# POLISH MARITIME RESEARCH
No 2(69)   2011   Vol 18

## CONTENTS

# Editorial

POLISH MARITIME RESEARCH is a scientific journal of worldwide circulation. The journal appears as a quarterly four times a year. The first issue of it was published in September 1994. Its main aim is to present original, innovative scientific ideas and Research & Development achievements in the field of :

**Engineering**, **Computing & Technology**, **Mechanical Engineering**,

which could find applications in the broad domain of maritime economy. Hence there are published papers which concern methods of the designing, manufacturing and operating processes of such technical objects and devices as : ships, port equipment, ocean engineering units, underwater vehicles and equipment as well as harbour facilities, with accounting for marine environment protection.

The Editors of POLISH MARITIME RESEARCH make also efforts to present problems dealing with education of engineers and scientific and teaching personnel. As a rule, the basic papers are supplemented by information on conferences , important scientific events as well as cooperation in carrying out international scientific research projects.

# Scientific Board

# Multi-objective optimization of high speed vehicle-passenger catamaran by genetic algorithm

## Part I
## Theoretical background on evolutionary multi-objective optimization

**Zbigniew Sekulski**, Ph. D.
West Pomeranian University of Technology in Szczecin

## ABSTRACT

*Real ship structural design problems are usually characterized by presence of many conflicting objectives. Simultaneously, a complete definition of the optimal structural design requires a formulation of size-topology-shape-material optimization task unifying the optimization problems from these four areas and giving an effective solution of this problem. So far, a significant progress towards the solution of this problem has not been obtained. An objective of the present paper was to develop an evolutionary algorithm for multi-objective optimization of the structural elements of the large spatial sections of ships. Selected elements of the multi-criteria optimization theory have been presented in details. Methods for solution of the multi-criteria optimization problems have been discussed with the focus on the evolutionary optimization algorithms. In the paper an evolutionary algorithm where selection takes place based on the aggregated objective function combined with domination attributes as well as distance to the asymptotic solution is proposed and applied to solve the problem of optimizing structural elements with respect to their weight and surface area on a high speed vehicle-passenger catamaran structure with several design variables, such as plate thickness, scantlings of longitudinal stiffeners and transverse frames, and spacing between longitudinals and transversal members. Details of the computational models were at the level typical for conceptual design. Scantlings were analyzed using the selected rules of a classification society. The results of numerical experiments with the use of the developed algorithm are presented. They show that the proposed genetic algorithm can be an efficient multi-objective optimization tool for ship structures optimization.*
*The paper will be published in three parts: Part I: Theoretical background on evolutionary multi-objective optimization, Part II: Computational investigations, and Part III: Analysis of the results.*

**Keywords**: ship structure; multi-objective optimization; evolutionary algorithm; genetic algorithm; Pareto domination; set of non-dominated solutions

## INTRODUCTION

Hull structures are fundamental load-carrying members of modern waterborne vessels integrating the whole structural system, similarly to air, railway and other transportation vessels. Except for specific solutions of the particular types of ships (e.g. bulk carrier, container carrier, ro-ro carrier), the hull of each merchant ship is constructed according to the common, basic design that involves a thin, tightly-closed, watertight coating stiffened by a set of orthogonally intersecting stiffeners and series of bulkheads, Fig. 1.

Fore part of the ship hull is referred to as a bow or fore peak, aft part – a stern or after peak. The lowest part of the hull is called a bottom, side walls – sides, and the top part closing the hull – a deck. Consistently the bottom part of hull

shell is referred to as a bottom shell, side part – side shell, deck part – deck shell. Structural components holding the watertight shell are referred to as framing. Hull framing is composed of intersecting transverse and longitudinal beams referred to as members joined with bottom, side and deck plating. Beams oriented across form transverse framing, while those oriented alongside form longitudinal framing. Vertical watertight bulkheads, dividing space limited by hull outer coating to series of sections and compartments, are referred to as watertight bulkheads: transverse and longitudinal. First ones are set in perpendicular planes in bow-stern direction and they divide hull inner space onto several compartments along ship. Longitudinal bulkheads, set parallel to that direction, divide inner volume to series of compartments in ship breadth. Watertight bulkheads are

situated between bottom and upper deck ensuring floatability in case of emergency flooding of specific compartments. Apart from vertical diaphragms or bulkheads there are horizontal diaphragms assembled inside the shell referred to as tweendecks or platforms, depending on spatial range and tightness.
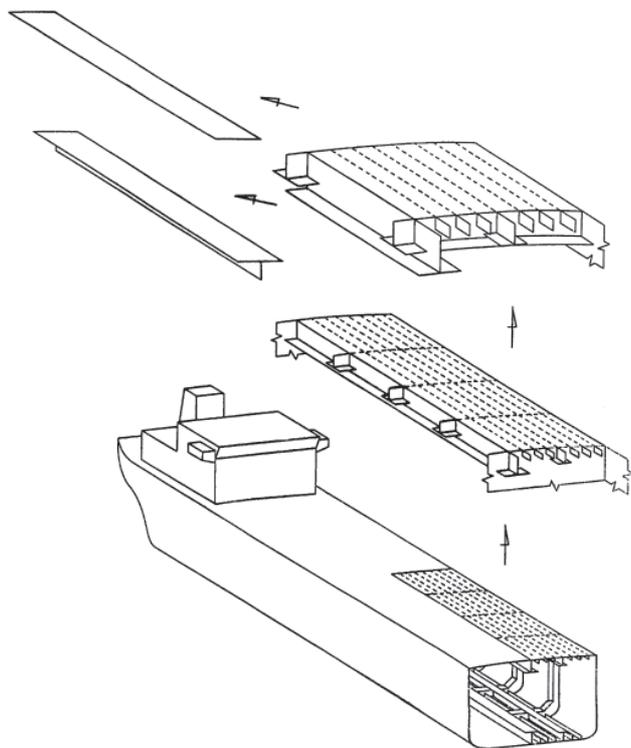


*Fig. 1. Structural members of ship hull structure*

Structural members are divided into two basic groups. Structural members made from rolled profiles[1] supporting only shell plating elements are referred to as secondary structural members or stiffeners. Structural members, usually of large size, prefabricated from elements carved from steel plates, supporting shell plating and secondary structural members (stiffeners) are referred to as primary structural members or frames (perpendicular oriented) or girders (longitudinal oriented). Stiffeners supporting bottom shell plating are referred to as bottom frames, whereas transverse oriented bottom primary members are referred to as floors and longitudinally oriented – bottom girders. Stiffeners supporting side shell plating are referred to as transverse or longitudinal side frames – depending on orientation. Transverse primary side members are referred to as web frames, those oriented longitudinally – side girders. Stiffeners supporting a deck are referred to as beams, transverse or longitudinal depending on the orientation, while primary members are referred to as deck transverse, if transverse, or deck girders, if running along the hull. Floors, web frames, and deck transverse are situated in the same intersections (planes) and stiffly joined form main frame.

Transversally oriented ship stiffeners are spaced in regular distances referred to as frame spacing. The frame spacing is determined in the hull structure design process and fundamentally influences its strength and operational characteristics as a whole.

Since the hull of the seagoing ship (1) is the main factor ensuring safety and structural strength, (2) has a very large weight, (3) imposes the framework for internal space arrangement and (4) dictates the basic construction strategy, ship structural design is a key factor in the ship design process. Ship hull structural design always begins with defining design objective. One is supposed to completely and fully specify tasks to fulfill by the structure being designed. It is very rare that the solution of the design process is strictly specified and no variations from initial assumptions are allowed. In most cases precise expression of tasks made in subsequent phase may influence changes in initial concept developed in foregoing phase. Afterwards the actual structural design process begins in which a proper set of component structures are selected and it is specified how they should act jointly. The structural design being developed is subject to continuous qualitative and quantitative analyses. As a result topological description of useful solutions and dimensions of particular component structures are obtained. In this way the ship hull structure is defined in all details. After an operation that is made up from certain number of iteration steps (where loops of these steps differ) a complete structural design is received as a result. After that it remains to expand the information concerning the use of proper technology. It is also needed to confirm that ship hull structure properties meet the requirements specified at the beginning, e.g. with use of numerical calculations or laboratory research. On the other hand detailed design documentation components must be prepared, according to witch the structure is to be realized.

The considered entire process of designing a ship structure is characterized by repeatability of different phases. That repeatability cannot be found in simple sequential form. For example, the analysis of solution concept can influence the task specification, which as a result, may undergo expanding or restriction. Also the design objective can change, e.g. by expanding the range of structure requirements. Structure design may be considered finished if it is presented and accepted by the builder and ordered body. Therefore designing and constructing the whole ship and its hull courses in phases and is a sequent-iterative process. Very often this process is represented in so referred to as "Design spiral" form, Fig. 2.

Thus seagoing ship design process has the sequent-iterative course including concept, structural and detailed design where many operations are repeated with simultaneous increase in work detail and gradual approaching from initial assumptions and basic structure ideas to detail design of structural joints. It is also a hierarchic process with descending structure from more general requirements to detailed design of structural regions. In most general form structure designing is a choice of material and its spatial arrangement in form of structure components (decks, bulkheads, hull sections), which are made of smaller structural elements (stiffeners, plating etc.), Fig. 1. It is supposed to be performed in such a way, that with fulfilled assumed design requirements the hull will ensure safe ship operation at the least costs.

Decisions fundamental for achieving design objectives are taken at the preliminary design phase because at this phase basic ship characteristics are being decided and it is here that the basic risk of making fundamental errors in designing the vessel. Errors made at this phase will influence utilitarian and operational characteristics and actually will not be able to be compensated in later optimization of details on lower hierarchic levels of structure. Not removed will cause the highest losses. At the same time quick and accurate decisions bring the highest benefits and decisively influence total success of the enterprise that is bound with investing substantial amount of money in building a new ship. In this paper it was assumed

---

[1]   Steel profiles are produced in rolling process, while profiles made from aluminum alloys are produced through extrusion.
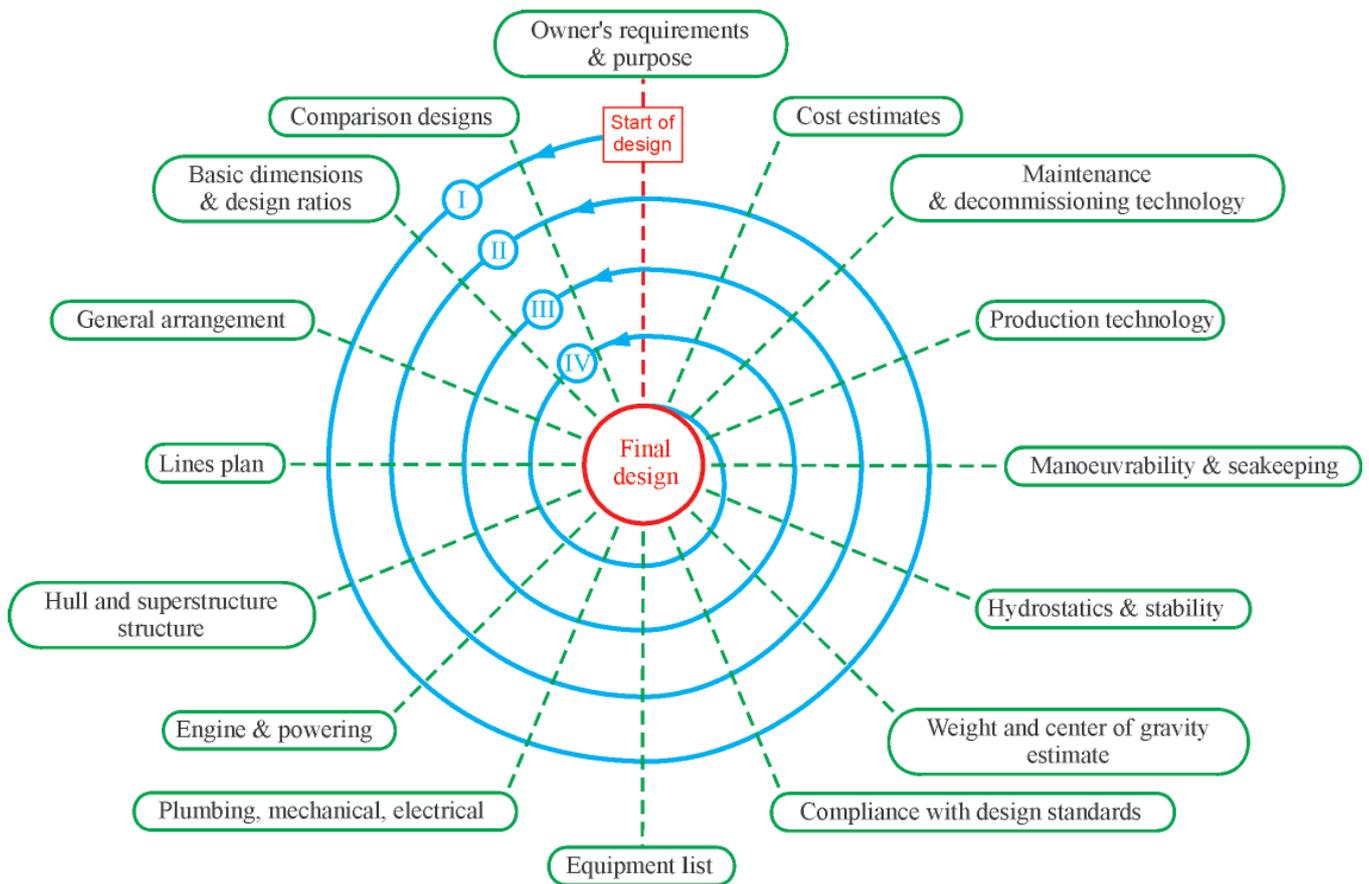
**Fig. 2.** *Ship design as the process of achieving the next design levels in several approximations - ship design spiral; designation of ship design phases:*
*I) concept design, II) preliminary design, III) contract design, IV) basic design, V) detailed engineering design,*
*VI) as fitted documentation or delivery documentation*

that the research will be conducted in a field proper for ship preliminary design.

Considering dimensioning of structural elements, the basis for the structural design are either requirements of classification societies (e.g. Det Norske Veritas, Lloyd's Register, Germanischer Lloyd) published as classification rules, containing simplified formulas for evaluating structural loads and dimensioning structural elements resisting these loads, or more rational approach based directly and fully on the structural mechanics. Regardless of occasionally formulated criticism of the classification rules as a tool not meeting the designers expectations in case of designing innovative solutions, they are, however, a set of recommendations derived from the good practice and their application in case of conventional ships is fully satisfactory. In the present paper the author assumes that dimensions of structural elements are determined only according to the requirements of a classification society enabling quick and automatic dimensioning of many structural variants by an optimization algorithm.

