psi}_0, \frac{\mathrm{d}\underline{I}^*/\varepsilon}{\mathrm{d}t}\right) \geqslant 0 \qquad \text{(A.1.13)}$$

Taking $\varepsilon > 0$ and letting ε descend to zero, which implies that $\underline{\psi}_\varepsilon$ tends to $\underline{\psi}_0$, it follows that

$$\left(\underline{\psi}_0, \frac{\mathrm{d}\underline{L}}{\mathrm{d}t}\right) \geqslant 0 \quad \text{according to (A.1.10)}$$

while

$$\left(\underline{\psi}_0, \frac{\mathrm{d}\underline{L}}{\mathrm{d}t}\right) \leqslant 0 \quad \text{in view of (A.1.12).}$$

So for $\varepsilon \downarrow 0$ we have $(\underline{\psi}_0, \mathrm{d}\underline{L}/\mathrm{d}t) = 0$ almost everywhere on $[t_a, \tau]$.

For ε approaching zero from the left, the arguments are analogous. (A.1.11) gives that $(\underline{\psi}_0, \mathrm{d}\underline{L}/\mathrm{d}t) \leqslant 0$ and (A.1.13) leads to $(\underline{\psi}_0, \mathrm{d}\underline{L}/\mathrm{d}t) \geqslant 0$. So the over-all-conclusion is that

$$(\underline{\psi}_0, \mathrm{d}\underline{L}/\mathrm{d}t) = 0 \qquad \text{(A.1.14)}$$

almost everywhere on $[t_a, \tau]$.

Let us now consider the expression (A.1.5c). Differentiation to t_2 and then omitting the index gives

$$\frac{\mathrm{d}(\delta\underline{x})}{\mathrm{d}t} = \frac{d}{\mathrm{d}t}(\lim_{\varepsilon \to 0} \underline{J}/\varepsilon) + \frac{\mathrm{d}\underline{L}}{\mathrm{d}t}.$$

In regard of (A.1.7) this becomes

$$\frac{\mathrm{d}(\delta\underline{x})}{\mathrm{d}t} = \sum_{j=1}^{2} \frac{\partial \underline{f}(\underline{x}(t, s_0), \alpha_0(t), t)}{\partial x^j} \delta x^j + \frac{\mathrm{d}\underline{L}}{\mathrm{d}t} \qquad \text{(A.1.15)}$$

For the gradient vector $\underline{\psi}(t)$, which is to be permanently perpendicular to a timefront, the relation

$$(\underline{\psi}(t), \delta\underline{x}(t)) = 0$$

must hold for all t past t_a, so it has to satisfy the requirement

$$\left(\frac{\mathrm{d}\underline{\psi}}{\mathrm{d}t}, \delta\underline{x}(t)\right) + \left(\underline{\psi}(t), \frac{\mathrm{d}(\delta\underline{x}(t))}{\mathrm{d}t}\right) = 0.$$

Working this out and taking account of (A.1.14 & 15), we find the components ψ_i of $\underline{\psi}$ to change along an extremal according to

$$\frac{\mathrm{d}\psi_i}{\mathrm{d}t} = -\sum_{j=1}^{2} \frac{\partial f^j(\underline{x}(t), \alpha(t), t)}{\partial x^i} \psi_j \qquad (i = 1, 2)$$

almost everywhere for $t \geqslant t_a$.

Summarizing this, we can say that the construction of an extremal amounts to the solution of the following system of differential equations:

For $i = 1, 2$

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = f^i(x^1, x^2, \alpha, t)$$

$$\frac{\mathrm{d}\psi_i}{\mathrm{d}t} = -\sum_{j=1}^{2} \frac{\partial f^j(x^1, x^2, \alpha, t)}{\partial x^i} \psi_j$$

Select α within the limits of feasibility so, that it maximizes the inner product

$$\sum_{j=1}^{2} \psi_j(t) f^j(x^1, x^2, \alpha, t).$$

The starting conditions for x^i are: $x^i(t_a) = x_a{}^i$.