A primary objective of the ship structural optimization is to find the optimum positions of structural elements, also referred to as topological optimization, shapes (shape optimization) and scantlings (sizing optimization) of structural elements for an objective function subject to constraints. Formally, selection of structural material can also be treated as a part of the optimization process (material optimization). The topological optimization means searching for the optimal existence and space localization of structural elements while the shape optimization is searching for the optimal shape of a ship hull body. The sizing optimization can also be expressed as a process of finding optimum scantlings of structural elements with

fixed topology and shape. Selection of the structural material is usually not an explicit optimization task but is rather done according to the experience and capability of a shipyard. Systematic optimization procedures for the selection of structural material are applied directly in rare cases.

Moreover the design of such a complex object as seagoing ship structure is a solution of the multi-objective optimization task including many optimization criteria often counteracting each other, e.g. small hydrodynamic resistance vs. large cargo deadweight, high structural strength and reliability vs. low structural weight. This requires a comprehensive search of the solution space, without a capability to select one solution unequivocally selected as the best one, as it is in the single-objective optimization tasks. This is because multi-objective optimization does not yield an unequivocal determination of the single variant proposed for further development, but a set of compromise solutions (infinite in general), which is used as a basis for taking the final design decision consisting in a selection of a solution (or solutions) to be further developed. a task of the multi-objective optimization is thus the appropriate identification of the set of "best possible compromises" or the single "best possible solution" as a result of the multi-objective seagoing ship structure design process.

Due to its high complexity, in spite of rising research and computational resources the multi-objective optimization of the seagoing ship structures is still held back by a number of obstacles hindering its application in practice, and the attempts with respect to this problem can be judged to be marginal. Most authors assume that an outcome of the multi-objective optimization task is a set of the Pareto-optimal solutions, while it is impossible to point the objectively best one among

them. Classical multi-objective optimization algorithms allow for finding in the best case a single solution in the single algorithm run, which makes them unsuitable for multi-objective optimization tasks involving the determination of the Pareto-optimal solution set. At the same time, the evolution-based algorithms for example allow for the determination of this set in the single algorithm run thanks to the fact that they process not single solutions but usually the large set of potential solutions which in their consecutive steps gradually evolve to a Pareto-optimal set.

A practical example of the application of the developed computational tool is presented, featuring the multi-objective optimization of the structure of fast passenger-vehicle ferry model design named Auto Express 82 m. a task of the two-objective optimization of the ship structure minimizing its weight and surface area for cleaning and painting has been formulated. Number of optimization objectives has been limited to two for the sake of simplicity of graphical presentation of results and their analysis. The precision of the developed computational model has been limited to the level typical for the preliminary phases of the design process for structures of a similar type. The formulated constraints account for structural strength values estimated using procedures laid down in classification rules. A computation application has been built for the solving of so formulated task, being based on a specialized genetic algorithm code. Appropriate models of the ship structures have been built and computational investigations carried out. The obtained results have led to the conclusion that the genetic algorithms can be considered as a method allowing for the solution of the topology-size multi-objective optimization problems formulated for the ship structures in the concept design stage.

Because the prepared paper has considerable sizes it was divided on three parts:
– Part I: Theoretical background on evolutionary multi-objective optimization,
– Part II: Computational investigations,
– Part III: Analysis of the results,
for the publication in the consecutive volumes of the journal.

## BASIC CONCEPTS OF MULTI-OBJECTIVE OPTIMIZATION

From mathematical point of view, the multi-objective optimization can be defined in the general way as a procedure consisting in selecting an element of the set on the basis of relations establishing some order in this set. In reference to the ship design the elements of this set are in general the representations of particular problem solutions, such as ship structural variants, various types of ship main propulsor control ensuring that a specified aim of control is attained (e.g. lowest ship operation cost) etc. This set, referred to as "set of possible solutions", is a subset of solution space $V_x$. As we know, the set of such solutions is limited by the introduction of various constraints and such a constrained set is than called "set of feasible solutions $\Phi$". For obvious reasons set of feasible solutions $\Phi$ is also subset of solution space $V_x$, and each element of this space is a vector of design variables $x \in V_x$, Fig. 3. Solution space $V_x$ may be a functional space or Euclidean space, if all its coordinates are numbers:

$$x = [x_1 \ x_2 \ ... \ x_i \ ... \ x_n]^T \in V_x \qquad (1)$$

The case where the solution space is a n-dimensional Euclidean space $\Re^n$ is most often encountered in practical applications. In the further part of the paper the solution space is an $\Re^n$ space.

An objective of the multi-objective optimization problem solving can be formulated in the following way: find a combination of design variable values $x = [x_1 \ x_2 \ ... \ x_i \ ... \ x_n]^T$, which optimizes at the same time all components of a given objective function vector $f(x) = [f_1(x) \ f_2(x) \ ... \ f_s(x) \ ... \ f_S(x)]^T$. With a possibility to impose constraints on the variability ranges of design variables. It is also assumed that all the functions occurring within the problem are real ones, and the number of constraints is finite. Taking into account the demand or computational resources and their cost, another requirement may be formulated, that the selection made could be implemented at the lowest possible cost. Exact definition of the meaning of term „optimize" has crucial significance in case of multi-objective optimization problem. In the further part of the paper this concept is going to be discussed in more detail. The general mathematical formulation of a multi-objective optimization problem can be presented as follows:
- for design variables:

$$x = [x_1 \ x_2 \ ... \ x_i \ ... \ x_n]^T:$$
$$x_{i,min} \leq x_i \leq x_{i,max}, \ i = 1, 2, ..., n \qquad (2a)$$

- optimize:

$$f(x) = [f_1(x) \ f_2(x) \ ... \ f_s(x) \ ... \ f_S(x)]^T$$
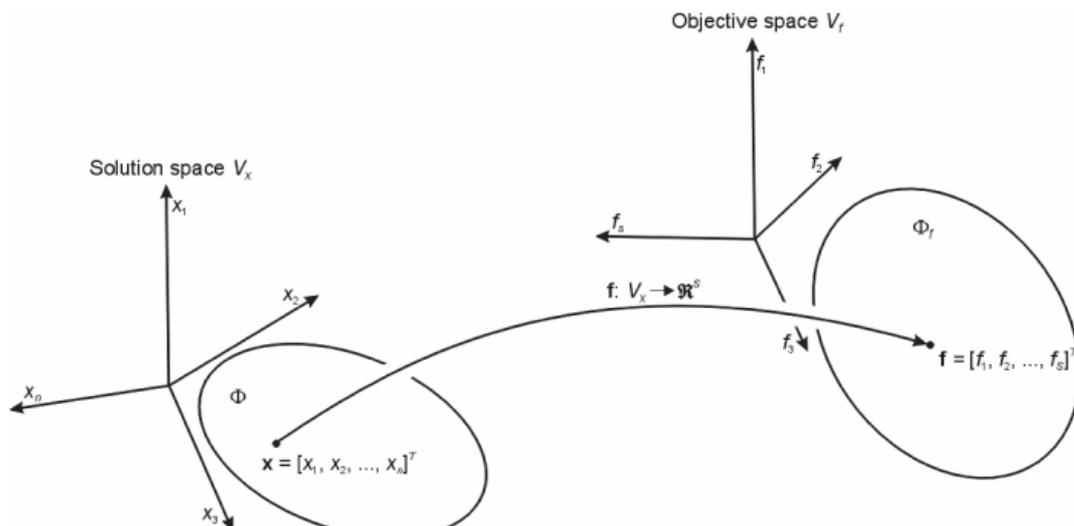$$s = 1, 2, ..., S \qquad (2b)$$



***Fig. 3.*** *Graphical illustration of multi-objective optimization task; x ∈ Φ – vector of design variables, Φ – set of feasible solutions, f(x) ∈ Φ_f – vector of optimization criteria, Φ_f – set of attainable objectives (goals, evaluations)*

- subject:

$$h_k(x) = 0 \quad k = 1, 2, ..., m_1 \qquad (2c)$$

$$g_j(x) \geq 0 \quad j = 1, 2, m_2 \qquad (2d)$$

where:

| | | |
|---|---|---|
| $f(x)$ | – | a single column objective function vector S, |
| $x$ | – | a single-column independent variable vector n (design variables), |
| $h_k(x)$ and $g_j(x)$ | – | constraint functions |
| $x_{i,min}$ and $x_{i,max}$ | – | respectively and upper and lower limit of variables. |
| Eq.s $h_k(x) = 0$ | – | referred to as equality constraints, while inequalities $g_j(x) \geq 0$ – inequality constraints. |

A solution of multi-objective decision making problem originally formulated by Francis Ysidro Edgeworth in 1881 [Edgeworth (1881)] and generalized then by Vilfredo Pareto [Pareto (1896)] is the commonly accepted measure of quality in the multi-objective selection problems. It is now referred to as Edgeworth-Pareto optimum or Pareto optimum. According to the definition, a design of the seagoing ship structure may be called Pareto-optimal under the condition that there are no other variants of the structure which are better with regard to at least one criterion while at the same time being equally good with regard to all the other optimization criteria. This means that the Pareto-optimal structural variant cannot be improved without simultaneous worsening of at least one criterion. Pareto-optimal designs are also referred to in literature as being non-dominated ones, trade-offs, noninferior or Pareto-efficient. The variant of the ship structure is not Pareto-optimal if there is any other variant improving at least one criterion while at the same time not worsening the values obtained for the remaining ones. Such variants are also called dominated ones or inferior ones.

Using the concept of domination formulated by Pareto we can say that a multi-objective optimal solution is each solution which has no other feasible solutions dominating it. We say that the solution $x_1$ dominates (is better than) the solution $x_2$, see Fig. 4, if the following two conditions are satisfied:

$$f_s(x_1) \leq f_s(x_2), \text{ for all } s = 1, 2, ..., S \qquad (3a)$$

$$f_{s'}(x_1) < f_{s'}(x_2), \text{ for at least one } s' = 1, 2, ..., S \qquad (3b)$$

There is then no such a solution in the set of feasible ones for which the value of all criteria would be "better" than their respective values for any multi-objective optimal one. In other words, the multi-objective optimal solution is such feasible solution, for which no better solution can be found in the set of feasible solutions. The word "better" should be understood here in the sense of Pareto domination.
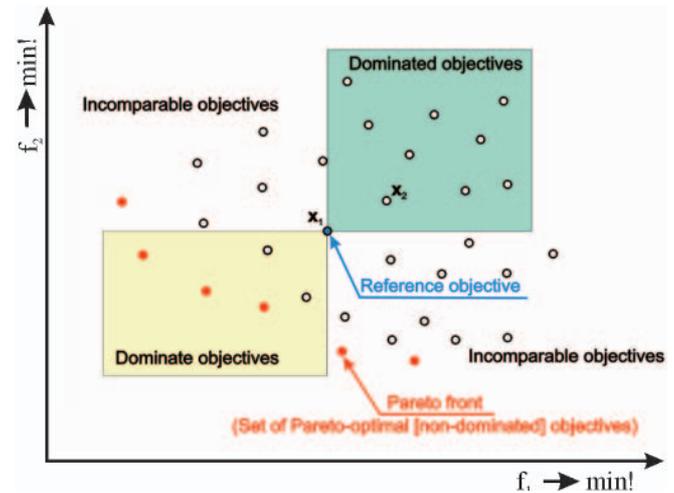


*Fig. 4. Graphical illustration of possible relations between solutions in the objective space; the objectives dominated by a reference objective and the objectives dominating it are highlighted; objectives non-dominated by any other ones belonging to the set constitute a set of Pareto-optimal objectives; $f_1 \to min!, f_2 \to min!$*

The concept of Pareto domination allows for introduction of the two-value measure of quality for solutions of the multi-objective optimization problem. It allows for dividing the set of feasible solutions into two subsets, Fig. 4: (1) subset of dominated solutions $\Phi_{fd}$, and (2) subset of non-dominated solutions, $\Phi_{fnd}$ which may be considered to be the solution of a multi-objective optimization problem. Two-value of this measure does not allow for a further evaluation of a feasible dominated solutions set[2], and, particularly, does not allow for relative estimation of distances between dominated solutions and the set consisting of non-dominated solutions (set of Pareto-optimal solutions) by any other feasible solution. In spite of this, the relation of Pareto domination is the one most often used for the definition of multi-objective optimal solution. In the further
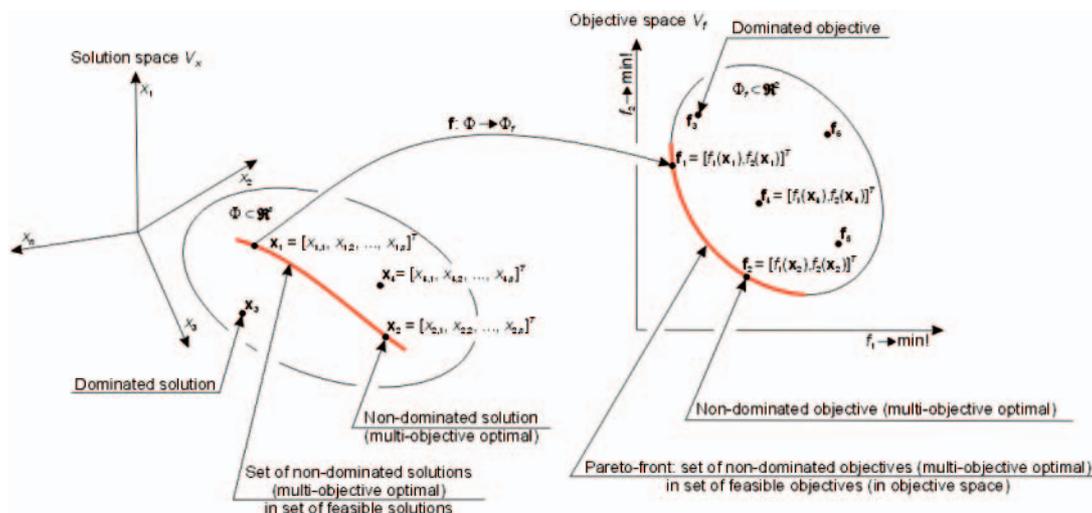


*Fig. 5. Graphical illustration of line of non-dominated objectives in the objective space and set of non-dominated solutions in solution space; **f** - vector objective function, **Φ** - set of feasible solutions, **Φ**_f - set of feasible objectives; $f_1 \to min!, f_2 \to min!$*

---

[2]  Most detailed analysis of set of feasible solutions in relation to domination. Relationship of domination does not „see" many details of the evaluation space. The only details it can "see" whether the solution is dominated or non-dominated.

part of the work, when talking about domination relation, we shall then understand it to be the relation of Pareto domination, and the earlier used phrase "optimize vector objective function f" shall be understood as a command: find the Pareto-non-dominated solutions within the feasible solutions set.