For the adjoint vector $\underline{\psi}$ no start is given, but the starting values of ψ_i have to be so that the above mentioned system generates a trajectory, that satisfies the requirements $x^i(t_b) = x_b{}^i$ for some, yet unknown time t_b.

APPENDIX 2

Modifications in case of co-ordinate restrictions

Let Γ_0 in figure A.2a be an extremal, starting from A at a time t_a and let Γ_0 touch the boundary at a point R, with co-ordinates (x_0^1, x_0^2), at the time t_0. Let the boundary near R be specified by the equation $\varphi(x^1,x^2) = 0$, while points of G have to satisfy the condition $\varphi(x^1,x^2) \leqslant 0$. Suppose the boundary curve near R to be smooth, so that $\varphi(x^1,x^2)$ is continuously differentiable with respect to x^1 and x^2. We now want to investigate, what conditions have to be satisfied for an arc RS of the boundary to be a "sub-extremal", i.e. the boundary arc RS is a minimal time curve, compared to other curves from R to S with no points outside of G.

Otherwise stated, we start at a time t_0 with $x^i(t_0) = x_0^i$ $(i = 1,2)$, while $\varphi(x_0^1, x_0^2) = 0$. We navigate according to

$$\frac{dx^i}{dt} = f^i(x^1, x^2, \alpha, t),$$

with the position co-ordinates satisfying $\varphi(x^1(t), x^2(t)) \leqslant 0$ for $t > t_0$. The destination S has co-ordinates (x_1^1, x_2^2), for which the equality $\varphi(x_1^1, x_1^2) = 0$ holds.

We also assume there is a class F of navigating functions $\{\alpha(t)\}$ that all take the ship from R to S, arriving there at a time t_1.

(N.B. This arrival time t_1 depends on the choice of the navigating function $\alpha(t)$).

Let the steering function $\tilde{\alpha}(t)$ be optimal, so the corresponding arrival time $\tilde{t}_1$ in S is minimal. If the boundary arc RS is a sub-extremal, then the solutions $\tilde{x}^i(t)$ of the above stated problem satisfy the equality $\varphi(\tilde{x}^1(t), \tilde{x}^2(t)) = 0$ for $t_0 \leqslant t \leqslant t_1$.

If $\tilde{\alpha}(t)$ is given a small variation $\delta\alpha(t)$, while $\tilde{\alpha}(t)+\delta\alpha(t)$ still belongs to class F, this varied navigator will generate the solutions $x^i = \tilde{x}^i(t)+\delta x^i(t)$ and an arrival time $\tilde{t}_1+\delta t_1$, with $\delta t_1 \geqslant 0$, will be the result. Substituting

$$\begin{array}{llll} x^1 = x, & x_0^1 = x_0, & x_1^1 = x_1, \\ x^2 = y^1, & x_0^2 = y_0^1, & x_1^2 = y_1^1, \\ t = y^2, & t_0 = y_0^2, & t^1 = y_1^2, \\ 1/f^1 = G^2 & f^2/f^1 = G^1, & \end{array}$$

we have for $i = 1,2$:

1. The steering equations

$$\frac{dy^i}{dx} = G^i(y^1, y^2, \alpha, x) \qquad \text{(A.2.1)}$$

2. The start and end conditions

$$y^i(x_0) = y_0^i, \quad y^i(x_1) = y_1^i,$$

where y_0^1, y_0^2 and y_1^1 are fixed and given, while the arrival time y_1^2 is variable.

3. The restriction

$$\varphi(x, y^1) \leqslant 0 \quad \text{for all} \quad y^2, \qquad \text{(A.2.2)}$$

while $\varphi(x, y_k^1) = 0 \quad \text{for} \quad k = 0,1.$

The time that it takes the ship to go from R to S is equal to

$$J(\underline{y}) = y_1^2 - y_0^2 = \int_{x_0}^{x_1} G^2(y^1, y^2, \alpha, x)\, dx \qquad \text{(A.2.3)}$$

Denoting the minimal solutions by $y^i = \tilde{y}^i(x)$ and introducing the abbreviative notations

$$\frac{\partial G^i(\tilde{y}^1, \tilde{y}^2, \tilde{\alpha}, x)}{\partial y^k} = G^i_{y^k} \text{ and } \frac{\partial G^i(\tilde{y}^1, \tilde{y}^2, \tilde{\alpha}, x)}{\partial \alpha} = G^i_\alpha,$$

the variations δy^i and $\delta\alpha$ have to satisfy the minimal condition

$$J(\underline{\tilde{y}}) \leqslant J(\underline{\tilde{y}} + \delta\underline{y}),$$

so

$$\int_{x_0}^{x_1} \{G^2(\underline{\tilde{y}} + \delta\underline{y}, \tilde{\alpha} + \delta\alpha, x) - G^2(\underline{\tilde{y}}, \tilde{\alpha}, x)\}\, dx \geqslant 0.$$

This means that the requirement

$$\int_{x_0}^{x_1} \left\{ \sum_{i=1}^{2} G^2_{y^i} \delta y^i + G^2_\alpha \delta\alpha \right\} dx \geqslant 0 \qquad \text{(A.2.4)}$$

has to be met with.

The x-derivatives of δy^i can be found to be

$$\frac{d(\delta y^i)}{dx} = \sum_{j=1}^{2} G^i_{y^j} \delta y^j + G^i_\alpha \delta\alpha \qquad \text{(A.2.5)}$$

The restriction to G, implying the inequality

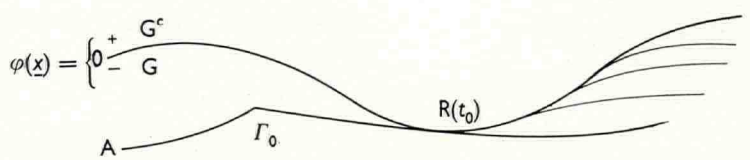


Fig. A.2a.

$\varphi(x,y^1) \leqslant 0$, can also be expressed by introducing a slack variable z, that has to be non-negative. The inequality can then be replaced by the equality

$$\varphi(x,y^1)+z = 0 \qquad \text{(A.2.6)}$$

where it can be remarked that $z = 0$ for $y^1 = \tilde{y}^1(x)$. The variations in $\tilde{y}^i(x)$ imply slack variations $\delta z(x)$, so variation of (A.2.6) gives

$$\varphi_{y^1}\delta y^1+\delta z = 0 \qquad \text{(A.2.7)}$$

Differentiating this to x and using (A.2.5) gives

$$\varphi_{xy^1}\delta y^1+\varphi_{y^1y^1}G^1\delta y^1+$$
$$+\sum_{j=1}^{2}\varphi_{y^1}G^1_{y^j}\delta y^j+\varphi_{y^1}G^1_\alpha\delta\alpha+\delta\left(\frac{dz}{dx}\right) = 0$$

Putting

$$\varrho = \varphi_{y^1}G^1_\alpha, \quad \sigma = 1/\varrho,$$

$$\chi_1 = \varphi_{xy^1}+\varphi_{y^1y^1}G^1+\varphi_{y^1}G^1_{y^1}, \quad \chi_2 = \varphi_{y^1}G^1_{y^2},$$

the last equation becomes

$$\sum_{j=1}^{2}\chi_j\delta y^j+\varrho\delta\alpha+\delta\left(\frac{dz}{dx}\right) = 0$$

or

$$\delta\alpha = -\sigma\sum_{j=1}^{2}\chi_j\delta y^j-\sigma\delta\left(\frac{dz}{dx}\right)$$