A basic feature of multi-objective optimal solutions is the fact that there are many (or even infinite number) of them exist in practical problems. In the case of the feasible set is continuous, and is a subset of $\Re^n$, then the set of feasible objectives shall also be continuous and a subset of $\Re^2$ (as a result of two quality criteria). As a result of the analysis of feasible objectives set we can obtain not several multi-objective optimal points but the whole curve of multi-objective optimal objectives, presented in Fig. 5.

## SOLVING OF MULTI-OBJECTIVE OPTIMIZATION PROBLEM: HOW FIND SET OF NON-DOMINATED SOLUTIONS

Monographies and books on the general problems of multi-objective optimization are e.g.: [Eschenauer et al. (1990)], [Statnikov et al. (1995)], [Sen and Yang (1998)] and [Stadler (1998)]. Rare works focusing on multi-objective optimization of ship structure include: [Shi (1992)], [Das (1993)], [Das et al. (1993)], [Trincas et al. (1994)], [Ray and Sha (1994)], [Sen and Yang (1995)], [Jianguo and Zuoshui (1996)] and [Parsons and Singer (2000)].

The methods used for the solving of tasks of this type, outlined in monographies and currently applied in practice, my be divided into two basic groups: (1) classical methods, and (2) methods inspired by natural systems, evolutionary methods in the peculiarity. [Cohon (1978)], [Stadler (1988)], [Statnikov et al. (1995)] discuss classical methods, which include two basic methods used for the solving such tasks: (1.1) optimization problems are solved with regard to all optimization criteria taken individually one by one while the remaining criteria are included in the set of constraints, (1.2) a substitute optimization criterion is formed of the adopted ones as a linear combination of the original component criteria multiplied by the appropriately selected weight coefficients, and then the optimization problem is solved with regard to such a newly formed aggregate criterion. For case (1.2) a series of calculations is usually carried out for variously adopted values of weight coefficients, and the best among the found solutions is taken as the solution of the problem. The methods based on a aggregation of the vector objective function have been used in wide-ranging applications also in the methods of evolution-based multi-objective optimization, as they allow for the use of well developed single-objective optimization algorithms. Fundamental disadvantages of methods from this group are: (a) seeking only a single point on the non-dominated solutions front and resulting necessity to make numerous calculation runs for the single optimization task, (b) sensitivity of some solutions to the shape of non-dominated solutions front, and (c) the fact that expert knowledge is required at the beginning to specify the weight coefficients used for component optimization criteria.

Classical methods used for the solving of multi-objective optimization problems based primarily on the aggregation of vector objective functions are easy to implement but ineffective in many cases. However, evolutionary multi-objective optimization algorithms developed in the recent years have been proven to be highly effective in this regard [Deb (2001)], [Osyczka (2002)], [Sarker et al. (2002)], [Abraham et al. (2005)] and [Coello Coello et al. (2007)].

Not numerous but highly promising results in the field of genetic algorithms use for multi-objective optimization tasks have been obtained lately including also the results in the field of ship structures: [Okada and Neki (1992)], [Hutchinson et al. (1998)], [Kitamura et al. (2000)], [Klanac et al. (2009)] and [Sekulski (2010)]. [Jang and Shin (1997)] have applied the Evolutionary Strategy method for the multi-objective optimization of ship structures.

Special evolutionary multi-objective optimization methods can also be applied as far as genetic algorithms are concerned: VEGA – [Schaffer (1985)], HLGA – [Hajela and Lin (1992)], FFGA – [Fonseca and Fleming (1993)], NPGA – [Horn et al. (1994)], NSGA – [Srinivas and Deb (1995)], RWGA – [Murata and Ishibuchi (1995)], MOBES – [Binh and Korn (1997)], SPEA – [Zitzler and Thiele (1998)], MOMGA – [Veldhiuzen (1999)], PAES – [Knowles et al. (1999)], NSGA-II – [Deb et al. (2000)], SPEA2 – [Zitzler et al. (2001)]. Fundamental advantages of these methods are: (1) effective search of solution space and (2) capability to illustrate the non-dominated solutions front in a single simulation run. Excellent presentation of evolutionary methods of multi-objective optimization can be found in recently published monographs [Deb (2001)], [Osyczka (2002)] and [Coello Coello et al. (2007)]. The principal elements of these algorithms are: (1) selection strategies based on the Pareto-domination relation, (2) niching strategies to preserve diversity in the consecutive populations, (3) elitist strategy to ensure survival of non-dominated solutions in the time of evolution. Methods based on aggregation of the objective function are considered less effective and are closed and no more developed stage of the history of the evolutionary multi-objective optimization. Despite it, we can, however, claim that the algorithms employing aggregation of the objective functions are efficient algorithms transient from the classic methods to the advanced algorithms employing the Pareto-domination relation for the variant selection. The researchers have reported for several years that if the number of the optimization criteria is greater than 3, the methods based on the domination relation turn to be ineffective since together with the increase of the number of optimization criteria the number of non-dominated variants decreases reducing the effectiveness of the selection operator [Hughes (2003)], [Purshouse and Fleming (2003)], [Jaszkiewicz (2004)] and [Hughes (2005)]. The aggregation methods have been found promising again with the hope to: (1) developing more simple and intuitive algorithms than algorithms based on the domination relation, obtaining expertise on the multi-objective ship structural optimization, (2) developing effective algorithms for problems with a large number of the optimization criteria. Of course, it needs to be considered what number of criteria is practically justified. With regard to capability of processing data by human beings and capability to work out decisions it seems that the number of the optimization criteria in practical problems should be between 5 and 7.

Due to the following practical problems (1) lack of information about the actual localization of non-dominated solutions set, and (2) necessity to deploy significant computational resources to solve the multi-objective optimization problem, main effort in the practical evolution-based multi-objective optimization is directed at determining the acceptable approximation of Pareto set instead of accurate composition of this set. With regard to this it can be assumed that in practice the result of multi-objective optimization process is a set of non-dominated solutions called shortly the approximation of Pareto set and not the exact Pareto-optimal solutions set. Practical formulation of multi-objective optimization problem and of attained results should follow this guideline.

Opposite to single-objective optimization problems, where the objective function and fitness function are often the same,
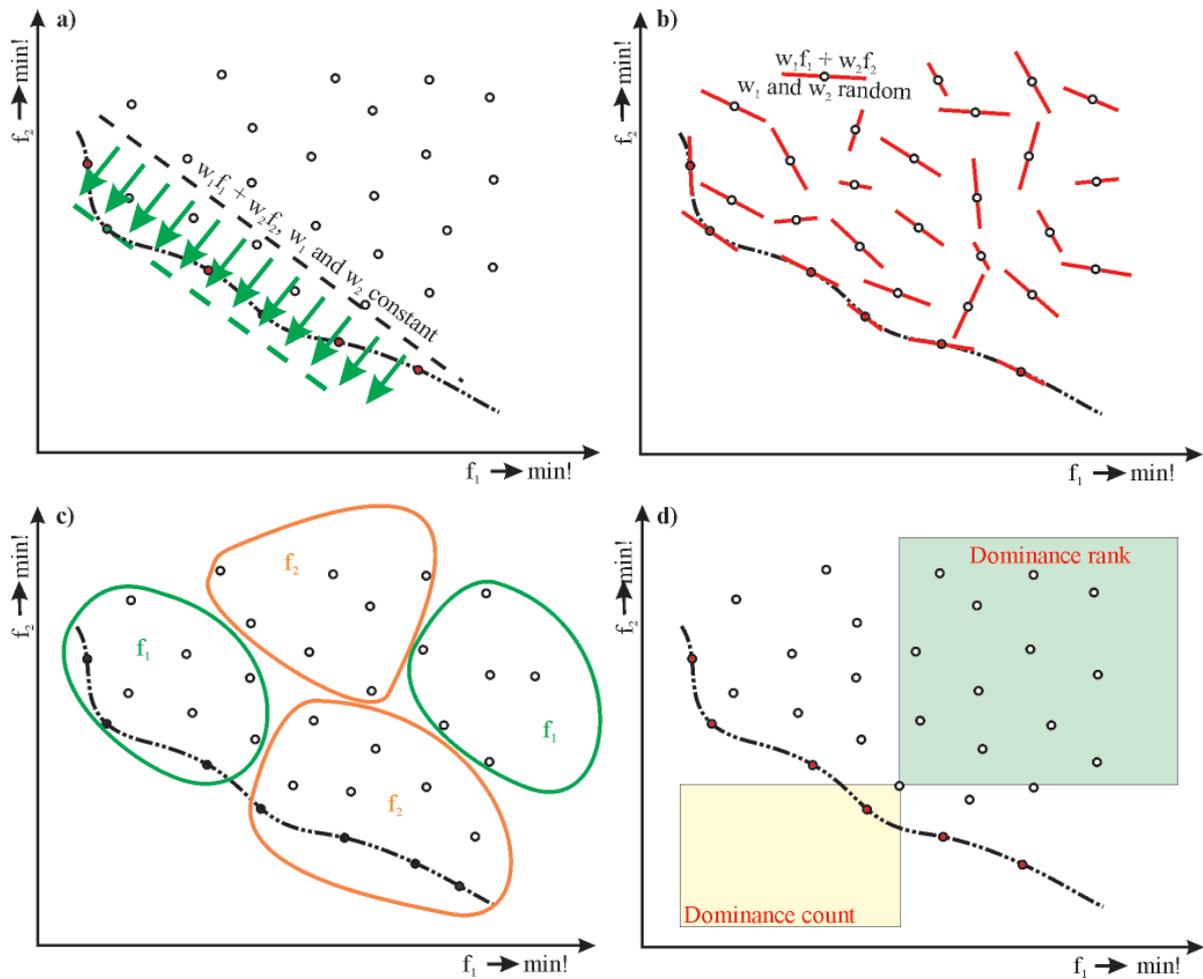
***Fig. 6.*** *Graphical illustration of selected strategies for taking into account the particular optimization criteria used in the multi-objective optimization algorithms:* ***a)*** *selection with respect to the scalar objective function with fixed weights of optimization criteria,* ***b)*** *selection with respect to the scalar objective function with random weights of optimization criteria,* ***c)*** *division of the variant set into sub-sets and selection in each of them with respect to single criteria,* ***d)*** *selection with respect to attributes of Pareto domination; $f_1 \to min!$, $f_2 \to min!$*

the evolution-based algorithms of multi-objective optimization feature both the fitness function and the selection process taking into account a number of criteria which are included in a single fitness function. From this point of view these methods can be in general divided by the type of fitness function used for calculations into the following classes: a) selection with respect to the scalar objective function with fixed weights of optimization criteria, b) selection with respect to the scalar objective function with random weights of optimization criteria, c) division of the variant set into sub-sets and selection in each of them with respect to single criteria, Fig. 6.

First proposal of Fig. 6a, stemming from classical methods used for the determination of compromise surface, consists in summing the criteria up and formulating a single, parameterized objective function. Parameters of this substitute objective function are fixed during the optimization run, which allows for finding the one non-dominated solution. The multi-objective optimization problem is reduced to the single-objective problems, with the criteria being usually referred to as objective functions. The simplest concept is the introduction of objective function F as a linear combination S of partial optimization criteria $f_s$:

$$F(x) = \sum_{s=1}^{S} w_s f_s(x) \qquad (4)$$

where:

$w_s$ – coefficients determining the weights given to particular criteria.

Next proposals, illustrated on Fig. 6b and Fig. 6c, are based on the weighted sum of optimization criteria, where weight coefficients represent the values changing in the process of evolution [Hajela and Lin (1992)], [Ishibuchi and Murata (1996)]. Weight coefficients of the substitute objective function change in a specific way during the optimization run, which allows for finding the non-dominated solutions set instead of a single compromise solution.

Methods based on selection with respect to the scalar objective function with random weights of optimization criteria Fig. 6b employ numerical procedures for setting random values of weight coefficients $w_s$. The simplest and most frequently applied implementation of the method is setting random values of uniform distribution in range [0, 1].