Considering the minimal time condition (A.2.4), we can now substitute for $\delta\alpha$ and find

$$\int_{x_0}^{x_1}\left\{\sum_{j=1}^{2}(G^2_{y^j}-\sigma G^2_\alpha\chi_j)\delta y^j-\sigma G^2_\alpha\delta\left(\frac{dz}{dx}\right)\right\}dx \geqslant 0 \qquad \text{(A.2.8)}$$

We now introduce the adjoint vector $p = (p_1,p_2)$ by defining

$$G^2_{y^i}-\sigma G^2_\alpha\chi_i = \frac{dp_i}{dx}+\sum_{j=1}^{2}p_j(G^j_{y^i}-G^j_\alpha\sigma\chi_i)$$

Before further development, let us consider this

$$\frac{d}{dx}\left\{\sum_{j=1}^{2}p_j\delta y^j\right\} = \sum_{j=1}^{2}\left\{\frac{dp_j}{dx}\delta y^j+p_j\frac{d(\delta y^j)}{dx}\right\}$$
$$= \sum_{j=1}^{2}\left\{\left(\frac{dp_j}{dx}+\sum_{k=1}^{2}p_kG^k_{y^j}\right)\delta y^j+p_jG^j_\alpha\delta\alpha\right\}$$
$$= \sum_{j=1}^{2}\left\{\left(\frac{dp_j}{dx}+\sum_{k=1}^{2}p_kG^k_{y^j}\right)\delta y^j+\right.$$
$$\left.+p_jG^j_\alpha\left(-\sigma\sum_{k=1}^{2}\chi_k\delta y^k-\sigma\delta\left(\frac{dz}{dx}\right)\right)\right\}$$

With

$$\lambda = \sum_{j=1}^{2}p_jG^j_\alpha\sigma$$

this becomes

$$\frac{d}{dx}\left(\sum_{j=1}^{2}p_j\delta y^j\right) = \sum_{j=1}^{2}\left(\frac{dp_j}{dx}+\sum_{k=1}^{2}p_kG^k_{y^j}-\lambda\chi_j\right)\delta y^j-$$
$$-\lambda\delta\left(\frac{dz}{dx}\right)$$

Now the inequality (A.2.8) can be replaced by

$$\int_{x_0}^{x_1}\left[\frac{d}{dx}\sum_{j=1}^{2}(p_j\delta y^j)+(\lambda-\sigma G^2_\alpha)\delta\left(\frac{dz}{dx}\right)\right]dx \geqslant 0,$$

or

$$\left[\sum_{j=1}^{2}p_j\delta y^j\right]_{x_0}^{x_1}+\int_{x_0}^{x_1}\mu\delta\left(\frac{dz}{dx}\right)dx \geqslant 0, \quad \text{with} \quad \mu = \lambda-\sigma G^2_\alpha$$

Remembering, that $\delta y^i(x_0) = 0$ and $\delta y^1(x_1) = 0$, this leads to

$$p_2\delta y^2(x_1)+\mu\delta z-\int_{x_0}^{x_1}\frac{d\mu}{dx}\delta z\,dx \geqslant 0$$

The requirements $\delta y^2(x_1) = \delta t_1 \geqslant 0$ and $\delta z \geqslant 0$ now mean that $p_2 \geqslant 0$, $\mu \geqslant 0$ and $d\mu/dx \leqslant 0$.

The last conclusion means, that the "boundary extremal", for which μ is positive, tends to a free extremal. As soon as μ becomes zero, the boundary condition has lost its weight and the extremal can be constructed without having to consider the $\varphi \leqslant 0$-restriction for the time being.

Summarizing the result, the variables y^i, α, μ and p_i have to be solved from the following equations

$$\frac{dy^i}{dx} = G^i(y^1,y^2,\alpha,x)$$

$$\frac{dp_i}{dx} = -\sum_{j=1}^{2}p_jG^j_{y^i}+\mu\chi_i+G^2_{y^i}$$

$$\varphi(x,y^1) = 0$$

$$G^2_\alpha+\mu\varphi_{y^1}G^1_\alpha-\sum_{j=1}^{2}p_jG^j_\alpha = 0$$

Transforming this back to a system with the time as independent variable and putting

$$\psi_2 = p_1$$

$$\psi_1 = -(\psi_2f^2+p_2-1)/f^1$$

one can find

$$\frac{dx^i}{dt} = f^i(x^1, x^2, \alpha, t)$$

$$\frac{d\psi_i}{dt} = \sum_{j=2}^{2} \{\psi_j f^j_{x^i} + \mu(\varphi_{x^j} f^j_{x^i} + \varphi_{x^i x^j} f^j)$$

$$\varphi(x^1, x^2) = 0$$

$$\sum_{j=1}^{2} (\psi_j - \mu\varphi_{x^j}) f^j_\alpha = 0$$

The last expression determines μ. This can be geometrically explained by means of figure A.2b.

In a point P of the boundary, characterized by $\varphi(x^1, x^2) = 0$, the stream vector $\underline{c}$ is drawn. The mean wae vdirection and the significant wave height determine the ship's velocity indicatrix.

The gradient vector $\nabla\underline{\varphi}$ is pointed to the outside of G.

The adjoint vector $\underline{\psi}$ has a direction, so that bare application of the maximum principle would produce a velocity vector $\underline{f}$, that has a positive inner product

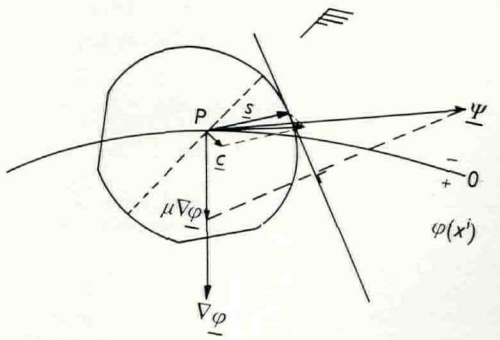


Fig. A.2b.

with $\nabla\underline{\varphi}$. This would mean a tresspassing into the prohibited region G^c. Subtraction of $\mu\nabla\underline{\varphi}$ from $\underline{\psi}$ – with $\mu > 0$ – has to deliver a velocity $\underline{f}$ perpendicular to $\nabla\underline{\varphi}$. This number μ can be found as follows:

Determine the ship's velocity $\underline{s}$ so that $\underline{c}+\underline{s}$ touches the boundary.

In the endpoint of $\underline{s}$ draw the tangent to the indicatrix. Now the vectors $\underline{\psi}$ and $\mu\nabla\underline{\varphi}$ must have equal projections on this tangent. This determines μ, as long as it is positive. As soon as μ becomes zero, the extremal leaves the boundary and leads to points inside G, with $\varphi(x^1, x^2) < 0$.

PUBLICATIONS OF THE NETHERLANDS SHIP RESEARCH CENTRE TNO

PUBLISHED AFTER 1963 (LIST OF EARLIER PUBLICATIONS AVAILABLE ON REQUEST)

PRICE PER COPY DFL. 10,—

M = engineering department S = shipbuilding department C = corrosion and antifouling department

Reports

57 M Determination of the dynamic properties and propeller excited vibrations of a special ship stern arrangement. R. Wereldsma, 1964.

58 S Numerical calculation of vertical hull vibrations of ships by discretizing the vibration system, J. de Vries, 1964.

59 M Controllable pitch propellers, their suitability and economy for large sea-going ships propelled by conventional, directly coupled engines. C. Kapsenberg, 1964.

60 S Natural frequencies of free vertical ship vibrations. C. B. Vreugdenhil, 1964.

61 S The distribution of the hydrodynamic forces on a heaving and pitching shipmodel in still water. J. Gerritsma and W. Beukelman, 1964.