Methods based on selection according to single criteria as it is illustrated on Fig. 6c consist of mechanisms switching between the criteria during the selection phase. In each case where the algorithm commences the execution of reproduction, some criterion (potentially other) decides which member of population is going to be copied to the set of variants earmarked for crossbreeding. For example, [Schaffer (1985)] proposed an algorithm, where the population is divided in advance to identical parts and then a different, single criterion is used on the members of each of groups one by one and [Kursawe (1991)] proposed a different method consisting in the random selection of a single optimization criterion to be used in the next step of selection process, with probabilities to be set by the user or randomly adjusted during the evolution.
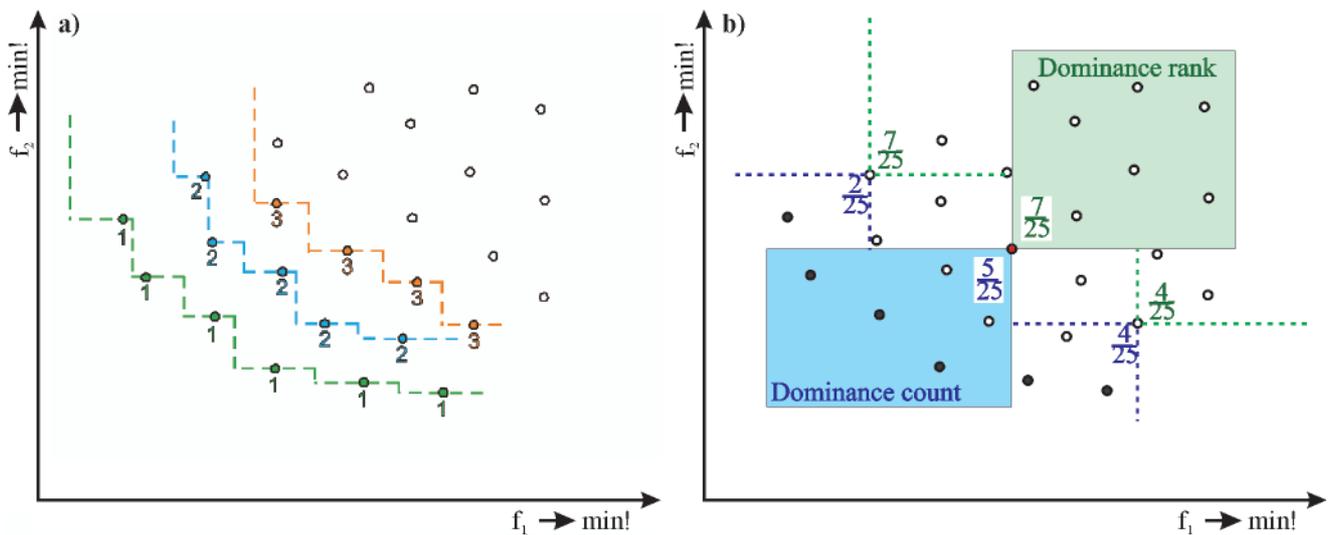
***Fig. 7.*** *Graphical illustration of basic concepts used for taking into account of variant domination in the multi-objective optimization algorithms:*
***a)*** *domination depth,* ***b)*** *dominance rank, and dominance count; for non-dominated solutions (Pareto-optimal) dominance count equals zero; dominance rank for feasible solutions not dominating any other solution equal zero;* $f_1 \rightarrow min!$, $f_2 \rightarrow min!$

Assigned to feasible variants two-argumental domination attribute is a rather general information and does not refer to inner structure of feasible set. Particularly it does not provide knowledge e.g. about (1) what number of variants is dominated by every of feasible variants, (2) what number of solutions is dominated by a given not-dominated variant, (3) how distant is a given dominated feasible variant from Pareto front, etc. Detail study of feasible set structure enables the use of additional knowledge and elaboration of very refined and useful tools supporting desirable convergence of computing algorithm. Different detailed solutions using this concept can be found in bibliography, for example: [Fonseca and Fleming (1993)], [Zitzler and Thiele (1999)], [Deb (2001)], [Osyczka (2002)], [Abraham et al. (2005)], [Coello Coello et al. (2007)].

The concept of calculating the fitness of variants while accounting for the Pareto-type domination, Fig. 6d, has already been proposed by [Goldberg (1989)] and since then taking the advantage of partial order being present in the population, Fig. 7a, has been facilitated by many various methods developed specifically for purpose. Basing on the concept of dominance depth proposed by [Goldberg (1989)], the feasible variants appearing in the consecutive generations are divided into consecutive fronts of non-dominated variants located deeper and deeper in the feasible set, or in other words, further and further from Pareto front. Other proposals are based on the concept of dominance rank, Fig. 7b, which is based on use of the number of variants dominated by a selected variant for the calculation of fitness function value [Fonseca and Fleming (1993)]. Also a dominance count (which is a number of variants dominating a selected variant), Fig. 7b, may be taken into account. For example SPEA strategies [Zitzler and Thiele (1999)] and SPEA2 [Zitzler et al. (2001)] for the calculations of fitness function are based on both approaches, i.e. the rank and the dominance count. Disregarding the specific strategy applied, in case of these methods the value of fitness depends on the characteristics of the variants remaining in the population, in contrast with strategies based on aggregation of objectives or the selection of variants on the basis of less or more arbitrary selected single criteria, where the adaptation values are independent from the characteristics of other variants in the population.

The multi-objective evolution-based optimization algorithms outlined above have been tested by other authors on simple problems of multi-objective optimization eg. [Zitzler

(1999)], [Zitzler et al. (1999)], [Zitzler and Thiele (1999)]. As no systematic research into the suitability of these algorithms for the solving of optimization problems involved in the design of seagoing ship structures has been carried out so far, then the application of a particular method should be preceded by systematic research into its effectiveness in the problems involved in the design of such structures.

## LOOKING INSIDE THE FEASIBLE SOLUTIONS SET: DOMINANCE RANK AND DOMINANCE COUNT

Let us recall that assignation of a two-argument dominance attribute (0 or 1) to the feasible variants allows for dividing the feasible variants set into two subsets: (0) subset of dominated variants and (1) subset of non-dominated variants, Fig. 8a. As can be seen, this information is quite general and does to refer to the internal structure of the feasible set, and in particular does not provide any knowledge about e.g. the following: (1) what number of variants is dominated by each of feasible variants, (2) what number of solutions is dominated by a given non-dominated variant, (3) how far away from Pareto front lies a given dominated feasible variant, etc. [Goldberg (1989)] proposed for ordering the feasible solutions depending on the depth of consecutive fronts of non-dominated solutions, Fig. 8b. As can be seen, the dominance attribute which in this case is the domination depth takes discrete values: 1, 2, 3, etc. More detailed analysis of the structure of feasible set allows for the use of following concepts as dominance attributes: (a) rank of feasible variants, dominance rank, Fig. 8c, and (b) feasible variant evaluation dominance count, Fig. 8d. Values of both attributes vary in continuous way over the set of feasible solutions, depending on the strategies adopted for their determination.

The dominance rank for a given variant is proportional to the number of feasible variants dominated by a given variant, Fig. 8c. The value of dominance rank is then highest for non-dominated variants approximating the Pareto front. For the variants lying outside of Pareto front (or its approximation), deeper and deeper in the feasible variants set, the values of dominance rank are a falling measure of their distance from the front. The dominance count of a feasible variant is proportional to the number of other feasible variants dominating a given variant, Fig. 8d. Its value is then lowest and equals zero in
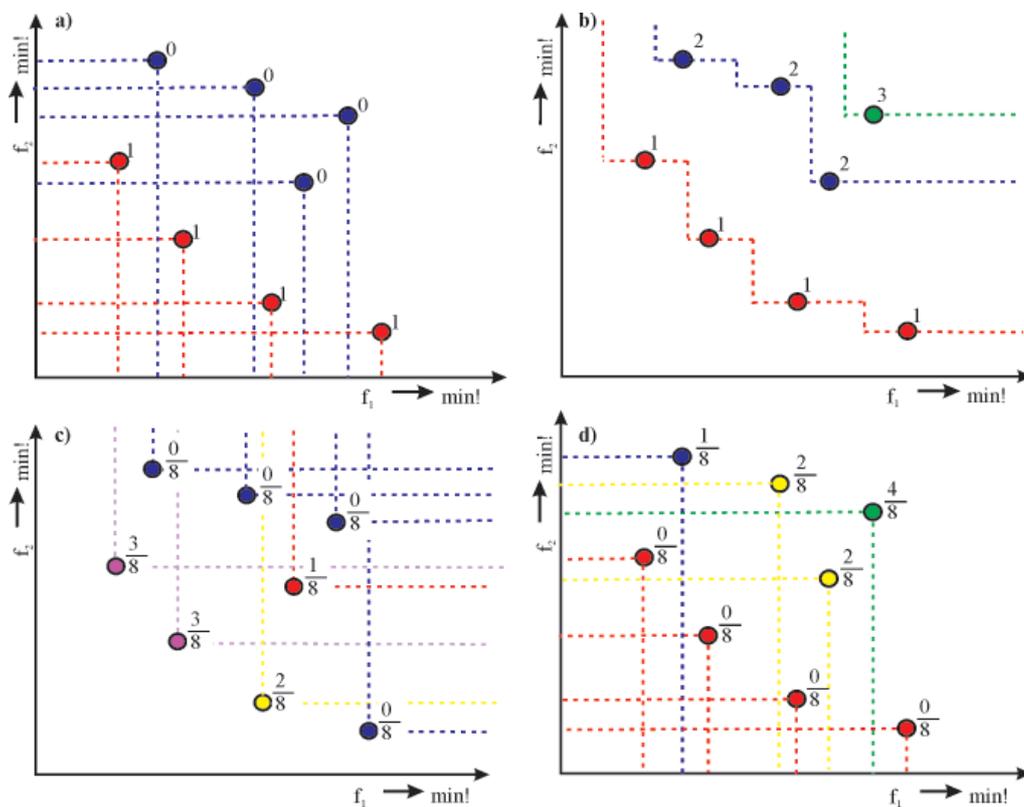
**Fig. 8.** *Graphical illustration of the groups of dominance attributes: **a)** classifying the feasible variants basing on Pareto-domination relation, **b)** classifying the feasible variants basing on domination depth, **c)** dominance rank (proposed strategy for the determination of a number of variants dominated by a given variant) in the objective space, and **d)** dominance count (proposed strategy for the determination of a number of variants dominating a given variant); $f_1 \rightarrow$ min!, $f_2 \rightarrow$ min!*

case of non-dominated variants which approximate the Pareto front. For the variants lying outside of Pareto front (or its approximation), deeper and deeper in the feasible variants set, the values of count are a rising measure of their distance from the front. Other schemes for dominance attributes have been proposed by e.g. [Fonseca and Flaming (1993)], [Zitzler and Thiele (1999)]. Proper application of both attributes allows for development of very subtle and useful tools supporting the desired convergence of the calculation algorithm.

Attention should be paid to the fact that the values of objective function for each variant are determined after its characteristics (evaluation of design criteria and compliance with imposed constraints) are calculated, and the distances between variants in objective space of already determined objective function values are constant and may be used as a measure of absolute mutual similarity of solutions. It is however a different story in case of dominance attribute. It is not an absolute measure of variant location in the objective space, but rather a relative measure of variant location in this space determined in relation to other feasible solutions of the set. As the feasible set continuously evolves, then the dominance attribute of its elements, feasible solutions, evolves as well, because in consecutive generations it is determined over a different set of feasible solutions, in spite of the fact that the objective function values remain unchanged. This is quite obvious: in the population initiating the simulation the first solution (non-dominated) shall be granted dominance depth 1 according to the proposal by [Goldberg (1989)]. As a result of the progress of simulation and the feasible solutions gradually approaching the Pareto optimal front, this solution is going to be overtaken by other ones and its dominance depth is going to drop lower and lower. Moreover, the difference in dominance depth between two solutions is not fixed and may not be used as a relative measure of value for two compared solutions. Of course, in each case the non-dominated solution

is better than any dominated one, but it cannot be simply claimed that in case solution B has a dominance depth lower by 1 from solution A, and solution C has a dominance depth lower by 3 from solution A, then the solution B is going to be better than solution C with regard to optimization criteria. Let's analyze the example illustrated in Fig. 9, with initial set from Fig. 9a consisting of solutions A, B and C, where solution a dominates the others, but they do not dominate each other. Solution a has dominance depth equal 1 while solutions B and C have dominance depth 2. Let's then assume that two new feasible solutions D and E are added to the feasible set in the next generation as shown in Fig. 9b. Solution D is dominated by a and at the same time dominates the solutions C and E. Solution E is dominated by solutions a and D and at the same time dominates the solution C. Now, solution C obtains a new value of domination depth, equaling depth 4 (used to have depth 2 before), while the solution B retains dominance depth 2, which takes place in spite of the fact that solutions C and B did not change their locations in the objective space.

Like dominance depth neither dominance rank nor dominance count of a given solution are then objective measures of its value, yet they are practicable and widely used in practice for the evaluation of solutions sought for in the problems of multi-objective optimization.

Various particular methods based on application of these concepts of dominance attributes are found in literature, for example [Fonseca and Fleming (1993)], [Zitzler and Thiele (1999)]. Each such a method has its own particular characteristics, but the early experiences of the author and also other researchers, e.g. [Leyland (2002)] suggest that the influence of particular dominance rank strategies on the convergence of optimization process may be insignificant. In particular, while applying modern elite strategies the solutions quickly converge to non-dominated solutions set.
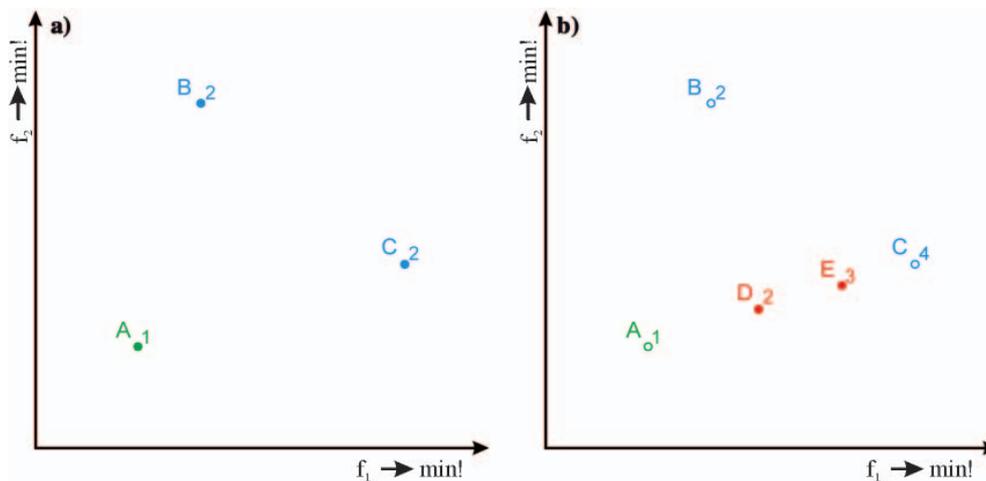
*Fig. 9. Illustration of the concept of dominance depth shift without the relocation of solutions in the objective space, for min! type criteria; a) dominance depth values for solutions A, B and C in initial population (preceding one); b) extension of feasible solutions set by solutions D and E does not cause a change of dominance depth of solutions a and B but does cause a change of dominance depth of solution C without causing its relocation in the objective space: what is changed only is the relative location of this solution in relation to the remaining feasible solutions*

## CALCULATION TOOL FOR EVOLUTIONARY MULTI-OBJECTIVE OPTIMIZATION OF SHIP STRUCTURE

The essence of the classic Darwin's theory of evolution [Darwin (1859)] is an idea that the evolution of biological organisms is driven by natural selection which affects hereditary traits ensuring survival (in the environment). According to the concept of the natural selection some organisms in certain environmental conditions manage better than others: they survive more often, have more numerous offspring and in time there are more and more of them. Therefore "the environment" selects beings better adjusted to certain situation. If environmental conditions were to change, the organism with a set of traits best suited for adapting to them, it will start to dominate. The concept describing this occurrence in evolutionary biology is adaptation or fitness, which means probability of surviving or reproduction in certain environment. There are mutations in every generation that provide the population with new variants. Natural selection sifts them: environmental severity decreases the number of "bad" ones (relatively unfit) variants and increases the number of "good" ones (relatively fit). Therefore adaptive evolution is a two-phase process in which tasks are strictly divided between mutation and selection. It is worth to notice that there are many variants in the population that can help it adapt to change in environmental conditions simultaneously. Consequently, the better the individual is adapted to environment, the faster its share in population will grow. Natural selection has very good "eyesight", which allows to perceive slight differences in adaptation between individuals. Modern theory of evolution, Neo-Darwinism, performs a synthesis of classic Darwin theory and genetics, considering natural selection for basis of species evolution, in which the key role is played by variations (mutations, also random) of genetic code.