62 C The mode of action of anti-fouling paints: Interaction between anti-fouling paints and sea water. A. M. van Londen, 1964.

63 M Corrosion in exhaust driven turbochargers on marine diesel engines using heavy fuels. R. W. Stuart Mitchell and V. A. Ogale, 1965.

64 C Barnacle fouling on aged anti-fouling paints; a survey of pertinent literature and some recent observations. P. de Wolf, 1964.

65 S The lateral damping and added mass of a horizontally oscillating shipmodel. G. van Leeuwen, 1964.

66 S Investigations into the strength of ships' derricks. Part I. F. X. P. Soejadi, 1965.

67 S Heat-transfer in cargotanks of a 50,000 DWT tanker. D. J. van der Heeden and L. L. Mulder, 1965.

68 M Guide to the application of method for calculation of cylinder liner temperatures in diesel engines. H. W. van Tijen, 1965.

69 M Stress measurements on a propeller model for a 42,000 DWT tanker. R. Wereldsma, 1965.

70 M Experiments on vibrating propeller models. R. Wereldsma, 1965.

71 S Research on bulbous bow ships. Part II. A. Still water performance of a 24,000 DWT bulkcarrier with a large bulbous bow. W. P. A. van Lammeren and J. J. Muntjewerf, 1965.

72 S Research on bulbous bow ships. Part II. B. Behaviour of a 24,000 DWT bulkcarrier with a large bulbous bow in a seaway. W. P. A. van Lammeren and F. V. A. Pangalila, 1965.

73 S Stress and strain distribution in a vertically corrugated bulkhead. H. E. Jaeger and P. A. van Katwijk, 1965.

74 S Research on bulbous bow ships. Part I. A. Still water investigations into bulbous bow forms for a fast cargo liner. W. P. A. van Lammeren and R. Wahab, 1965.

75 S Hull vibrations of the cargo-passenger motor ship "Oranje Nassau", W. van Horssen, 1965.

76 S Research on bulbous bow ships. Part I. B. The behaviour of a fast cargo liner with a conventional and with a bulbous bow in a seaway. R. Wahab, 1965.

77 M Comparative shipboard measurements of surface temperatures and surface corrosion in air cooled and water cooled turbine outlet casings of exhaust driven marine diesel engine turbochargers. R. W. Stuart Mitchell and V. A. Ogale, 1965.

78 M Stern tube vibration measurements of a cargo ship with special afterbody. R. Wereldsma, 1965.

79 C The pre-treatment of ship plates: A comparative investigation on some pre-treatment methods in use in the shipbuilding industry. A. M. van Londen, 1965.

80 C The pre-treatment of ship plates: A practical investigation into the influence of different working procedures in over-coating zinc rich epoxy-resin based pre-construction primers. A. M. van Londen and W. Mulder, 1965.

81 S The performance of U-tanks as a passive anti-rolling device. C. Stigter, 1966.

82 S Low-cycle fatigue of steel structures. J. J. W. Nibbering and J. van Lint, 1966.

83 S Roll damping by free surface tanks. J. J. van den Bosch and J. H. Vugts, 1966.

84 S Behaviour of a ship in a seaway. J. Gerritsma, 1966.

85 S Brittle fracture of full scale structures damaged by fatigue. J. J. W. Nibbering, J. van Lint and R. T. van Leeuwen, 1966.

86 M Theoretical evaluation of heat transfer in dry cargo ship's tanks using thermal oil as a heat transfer medium. D. J. van der Heeden, 1966.

87 S Model experiments on sound transmission from engineroom to accommodation in motorships. J. H. Janssen, 1966.

88 S Pitch and heave with fixed and controlled bow fins. J. H. Vugts, 1966.

89 S Estimation of the natural frequencies of a ship's double bottom by means of a sandwich theory. S. Hylarides, 1967.

90 S Computation of pitch and heave motions for arbitrary ship forms. W. E. Smith, 1967.

91 M Corrosion in exhaust driven turbochargers on marine diesel engines using heavy fuels. R. W. Stuart Mitchell, A. J. M. S. van Montfoort and V. A. Ogale, 1967.