Genetic algorithms are an example of concepts of adaptation and evolutionary mechanism such as inheritance, mutation, natural selection, and recombination (or crossover) to solve complex optimization problems see eg. [Goldberg (1989)], [Michalewicz (1996)], [Coley (1999)] and [Man et al. (1999)], Population of rapidly reproducing creatures/individuals (representing models of optimized objects) is placed in the artificial environment (represented by optimization criteria and constraints), in which adaptation differences are exaggerated in order to speed up the evolution. The size of the population is determined in a way to make it large enough to ensure steady flow of high number of newly occurring mutations. Applied in experimental evolution selective pressure can be abnormally strong – probably much harder than in nature. Individuals with feature variants that are favorable in inhabited environment will survive longer and have more offspring than those with unfavorable features. Therefore favorable features will accumulate over time in process of inevitable "natural selection". That kind of lab procedure simulated with computers allows to quickly obtain variants of optimized objects (in our case ship structures) with desirable features and produce better and better solutions.

Genetic algorithms (GA) have already been extensively described in literature discussing their theoretical foundations, details of calculation procedures and their practical applications, so these problems are not going to be discussed here again and the reader is referred to respective literature e.g. [Goldberg (1989)], [Davis (1991)], [Michalewicz (1996)].

The genetic algorithm is typically implemented in the form of computer simulations where a population of abstract representations (called chromosomes) of candidate solutions (called individuals) to an optimization problem evolves gradually towards better solutions. Traditionally, solutions are represented in the binary system as strings of 0s and 1s but different encodings are also possible. The evolution starts from a population of completely random individuals and is continued in subsequent generations. In each generation, the fitness of the whole population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), modified (mutated or recombined) to form a new population which becomes current in the next generation. Procedures of creation and evaluation of the successive generations of trial solutions are repeated until the condition of termination of computations is fulfilled, e.g. forming a predefined number of generations or lack of correction of the fitness function in a number of successive generations. The best variant found is then taken as the solution of the optimization problem.

The most important point of calculation tool for multi-objective optimization of ship structure is appropriate formulation of fitness function which governs the optimization process. In the first step a single, parameterized objective function was formulated, consists in summing the objective criteria with proper weight coefficients. The weight coefficients values of this substitute objective function are taken accordingly corresponding to multi-objective optimization strategy. The

simplest concept is the introduction of objective function F(x) as a linear combination S of partial objective criteria $f_s(x)$:

$$F(x) = \sum_{s=1}^{n_o} w_s f_s(\mathbf{x}) \qquad (5)$$

objective = criteria

where:
$w_s$ – coefficients determining the weights given to particular criteria,
$n_o$ – the number of partial optimization criteria.

The multi-objective optimization problem is then reduced to the single-objective problems, with the criteria being usually called objective functions.

Formulating of fitness function in the form of scalar substitute optimization criteria in form of Eq. 5 is a commonly accepted practice. Written in Eq. 5 partial optimization criteria were replaced by properly formulated utility functions of these criteria: $f_s(x) \rightarrow u_s[f_s(x)]$:

$$u_i(x) = \left(\frac{f_i(x)}{f_{i,max}}\right) \rightarrow max! \Leftrightarrow f_i(x) \rightarrow max! \qquad (6a)$$

$$u_j(x) = \left(\frac{f_{j,max} - f_j(x)}{f_{j,max}}\right) \rightarrow max! \Leftrightarrow f_j(x) \rightarrow min! \qquad (6b)$$

where:
$f_{i,max}$ and $f_{j,max}$ – the greatest values of respective optimization criteria anticipated in computations. These values are in practice set during test computations.

After assuming utility function in form of Eq. 6 the scalar substitute optimization criterion can be written down in form:

$$F(x) = \sum_{s=1}^{n_o} w_s u_s(\mathbf{x}) \qquad (7)$$

objective = criteria

Advantages of proposed substitutions are:
1. scalar substitute optimization criterion is maximized for any types of partial optimization criteria, that means $F(x) \rightarrow max!$,
2. values of utility function are dimensionless and normalized to a unit, that means $u_s[f_s(x)] \rightarrow [0, 1]$.

The calculation tool developed for use in optimization of ship structure should of course allow for accounting for a series of constraints imposed by design, local strength and overall strength. On the other hand, implementation of genetic algorithms requires that the equivalent problem is formulated without any constraints. Observing that genetic algorithms do not require continuity nor the existence of derivative functions, an external penalty function has been used [Fox (1971)], [Ryan (1974)], [Reklaitis et al. (1983)], [Vanderplaats (1984)]. The augmented objective function of unconstrained maximization problem f(x), has been formulated as a penalty function:

$$f(x) = \sum_{s=1}^{n_o} w_s u_s(\mathbf{x}) + \sum_{k=1}^{n_c} w_k P_k(\mathbf{x}) \qquad (8)$$

objective = criteria + constraints

where:
$u_s(x)$ – utility function in the constrained problem (Eq. 7),
$n_o$ – number of optimization criteria,

$P_k(x)$ – component of penalty function for the violation of k-th constraint,
$w_k$ – penalty coefficient for the violation of k-th constraint,
$n_c$ – number of constraints.

To adjustment mathematical form of penalty function $P_k(x)$ with dimensionless and normalized utility functions $u_s(x)$ similar requirements were put before penalty functions. Mathematical form of the penalty function was therefore proposed in form:

$$Pk(x) = \begin{cases} e^{|x-x_k|}, & \text{for} \quad \mathbf{x} < \mathbf{x}_k \\ e^{(x_k-x)}, & \text{for} \quad \mathbf{x} \geq \mathbf{x}_k \end{cases} \qquad (9)$$

where:
$P_k(x)$ – appropriate component of penalty function,
x – vector of current values of design variables,
$x_k$ – vector of design variables ensuring the least permissible value of appropriate constrain condition, (e.g. vector containing the least required by classification rules thickness of hull plate or the least required value of section modulus of stiffener).

Assumed forms of penalty functions components ensure normalization of penalty function components to a unit and introduce „soft" selection to the search, from a point of view of a factor defined by an appropriate component of penalty function $P_k(x)$.

As the augmented objective function f(x) expressed by the relation Eq. 8 with penalty components expressed by Eq. 9 is: (1) defined, (2) single-valued, (3) ascending, having real values and positive in the search space, it has been adopted directly as the fitness function.

As we already know the scheme of multi-objective optimization proposed in Eq. 8 allows only for rough differentiation of feasible solutions with regard to domination relation in Pareto sense, see Fig. 8a, and does not account for information about how many solutions are dominated by a given solution.

For the solving of the mentioned problem the author proposed a scheme in which the feasible solutions are ranked by the number of other solutions dominated by them, relative to the number of feasible solutions in the current population. Therefore, dominance rank $R_{fi}$ of i-th feasible solution is specified by an equation:

$$R_{fi}(i) = \frac{\sum_{j=1, j\neq i}^{N_{fi}} dm(i, j)}{N_{fi}} \qquad (10)$$

where:
dm(i, j) = 1 when i dominates j,
dm(i, j) = 0 in other cases,
i, j – indices of verified feasible solutions,
$N_{fi}$ – number of feasible solutions in the current population.

Fig. 8c and Fig. 10a. The advantages of proposed strategy are: (1) ease of calculations, (2) standardization of dominance rank values in [0, 1] range, and (3) ascending values of dominance rank for solutions approaching the Pareto front (lying at the edge of feasible set). Thanks to properties (2) and (3) the value of dominance rank calculated in a proposed way may be directly included in the fitness function. In such a case selection is going to promote feasible solutions located close to Pareto front, while the solutions lying gradually further and

further from the Pareto front are going to be promoted weaker and weaker, which is a numerical realization of selection pressure exerted on solutions located close to Pareto front and which thus enhances the exploitational performance of the algorithm. A disadvantage of the proposed strategy is computational complexity $N^2$.

Similarly, feasible solutions may be classified by the number of solutions dominating them, relative to the number of feasible solutions. Thus, evaluation dominance count $C_{fi}$ of i-th feasible solution is expressed by the formula:

$$C_{fi}(i) = \frac{\sum_{j=1, j \neq i}^{N_{fi}} dm(j, i)}{N_{fi}} \qquad (11)$$

where:
$dm(i, j) = 1$ when i dominates j,
$dm(i, j) = 0$ in other cases,
i, j – indices of verified feasible solutions,
$N_{fi}$ – number of feasible solutions in the current population.

Fig. 8d and Fig. 10b. The dominance count defined in this way has the above mentioned properties (1) and (2), and property (3) ascending values of dominance count for the variants lying further and further from Pareto front (located deep inside the feasible set). Thanks to properties (2) and (3) the value of dominance count calculated in the proposed way may also be directly included in the feasible function - in such a case selection is going to promote feasible solutions located far from Pareto front, while the solutions approaching the Pareto front are going to be promoted weaker and weaker, which is a numerical realization of selection pressure exerted on solutions located far from Pareto front and which thus enhances the exploratory properties of the algorithm. Same as in the previous case, the disadvantage of the proposed strategy is computational complexity $N^2$.

As it has already been mentioned, the strategies for dominance ranking and the dominance count the feasible variants proposed by the author allows for their inclusion directly in the earlier formulated (Eq. 8) extended objective function of a unconstrained maximization problem f(x):

$$f(\mathbf{x}) = \sum_{s=1}^{n_o} w_s u_s(\mathbf{x}) + w_r R_{fi}(\mathbf{x}) + \\ + w_c C_{fi}(\mathbf{x}) + \sum_{k=1}^{n_c} w_k P_k(\mathbf{x}) \qquad (12)$$

combined objective = criteria + rank + count + constraints

where:
$u_s(x)$ – utility function for the constrained problem (Eq. 7),
$n_o$ – number of optimization criteria,
$R_{fi}(x)$ – dominance rank of feasible variant,
$w_r$ – dominance rank weight coefficient,
$C_{fi}(x)$ – dominance count of feasible variant,
$w_c$ – dominance count weight coefficient,
$P_k(x)$ – penalty function component for the violation of k-th constraint,
$w_k$ – penalty coefficient for the violation of k-th constraint,
$n_c$ – number of constraints. Giving a zero value to the respective weight coefficient lets the user deactivate a given component of combined fitness function.

For indicate a single solution which may be considered to be „the best" solution of a multi-objective optimization problem and the monitoring of evolution of non-dominated solution in the direction of theoretically lowest values of optimization criteria $f_i \rightarrow 0!$ as moving the set of non-dominated solutions in the desired direction the author has then introduced a concept of a asymptotic objective (asymptotic solution), which represents an objective/solution corresponding to asymptotical values of optimization criteria: $f_1 \rightarrow 0!$ $f_2 \rightarrow 0!$[3], Fig. 12b. Such a definition of a asymptotic solution it is possible to: (1) determine the distance between each non-dominated solution from and this point, and then choose a asymptotic closest solution, and additionally (2) monitor the evolution of set of non-dominated solutions distance from this solution during the simulation.
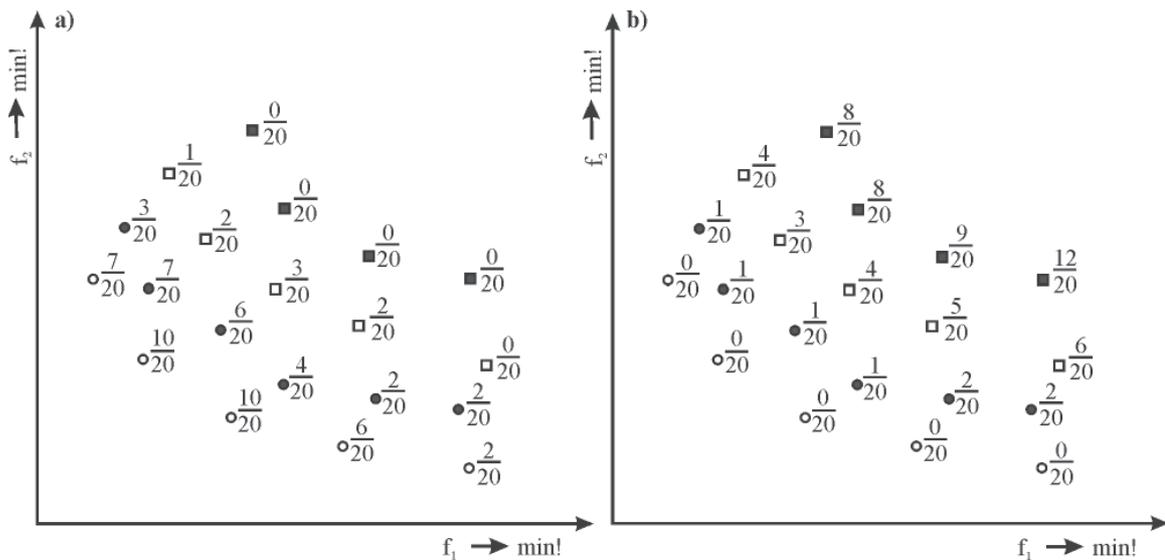


**Fig. 10.** *Illustration proposed strategy of value calculation of:*
***a)** dominance rank, **b)** dominance count for feasible solutions; $f_1 \rightarrow min!$, $f_2 \rightarrow min!$*

[3] It is then assumed that in case of min! type criteria the respective objectives (co-ordinates) are equal to zero. However, in case of max! type criteria the user shall set the value of these objectives (co-ordinates) as some known from experience values of respective criteria which are impossible to attain, but which are going to be approached asymptotically by the non-dominated solutions.

The method of calculating the distances between feasible solutions and a asymptotic solution allow for including a distance from an asymptotic solution directly in the earlier formulated extended objective function used for unconstrained maximization problem, f(x):

$$f(x) = \sum_{s=1}^{n_o} w_s u_s(\mathbf{x}) + w_r R_{fi}(x) + w_c C_{fi}(x) +$$
$$+ w_d[1 - d_{fi}(x)] + \sum_{k=1}^{n_c} w_k P_k(x) \tag{13}$$

combined objective = criteria + rank +
+ count + distance + constraints

where:

$u_s(x)$ – objective function for the constrained problem (Eq. 7),
$n_o$ – number of optimization criteria,
$R_{fi}(x)$ – dominance rank of feasible solution,
$w_r$ – dominance rank weight coefficient,
$C_{fi}(x)$ – dominance count of feasible solution,
$w_c$ – weight coefficient of dominance count,
$d_{fi}(x)$ – distance between the feasible solution and the asymptotic solution,
$w_d$ – distance weight coefficient,
$P_k(x)$ – penalty function component for the violation of k-th constraint,
$w_k$ – penalty coefficient for the violation of k-th constraint,
$n_c$ – number of constraints.

The combined fitness function used in the proposed form (Eq. 13), includes instruments which provide for effective solving of constrained multi-objective optimization ship structure problem. These are: combined fitness = objectives + rank + count + distance + constraints; objectives – represents selective pressure exerted in the direction of the desired values of optimization criteria, dominance rank as well as dominance count – represents selective pressure related to the location of an feasible variant with regard to non-dominated solutions set, measured with appropriate dominance attributes, distance – represents selective pressure related to the distance between an feasible solution and the asymptotic solution, constraints – represents the reduction of solution quality and the related reduction of selection probability caused by the violation of constraints. As the values of dominance rank, dominance count and the distance to asymptotic solution are calculated only for feasible variants, this involves also additional promotion of such variants. Proper use of the proposed components of the combined fitness function by the experienced user makes for a highly flexible and effective solving of multi-objective optimization problems based on a genetic algorithm or other evolutionary algorithms in general.

Domination of rank component is used for promoting the variants belonging to non-dominated solutions set which increases their chances for participation in reproduction and transmission of their characteristics to a larger number of descendant specimens, and thus boosts the propensity for exploration of solution space areas which contain good solutions. Variants lying far from non-dominated solutions set have lesser chances for reproduction. Such a strategy is going to reinforce the existing composition of non-dominated solutions set but it is not going to be conducive to supply of new, potentially interesting/attracting/desired solutions. It is then a strategy biased towards exploitation of areas with non-dominated solutions set variants as opposed to the exploration of a whole solution space. Domination of count component is

going to have an opposite effect: reduction of reproduction probability for variants belonging to non-dominated solutions set (use of elitist strategies may however dampen this effect) and the amplification of chances for adding new variants to non-dominated solutions set, potentially having the desired characteristics. It is then a strategy biased towards exploration a whole solution space as opposed to the exploitation of areas containing variants of the desired characteristics. The choice of strategy depends on a specific problem at hand and should be preceded with testing calculations. The author proposes to use the domination of dominance count in cases where the number of variants created during a simulation run is a small fraction of a number all variants which may be created within a given optimization model, and the requirements concerning the accuracy obtained in pinpointing the compromise solution are not very high. In such a case intensive exploration of solution space boosts the chances for finding a satisfactory solution. However, in case the user emphasizes the requirement for the accuracy of a located non-dominated solution approximating the Pareto front, rank component domination is going to be recommended.

The proposed combined fitness multi-objective evolutionary algorithm (CFMOEA) using in the selection process a scalar objective function expressed in the Eq. 13, which is a weighted sum of elements representing: (1) optimization criteria, (2) ranking in relation on dominance, (3) dominance count, (4) distance from the asymptotic solution, and (5) the degree of constraint violation, shall be tested in the next parts of the paper using a fast passenger-car ferry structure as an example of a multi-objective optimization.

The computer code used for multi-objective optimization of ship structures with combined fitness function has been built on the basis of a software package for multi-objective optimization while supplementing it with a series of calculation procedures providing for combined fitness multi-objective optimization of the ship hull structure. The most important of them are the procedures for: (1) encoded of genotype and generation of the ship hull structure variants, (2) analysis of constraints, (3) analysis of feasible solutions set with regard to dominance relation, (4) control of domination, including the determination of domination attribute values, dominance rank and dominance count, and the building of non-dominated solutions set approximating the Pareto set, (5) calculations of the distance of feasible solutions to the asymptotic one, and (6) calculation of combined fitness function values, while accounting for the domination attributes and distance to asymptotic solution of feasible solutions. Block-type diagram presenting the main concepts of developed calculation software is shown in Fig. 11.

The code carries out the calculations automatically, starting from an input data set prepared by the user, which the following data regard to multi-objective optimization strategies: (1) values of switches specifying the selection of strategy which is going to be used for the calculation of fitness function values basing on optimization criteria values, and (2) the values of weight coefficients assigned to combined fitness function components. In the computer code a population of individuals of a fixed size is randomly generated. Each individual is characterized by a string of bits and represents one possible solution to the ship structure. Each new created variant of solution (an individual being a candidate to the progeny generation) is analyzed by the pre-processor. In the pre-processor binary strings of chromosomes (genotypes) are decoded into the corresponding strings of decimal values representing design variables (phenotypes). Then for the actual values of the design variables defining spatial layout of the structural elements (topology) and
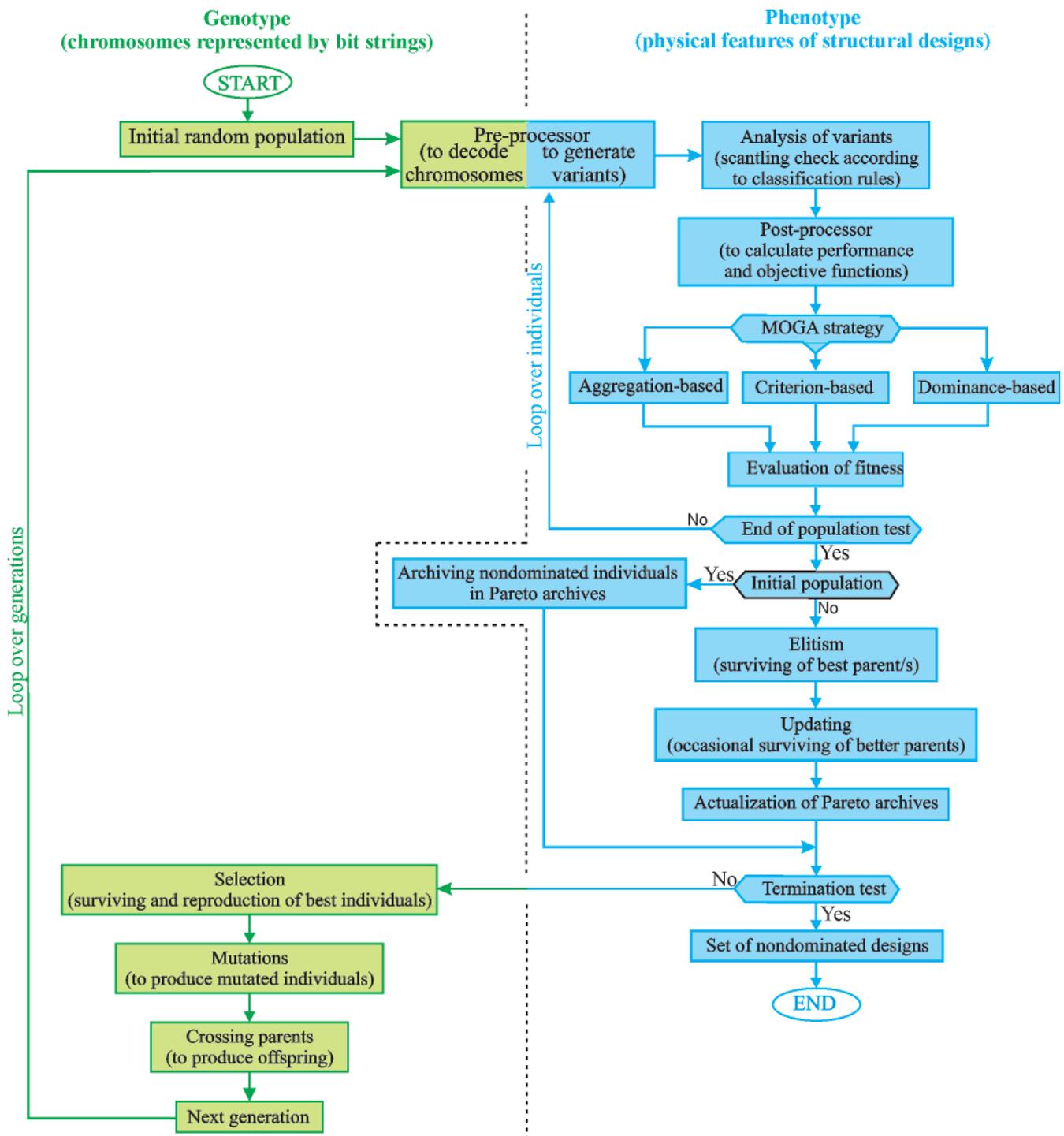
**Genotype**
**(chromosomes represented by bit strings)**

**Phenotype**
**(physical features of structural designs)**

START

Initial random population

Pre-processor
(to decode chromosomes to generate variants)

Analysis of variants
(scantling check according to classification rules)

Post-processor
(to calculate performance and objective functions)

MOGA strategy

Aggregation-based     Criterion-based     Dominance-based

Evaluation of fitness

End of population test — No

Yes

Archiving nondominated individuals in Pareto archives — Yes — Initial population

No

Elitism
(surviving of best parent/s)

Updating
(occasional surviving of better parents)

Actualization of Pareto archives

Loop over individuals

Loop over generations

Selection
(surviving and reproduction of best individuals) — No — Termination test

Yes

Set of nondominated designs

Mutations
(to produce mutated individuals)

Crossing parents
(to produce offspring)

Next generation

END

*Fig. 11. Diagram presenting the main concepts of the developed computer software*
*for the multi-objective optimization of seagoing ship structures, based on a CFMOGA*

their scantlings it is checked whether the actual configuration complies with the rules of the classification society. In the next step performance of solution is evaluated and it is checked whether the variant meets the constraints.

The domination relation is checked over a set of individuals considered to be feasible . The set of non-dominated individuals, updated for each generation, constitutes an approximation of Pareto front. At the end, the value of fitness function, proportional to the individual selection probability and its participation in the genetic operations of the next generation, is calculated. Calculation of combined fitness function value takes into account the of multi-objective optimization strategy, selected by the user. The computer software developed herein implements three selection strategies, Fig. 6: (1) strategy

based on the aggregation of vector objective function, with appropriate weight coefficients (fixed values for (1a) strategy or random values for (1b) strategy), (2) strategy executing the selection of variants on the basis of randomly selected single optimization criteria, and (3) strategy based on selection with respect to attributes of Pareto domination.

The code analyses the feasible individuals of successive generations, selects the non-dominated solutions and this way builds the archive of non-dominated solutions (Pareto-optimal ones). The archive of non-dominated variants is an external set and the variants which are recorded in it do not take part in the „breeding" of successive generation solutions, unless they have managed to survive in successive populations in a natural way, remaining not destroyed and not excluded by genetic operators.

Process of building and evaluation of descendant populations of tested variants and the updating of non-dominated variants archive is repeated until a fixed number of generations used as a simulation halt condition is reached. After a simulation is completed, the non-dominated solutions preserved in the archive are recognized to be the sought solutions of a multi-objective optimization problem.

At the end the value of the fitness function is calculated which is used for ordering the variants necessary to starting of selection. Variants are ordered with respect to this value. Knowing adaptation of each variant the random process is restarted to select variants of the successive progeny generation.

After selection the code determines randomly which genes of these whole population will mutate. This population is then mutated where small random changes are made to the mutants to maintain diversity. After that the mutate pool is created. Then decision is made how much information is swapped between the different population members. The mutated individuals are then paired up randomly and mated in the process commonly known as crossover. The idea is to derive better qualities from the parents to have even better offspring qualities. That is done by creating, with fixed probability, „cutting points" and then the parts of the chromosomes located between "cuts" are interchanged. The mating process is continued until the full population is generated. The resulting population member is then referred to as an offspring. The newly generated individuals are then re-evaluated and given fitness score, and the process is repeated until it is stopped after a fixed number of generations.

All genetic parameters are specified by the user before the calculations. The population size, number of design variables and number of bits per variable, the total genome length, number of individuals in the population are limited by the available computer memory.

The code carries out the calculations automatically, starting from an input data set prepared by the user, which has been supplemented with the following data with regard to its aggregated multi-objective predecessor: (1) values of switches specifying the selection of strategy which is going to be used for the calculation of fitness function values basing on optimization criteria values, and (2) the values of weight coefficients assigned to combined fitness function components and depending on dominance attributes and distance to asymptotic solution. The domination relations allowing for the determination of dominance rank and dominance count values for particular feasible variants are checked over a set of variants considered to be feasible. The set of non-dominated variants, updated for each generation, constitutes an approximation of the Pareto-optimal front. At the end, the value of fitness function, proportional to the variant selection probability and its participation in the genetic operations of the next generation, is calculated. Calculation of fitness function value takes into account the strategy of combined fitness multi-objective evolutionary algorithm, selected by the user. The computer software developed herein implements several multi-objective genetic algorithm selection strategies, Fig. 6. First of all, optimization is possible only with regard to optimization criteria: (1) strategy based on the aggregation of vector objective function, with appropriate weight coefficients, (2) strategy executing the selection of variants on the basis of randomly selected single optimization criteria. By assuming different from zero values of weight coefficients it is possible in optimization process to also take into consideration domination attributes as well as distance to asymptotic solution.

The values of parameters controlling the genetic operators are specified by the user before the start of simulation. Proper setting of these values is very important and requires extensive experience on the part of the user, but it is crucial for the attainment of the desired calculations convergence expressed in (1) quality of found solutions, (2) rapidity of their finding, and (3) required computation resources.

In case of multi-objective optimization problem we have to consider how to collect and present the information about the determined non-dominated solutions (Pareto-optimal) and how to archive them. It is commonly accepted to graphically present all the feasible solutions as points in the objective space. Only a part of them are going to be non-dominated (dominating, Pareto-optimal) solutions, and their set is going to be called a non-dominated set or a Pareto front containing trade-off solutions (actually we know this is going to be a set of non-dominated solutions approximating the Pareto set).

Non-dominated solutions produced during the simulation are recorded in a separate set (file) which is continuously supplemented and updated during the simulation. The solutions collected in the non-dominated solutions set may be dealt with in two possible ways: (1) set membership has no influence on the selection of individuals (egalitarian strategy) (2) individuals from the set (non-dominated solutions) enjoy a guaranteed participation in selection (elite strategy). Egalitarian strategy has been adopted in the underlying paper.