92 M Residual fuel treatment on board ship. Part II. Comparative cylinder wear measurements on a laboratory diesel engine using filtered or centrifuged residual fuel. A. de Mooy, M. Verwoest and G. G. van der Meulen, 1967.

93 C Cost relations of the treatments of ship hulls and the fuel consumption of ships. H. J. Lageveen-van Kuijk, 1967.

94 C Optimum conditions for blast cleaning of steel plate. J. Remmelts, 1967.

95 M Residual fuel treatment on board ship. Part I. The effect of centrifuging, filtering and homogenizing on the unsolubles in residual fuel. M. Verwoest and F. J. Colon, 1967.

96 S Analysis of the modified strip theory for the calculation of ship motions and wave bending moments. J. Gerritsma and W. Beukelman, 1967.

97 S On the efficacy of two different roll-damping tanks. J. Bootsma and J. J. van den Bosch, 1967.

98 S Equation of motion coefficients for a pitching and heaving destroyer model. W. E. Smith, 1967.

99 S The manoeuvrability of ships on a straight course. J. P. Hooft, 1967.

100 S Amidships forces and moments on a $C_B = 0.80$ "Series 60" model in waves from various directions. R. Wahab, 1967.

101 C Optimum conditions for blast cleaning of steel plate. Conclusion. J. Remmelts, 1967.

102 M The axial stiffness of marine diesel engine crankshafts. Part I. Comparison between the results of full scale measurements and those of calculations according to published formulae. N. J. Visser, 1967.

103 M The axial stiffness of marine diesel engine crankshafts. Part II. Theory and results of scale model measurements and comparison with published formulae. C. A. M. van der Linden, 1967.

104 M Marine diesel engine exhaust noise. Part I. A mathematical model. J. H. Janssen, 1967.

105 M Marine diesel engine exhaust noise. Part II. Scale models of exhaust systems. J. Buiten and J. H. Janssen, 1968.

106 M Marine diesel engine exhaust noise. Part III. Exhaust sound criteria for bridge wings. J. H. Janssen en J. Buiten, 1967.

107 S Ship vibration analysis by finite element technique. Part I. General review and application to simple structures, statically loaded. S. Hylarides, 1967.

108 M Marine refrigeration engineering. Part I. Testing of a decentralised refrigerating installation. J. A. Knobbout and R. W. J. Kouffeld, 1967.

109 S A comparative study on four different passive roll damping tanks. Part I. J. H. Vugts, 1968.

110 S Strain, stress and flexure of two corrugated and one plane bulkhead subjected to a lateral, distributed load. H. E. Jaeger and P. A. van Katwijk, 1968.

111 M Experimental evaluation of heat transfer in a dry-cargo ships' tank, using thermal oil as a heat transfer medium. D. J. van der Heeden, 1968.

112 S The hydrodynamic coefficients for swaying, heaving and rolling cylinders in a free surface. J. H. Vugts, 1968.

113 M Marine refrigeration engineering. Part II. Some results of testing a decentralised marine refrigerating unit with R 502. J. A. Knobbout and C. B. Colenbrander, 1968.