It is a well-known fact that in case many optimization criteria are used, it is going to be impossible to find a single best solution, as such a solution does not exist. In practice however the user awaits automatic or quasi-automatic determination of a single solution or a few solutions, which could be taken as a solution of the problem. Moreover, users are accustomed to the monitoring of evolution of a single value, which lets them evaluate the correctness of the calculation run, the convergence of solutions and the quality of solutions being found. In single-objective cases it is natural to monitor the values of fitness function and optimization criterion. In case of multi-objective optimization simultaneous evaluation of the evolving criteria is difficult to realization and interpretation. In order to alleviate this problem, the author has used the concept of ideal or utopia solution, well-known in literature [Cohon (1978)], [Stadler (1988)], [Statnikov and Matosov (1995)], see Fig. 12. In the generally accepted understanding an ideal point refers to the lowest values of all criteria analyzed singly and not together. It means that if $f^*_1, f^*_2, \ldots f^*_s, \ldots, f^*_s$ will be used to denote the individual minima of each respective objective function, and the ideal solution (in objective space) is defined as $f^* = [f^*_1 \; f^*_2 \ldots f^*_s \ldots f^*_s]^T$. As $f^*$ simultaneously minimizes all objectives, it is an ideal solution that is rarely feasible. In such a case however it is then possible to locate a solution closest to the ideal point (nearest to the ideal solution):

$$f^{\approx}_m(x) = [f^{\approx}_{1;m}(x) \; f^{\approx}_{2;m}(x)]^T \qquad (14)$$

with the concept of closeness being understood here usually in the sense of Euclidean metrics where $f^{\approx}_m(x)$ is the closest solution found in the m generation.

The concept of non-dominated feasible solution nearest to the ideal objective $f^{\approx}(x)$ is sufficient to find a single solution which may be considered to be „the best" solution of a multi-objective optimization problem. It is however inappropriate for the monitoring of evolution of non-dominated solution in the direction of theoretically lowest values of optimization criteria $f_i \rightarrow 0!$ as moving the set of non-dominated solutions in the desired direction may also take place with unchanging distances of these solutions from ideal solution, Fig. 13. The author has
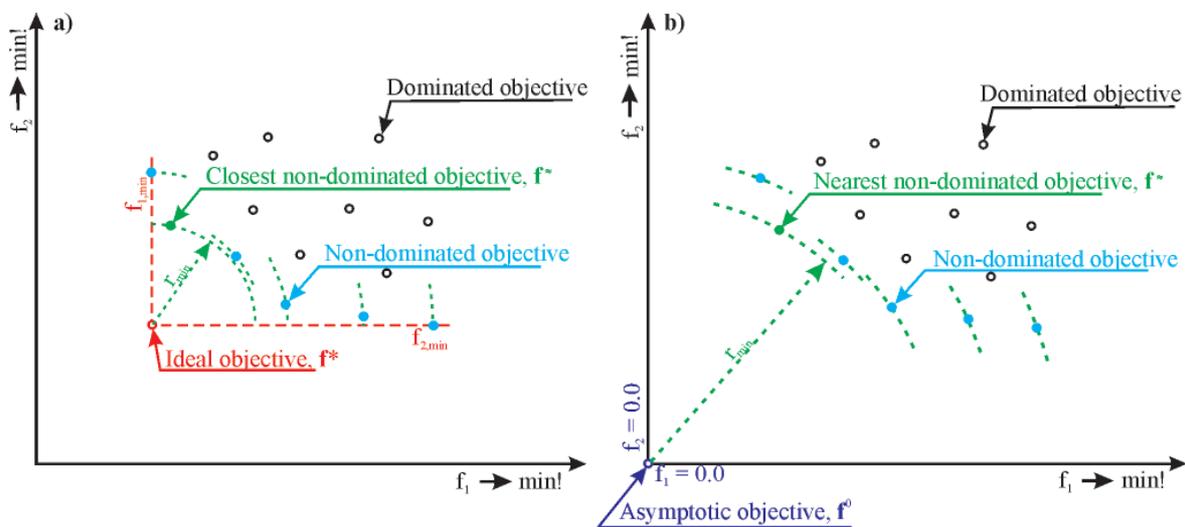
**Fig. 12. a)** *Graphical illustration of a non-dominated solution nearest to the ideal one f\*, for the case of min! type criteria.*
**b)** *Graphical illustration of the concept of a non-dominated feasible solution nearest to a asymptotic objective f⁰, in case of min! type criteria*
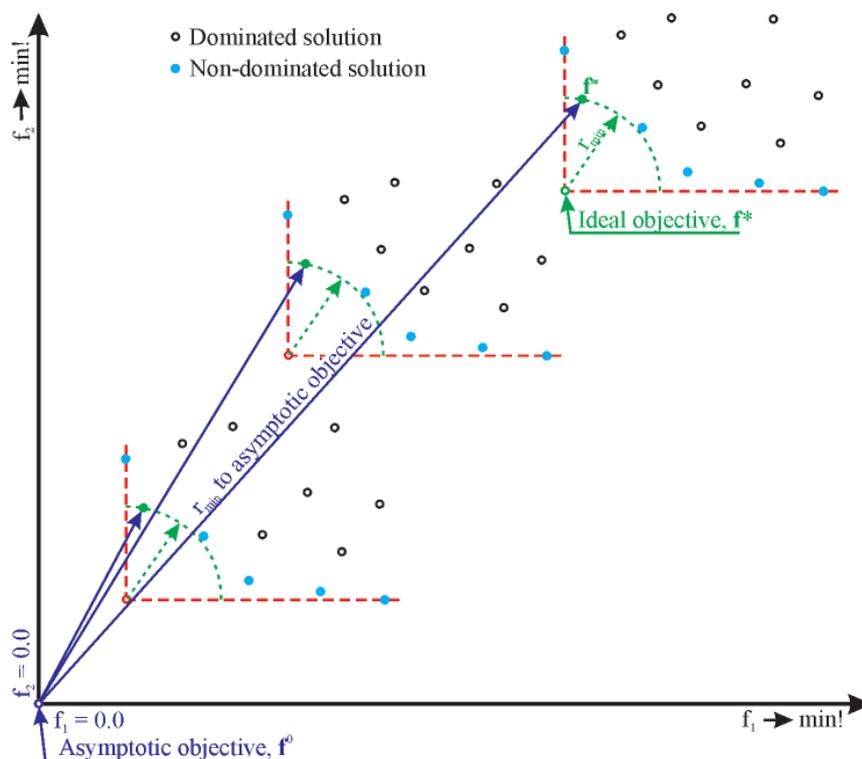


*Fig. 13. Graphical illustration of the movement of solutions nearest to the ideal solution f\* in the direction of asymptotic solution during the evolution, in case of min! type criteria; for constant distances from the ideal solution the distance from asymptotic one changes (evolves) instead*

then introduced a concept of a asymptotic objective (asymptotic solution) $f^0$, which represents an objective corresponding to asymptotical values of optimization criteria: $f_1 \rightarrow 0!$ $f_2 \rightarrow 0!$, Fig. 13. According to this definition, the location of ideal objective $f^*$ in the objective space is not fixed, Fig. 13, and this means that its location changes during simulation, so the distance to such a moving objective is not a good indicator of solution quality. Therefore, a definition of asymptotic objective $f^0$ has been adopted in such a way as to make it stationary. It is then necessary to have such a definition which would make the co-ordinates of this objective constant in the objective space, which means fixed values of optimization criteria. It is then assumed that in case of min! type criteria the respective objectives (co-ordinates) are equal to zero. However, in case of max! type criteria the user shall set the value of these objectives (co-ordinates) as some known from experience values of respective criteria which are impossible to attain, but which are

going to be approached asymptotically by the non-dominated solutions. After defining the asymptotic objective $f^0$ let's assume that as the nearest solution $f^*(x)$ we mean nearest solution with regard to the asymptotic objective $f^0$.

## REFERENCES

The complete list of references will be insert in the end of third part of the paper.

_____

**CONTACT WITH THE AUTHOR**
Zbigniew Sekulski, Ph. D.
West Pomeranian University of Technology, Szczecin
Faculty of Marine Technology
Al. Piastów 41
71-065 Szczecin, POLAND
e-mail: zbigniew.sekulski@zut.edu.pl

# Modelling of seakeeping qualities of open-top container carriers in the preliminary design phase

**Tomasz Cepowski**, Assoc. Prof.
Szczecin Maritime University

## ABSTRACT

*In this paper are presented problems of modelling seakeeping qualities of open-top container carriers in the preliminary design phase. Approximations of accelerations and occurrence rate of green water ingress to holds are presented in function of main ship design and wave parameters. The approximations have been elaborated by applying theory of artificial neural networks in the wide range of ship hull dimensions and forms, while ship motion and wave parameters have been limited to real operational conditions described by means of the so called operational scenario. Such approach has made it possible to reach high accuracy of approximation and simple structure of mathematical model simulatneously.*

## INTRODUCTION

In the work [4] a systematization of transport ships has been proposed from the point of view of their seakeeping qualities. In it, a.o. a group of ships having design features which increase their susceptability to weather conditions can be distinguished. To the group can be numbered the ships which usually operate in normal weather conditions however it should be expected that for some time of service they would sail in heavy weather conditions. From the point of view of seakeeping modelling problems the group of ships is characterized by design features which elevate their susceptibility to weather conditions with a view of safety of ship itself or its cargo. To the group can be numbered the following kinds of ships intended for the carrying of solid cargo vertically loaded through deck hatches, and fitted with open main deck which is:

1. either closed by means of hatch covers:
   a. universal wide-hatch bulk carriers,
   b. general cargo ships,
   c. container carriers,
   d. con-ro ships,
2. or not closed by means of hatch covers:
   a. open-top container carriers,
   b. con-ro open-top ships.

Main design problems for the group of ships concern ensuring an appropriate transport capacity as to amount of transported cargo, ship speed and required operational autonomy. In view of specificity of the transported cargo and design features of the ships in question, for them a.o. seakeeping qualities and in particular the necessity of:

- minimization of accelerations which may cause shifting or damaging the cargo and hull structure failure,
- minimization of amount of green water entering the deck which may lead to its ingress to holds constitute technical constraints.

The open-top container carrier represents the group of ships. Specific configuration of holds of open-top container carrier solves many general design problems characteristic for classical containership, such as:

- the problem of ship stability which results from high location of its centre of gravity,
- the problem of arrangement of containers which results from necessity of their arranging on the deck is such a way as to avoid necessity of displacing a number of them to make it possible to remove certain ones out of holds,
- the problem of proper securing containers located on the deck in the highest tiers against dynamic loads due to accelerations which can lead to their losing,
- the problem of proper design of hatch covers which transfer large loads and are to be highly reliable.

However on the other hand, the lack of hatch covers and carrying the containers in one stock generates additional problems:

- those resulting from large weight of stock of containers:
  - double bottom of open-top containerships is to be more strong,
  - large mass inertia forces which result from vertical and transverse accelerations may lead to the damaging of containers or hull structure.
- green water entering the deck and its subsequent ingress to holds.

On 26 February 2007 on the open-top containership „Annabella" during its voyage to Helsinki an accident associated with collapsing containers in an open hold, took place [7]. Consequently, 7 containers located in the hold close to ship's plane of symmetry, sufferred damage. As results from the report [7] the formal cause of the event was an inappropriate arrangement of containers. However its physical cause was an incorrect value of initial metacentric height that led to accelerations whose resultants generated large mass inertia forces.

Another formal problem associated with lack of hatch covers was non-fulfilment of the requirements resulting from International Convention on Loadlines, which deal with construction of freeboard deck. Because of the non-fulfilment of the requirements the Maritime Safety Committee adopted, in 1993, additional guidelines dealing with construction of open-top containerships. In particular there were defined the requirements concerning seakeeping model tests, which dealt with a.o. the following, [6]:
- all the model tests are to be performed under the assumption on irregular waves described by the wave spectrum of JONSWAP, Pierson-Moskovitz, or Bretschneider type, with the following parameters: the wave significant height $H_S = 8.5$ m and the most unfavourable characteristic wave frequencies,
- all the model tests are to be performed at least in the following conditions:
  - wave encounter angles:
    - following wave $\beta = 0°$,
    - head wave $\beta = 180°$,
    - oblique following wave $\beta = 45°/315°$,
    - oblique head wave $\beta = 135°/225°$,
    - beam wave $\beta = 90°/270°$,
  - ship speeds:
    - maximum speed on head or oblique head wave,
    - minimum speed on following or oblique following wave,
    - zero-speed (the so called dead ship condition) on beam wave,
- the model tests should be conducted for the ship in maximum draught load conditions.

Amount of water flooding open-top containership holds depends a.o. on: freeboard height, wave parameters, ship motion parameters, hydromechanic parameters as well as geometrical ship hull parameters such as: underwater hull form and dimensions (ship length in particular), draught, form of aft- and fore- hull parts, deck shape and bulwark structure [1, 3, 5, 8].

## DESIGN PARAMETERS WHICH AFFECT SEAKEEPING QUALITIES OF OPEN-TOP CONTAINER CARRIERS

With a view of the above specified aspects among the crucial seakeeping qualities of open-top container carriers the following can be numbered:
- phenomenon of green water entering the deck, which influences green water ingress to holds,

- transverse and vertical accelerations which may result in generating large mass forces leading to:
  - crushing containers placed in the lower part of stocks,
  - hull structure damage.

Occurrence of the above mentioned effects of wave action depends a.o. on the main geometric hull parameters of containership, which are modeled in the parametric ship design phase.

Out of the above mentioned parameters the following ones affect phenomenon of green water entering the deck:
- the freeboard height Fb,
- the ship length L,
- the block coefficient of underwater hull part, CB,
- the block coefficient of fore underwater hull part, CBF.

According to [3, 8] the above mentioned are crucial for occurrence of phenomenon of green water entering the deck, hence they may also greatly influence green water ingress to holds.

And, as results from [4] the phenomenon of green-water ingress to holds is affected by the following:
- freeboard height,
- middlebody length,
- block coefficients of fore and aft part of underwater hull.

And, vertical and transverse accelerations are affected by the following design parameters:
- the geometric underwater hull parameters:
  - the ship length L,
  - the ship length/breadth ratio L/B,
  - the ship breadth/draught ratio B/d,
- height of ship gravity centre, which influences initial transverse metacentric height and initial longitudinal metacentric height.