114 S The steering of a ship during the stopping manoeuvre. J. P. Hooft, 1969.

115 S Cylinder motions in beam waves. J. H. Vugts, 1968.

116 M Torsional-axial vibrations of a ship's propulsion system. Part I. Comparative investigation of calculated and measured torsional–axial vibrations in the shafting of a dry cargo motorship. C. A. M. van der Linden, H. H. 't Hart and E. R. Dolfin, 1968.
117 S A comparative study on four different passive roll damping tanks. Part II. J. H. Vugts, 1969.
118 M Stern gear arrangement and electric power generation in ships propelled by controllable pitch propellers. C. Kapsenberg, 1968.
119 M Marine diesel engine exhaust noise. Part IV. Transferdamping data of 40 modelvariants of a compound resonator silencer. J. Buiten, M. J. A. M. de Regt and W. P. H. Hanen, 1968.
120 C Durability tests with prefabrication primers in use of steel plates. A. M. van Londen and W. Mulder, 1970.
121 S Proposal for the testing of weld metal from the viewpoint of brittle fracture initiation. W. P. van den Blink and J. J. W. Nibbering, 1968.
122 M The corrosion behaviour of cunifer 10 alloys in seawaterpiping-systems on board ship. Part I. W. J. J. Goetzee and F. J. Kievits, 1968.
123 M Marine refrigeration engineering. Part III. Proposal for a specification of a marine refrigerating unit and test procedures. J. A. Knobbout and R. W. J. Kouffeld, 1968.
124 S The design of U-tanks for roll damping of ships. J. D. van den Bunt, 1969.
125 S A proposal on noise criteria for sea-going ships. J. Buiten, 1969.
126 S A proposal for standardized measurements and annoyance rating of simultaneous noise and vibration in ships. J. H. Janssen, 1969.
127 S The braking of large vessels II. H. E. Jaeger in collaboration with M. Jourdain, 1969.
128 M Guide for the calculation of heating capacity and heating coils for double bottom fuel oil tanks in dry cargo ships. D. J. van der Heeden, 1969.
129 M Residual fuel treatment on board ship. Part III. A. de Mooy, P. J. Brandenburg and G. G. van der Meulen, 1969.
130 M Marine diesel engine exhaust noise. Part V. Investigation of a double resonatorsilencer. J. Buiten, 1969.
131 S Model and full scale motions of a twin-hull vessel. M. F. van Sluijs, 1969.
132 M Torsional-axial vibrations of a ship's propulsion system. Part II. W. van Gent and S. Hylarides, 1969.
134 M The corrosion behaviour of cunifer-10 alloys in seawaterpiping-systems on board ship. Part II. P. J. Berg and R. G. de Lange, 1969.
135 S Boundary layer control on a ship's rudder. J. H. G. Verhagen, 1970.
137 M Torsional-axial vibrations of a ship's propulsion system. Part III. C. A. M. van der Linden, 1969.
138 S The manoeuvrability of ships at low speed. J. P. Hooft and M. W. C. Oosterveld, 1970.
141 S Resistance and propulsion of a high-speed single-screw cargo liner design. J. J. Muntjewerf, 1970.
142 S Optimal meteorological ship routeing. C. de Wit, 1970.

Communications

11 C Investigations into the use of some shipbottom paints, based on scarcely saponifiable vehicles (Dutch). A. M. van Londen and P. de Wolf, 1964.
12 C The pre-treatment of ship plates: The treatment of welded joints prior to painting (Dutch). A. M. van Londen and W. Mulder, 1965.
13 C Corrosion, ship bottom paints (Dutch). H. C. Ekama, 1966.
14 S Human reaction to shipboard vibration, a study of existing literature (Dutch). W. ten Cate, 1966.
15 M Refrigerated containerized transport (Dutch). J. A. Knobbout, 1967.
16 S Measures to prevent sound and vibration annoyance aboard a seagoing passenger and carferry, fitted out with dieselengines (Dutch). J. Buiten, J. H. Janssen, H. F. Steenhoek and L. A. S. Hageman, 1968.
17 S Guide for the specification, testing and inspection of glass reinforced polyester structures in shipbuilding (Dutch). G. Hamm, 1968.
18 S An experimental simulator for the manoeuvring of surface ships. J. B. van den Brug and W. A. Wagenaar, 1969.
19 S The computer programmes system and the NALS language for numerical control for shipbuilding. H. le Grand, 1969.
20 S A case study on networkplanning in shipbuilding (Dutch). J. S. Folkers, H. J. de Ruiter, A. W. Ruys, 1970.