As results from [7], the loads generated in lashing system of containers and the event of crushing the containers could be affected by such factors as:
- initial transverse metacentric height,
- unfavourable wave parameters (wave period).

It is not possible today to take into account the above mentioned seakeeping qualities in the parametric design phase as there are no simple, and simultaneously accurate, analytical functions which could make it possible to predict the above mentioned seakeeping qualities on the basis of:
- main geometric parameters of underwater hull,
- parameters which describe load condition,
- parameters of waves occurring in a assumed water area of ship operation.

Formal requirements for the designing of open-top containerships impose the necessity of performing ship model tests aimed at determining amount of water flooding holds without hatch covers.

To perform such tests is not possible in the parametric design phase but only in that design phase when dimensions of ship hull and its general form are already konwn. The model tests of open-top containerships, presented for instance in [1], make only possible to compare given design solutions from the point of view of freeboard height or a form of selected parts of hull. The model tests do not account for the influence of main geometric dimensions of hull as well as wave parameters on occurrence rate of green-water entering the deck.

Formal formulae for calculating vertical and transverse accelerations to determine additional loads exerted to

containers, given in classification rules, are too general and do not take into account the following:
- influence of the parameters which describe ship loading conditions, such as ship weight or height of its gravity centre,
- influence of wave parameters.

Informal design guidelines are limited first of all to the tests dealing with green water entering the deck of a given ship and do not find any application in the parametric design phase of open-top containership.

## AIM AND METHOD OF THE RESEARCH

In the work [4] have been presented the methods for predicting green water ingress to holds of open-top containership in its initial design phase, based only on hull form parameters and set hull dimensions. And, the research presented in this publication has been aimed at prediction of accelerations and occurrence rate of green water ingress to holds with a view of hull form and dimensions.

It was assumed that the above defined aim can be reached by analyzing results of numerical calculations of ship motions in waves in conventional operational conditions described by operationaal scenarios. The research method scheme is presented in Fig. 1.

To make the above mentioned aim possible an additional aim concerning determination of an essential operational scenario for open-top containership, was formulated.

In approximating maximum significant amplitudes and occurrence rates of the assumed seakeeping qualities the following was taken into account:
- set of explaining variables, which contained:
  - hull design parameters,
  - wave parameters,
- set of to-be-explained variables, which comprised:
  - occurrence rates of green water ingress to holds,
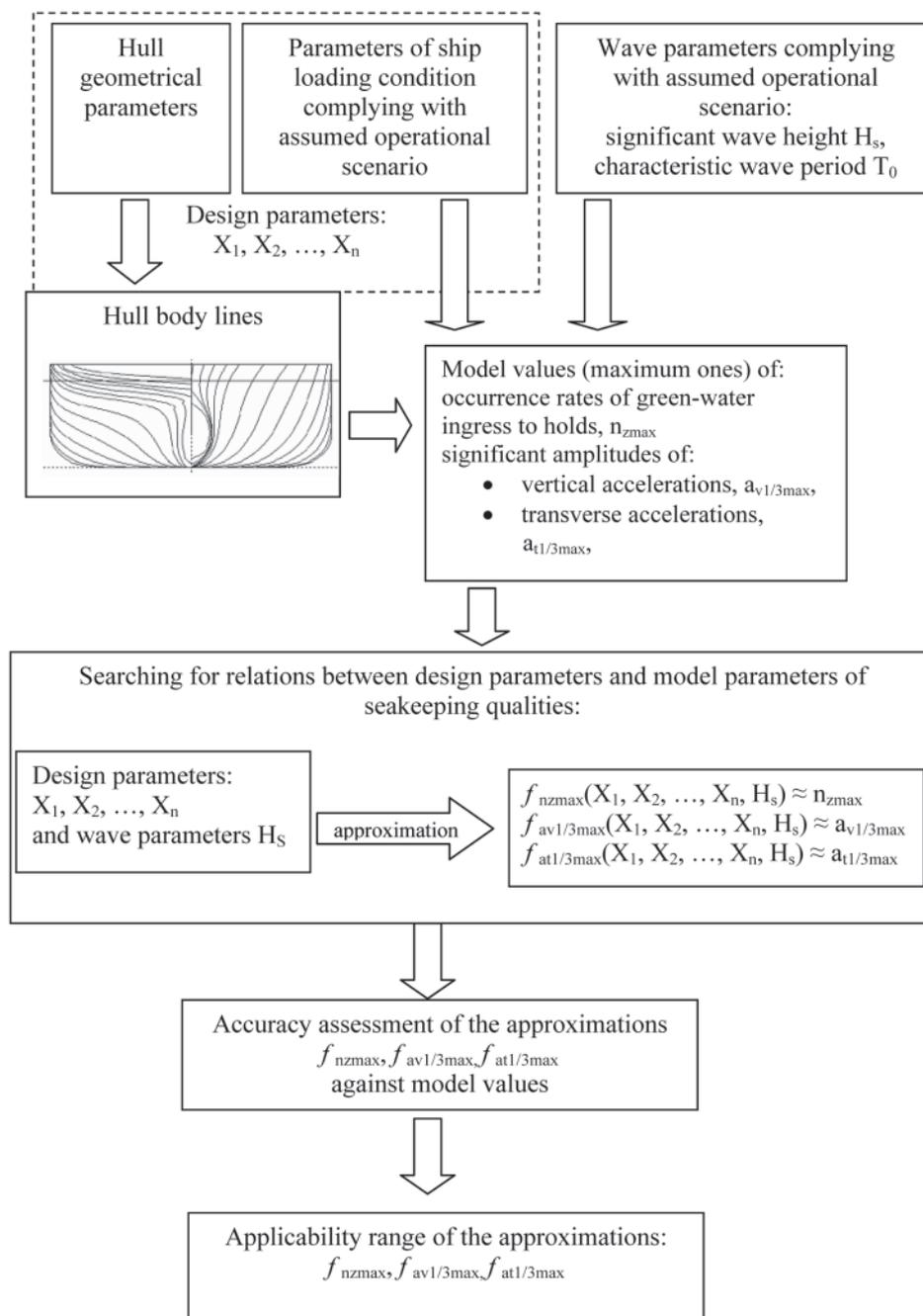  - significant amplitudes of transverse and vertical accelerations occurring in holds.



*Fig. 1. Algorithm of approximation process of assumed seakeeping qualities in function of design parameters of open-top containerships*

Values of the above specified variables were affected by assumptions resulting from the selected operational scenario.

Wave period greatly influences effects of wave-ship iteraction. Investigations carried out on FPSO ships demonstrated that the assumption of the characteristic wave period value equivalent to the so called 100-year wave, in design calculations, resulted in misfit of main dimensions of such ships to real wave conditions and hazardous escalation of phenomenon of green water entering the deck (Fig. 2).
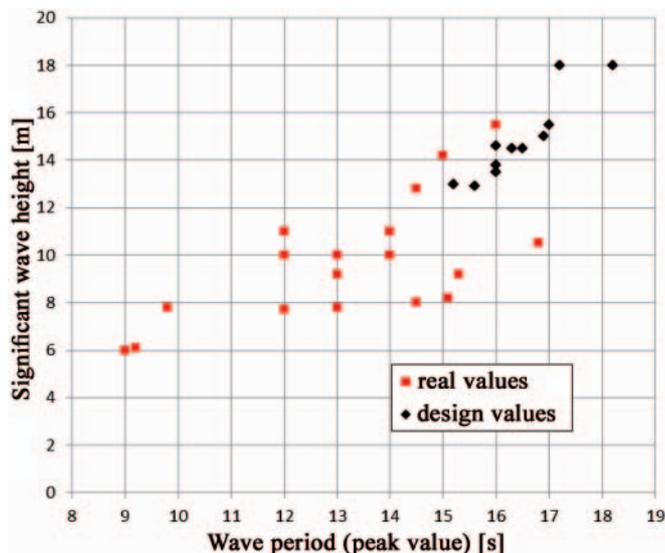


***Fig. 2.*** *Values of wave parameters at which hazardous green water entering the deck occurs, obtained for both design and real conditions [2]*

Therefore in order to avoid the problem in this research it was decided to approximate maximum values of the assumed seakeeping qualities by taking into account the characteristic wave period. The so elaborated approximations may provide over-estimated values but account for unfavourable influence of wave period on ship motions. An advantage achieved from such solution is elimination of wave period from the set of explaining variables and simplification of the approximation model.

## OPERATIONAL SCENARIO

For modeling seakeeping qualities of open-top containership it was assumed that the ship is intended for operating in the conditions described in [6]. To this end the following operational scenario was assumed:

The open-top containership sails, under its maximum draught, with the speed of 10 m/s in stormy weather. In view of necessity of conducting cargo segregation the hold no. 1 is fitted with hatch covers and the remaining ones are without covers. The ship is subjected to action of irregular statistical wave of the parameters complying with [6]. The wave has various characteristic periods and acts from various directions, causing the following:

- maximum occurrence rate of green water ingress to the open-top holds,
- maximum amplitudes of transverse and vertical accelerations acting onto containers placed in stocks in open-top holds.

The above mentioned scenario can be described by the following parameters:

- the ship speed v = 10 m/s,
- the wave encounter angle β ranging from 0° to 180°,
- the JONSWAP wave spectrum, the amplification factor γ = 3.3,
- the significant wave height Hs ranging from 1 to 8.5 m,
- the characteristic wave period which causes occurrence of maximum amplitudes of transverse and vertical accelerations.

## HULL DESIGN PARAMETERS

Open-top containerships are relatively small and few in number. Tab. 1 presents the set of main dimensions and their ratios of the representatives of the group of ships in question.

As results from Tab. 1, the design parameters of open-top containerships are contained within the following ranges:
- $L_{pp}/B$: 5 ÷ 7.2,
- $B/d$: 2.5 ÷ 4.7,
- $H/d$: 1.46 ÷ 2.11,
- $L_{pp}Bd$: 7182 ÷ 125 346 m³.

For the modelling of seakeeping qualities, were assumed the following ranges of values of hull geometric parameters, covered by the above specified ranges:
- for $L_{pp}Bd$: 7000 ÷ 130 000 m³,
- for $L_{pp}/B$: 5 ÷ 8,
- for $B/d$: 2.3 ÷ 4.7,
- for $H/d$: 1.3 ÷ 3.3,

as well as the underwater hull forms shown in Fig. 3. Their global coefficients are presented in Tab. 2.

On the basis of the above described assumptions the set of 2816 variants of hull forms and dimensions was elaborated. Value of the initial metacentric height GM was assumed, for each of the variants, from the range of 0.4 ÷ 1.8 m, every 0.2 m.

***Tab. 1.*** *Main dimensions and their ratios of representatives of open-top containerships*

| Name | $L_{pp}$ [m] | B [m] | d [m] | H [m] | H/d | $L_{pp}/B$ | B/d | $L_{pp}Bd$ [m³] |
|---|---|---|---|---|---|---|---|---|
| **178 type** | 155.40 | 26.80 | 9.60 | 14.00 | 1.46 | 5.80 | 2.79 | 39981 |
| **168 type** | 124.41 | 22.50 | 8.70 | 14.00 | 1.61 | 5.53 | 2.59 | 243536 |
| **161 type** | 136.80 | 25.90 | 9.61 | 14.00 | 1.46 | 5.28 | 2.70 | 34049 |
| **Atlantic Lady** | 160.00 | 28.80 | 9.00 | 16.80 | 1.87 | 5.56 | 3.20 | 41472 |
| **NB1020 Reestborg** | 124.40 | 20.00 | 6.30 | 12.60 | 2.00 | 6.22 | 3.17 | 15674 |
| **Nedlloyd Hong-Kong** | 265.00 | 37.75 | 12.53 | 23.25 | 1.86 | 7.02 | 3.01 | 125346 |
| **Norasia Hong-Kong** | 229.50 | 32.24 | 11.00 | 23.00 | 2.09 | 7.12 | 2.93 | 81389 |
| **Bell Pioneer** | 107.18 | 16.92 | 5.93 | 12.52 | 2.11 | 6.33 | 2.85 | 10753 |

**Tab. 2.** *Assumed variants of underwater hull form described by: **CB** – underwater hull block coefficent, **CB(L)** – underwater hull cylindrical coefficient, **CB(V)** – underwater hull vertical prismatic coeffcient, **CBA** –underwater hull aft part block coefficient, **CBF** – underwater hull fore part block coefficient, **CWL** – waterplane coefficient, **CM** – midship section coefficient, **XF** – percentage distance of waterplane geometric centre from aft perpendicular related to ship length between perpendiculars, **XB** – percentage distance of buoyancy centre from aft perpendicular related to ship length between perpendiculars*

| No. | CB [-] | CB(L) [-] | CB(V) [-] | CBA [-] | CBF [-] | CWL [-] | CM [-] | XF [%] | XB [%] |
|-----|--------|-----------|-----------|---------|---------|---------|--------|--------|--------|
| 1   | 0.58   | 0.59      | 0.78      | 0.60    | 0.55    | 0.73    | 0.97   | 45.82  | 48.49  |
| 2   | 0.55   | 0.59      | 0.76      | 0.43    | 0.48    | 0.73    | 0.94   | 44.59  | 46.92  |
| 3   | 0.70   | 0.74      | 0.90      | 0.59    | 0.59    | 0.78    | 0.94   | 50.02  | 50.10  |
| 4   | 0.58   | 0.61      | 0.85      | 0.53    | 0.46    | 0.68    | 0.97   | 47.81  | 47.94  |
| 5   | 0.63   | 0.67      | 0.76      | 0.64    | 0.58    | 0.82    | 0.97   | 40.66  | 45.65  |
| 6   | 0.61   | 0.64      | 0.79      | 0.53    | 0.65    | 0.78    | 0.98   | 44.59  | 47.52  |
| 7   | 0.57   | 0.60      | 0.81      | 0.58    | 0.55    | 0.71    | 0.98   | 47.35  | 48.12  |
| 8   | 0.62   | 0.65      | 0.73      | 0.63    | 0.62    | 0.85    | 0.98   | 41.41  | 48.71  |
| 9   | 0.56   | 0.59      | 0.80      | 0.56    | 0.53    | 0.70    | 0.97   | 45.20  | 47.66  |
| 10  | 0.58   | 0.61      | 0.77      | 0.61    | 0.52    | 0.75    | 0.98   | 43.34  | 46.76  |
| 11  | 0.54   | 0.58      | 0.80      | 0.55    | 0.45    | 0.67    | 0.93   | 43.51  | 45.99  |
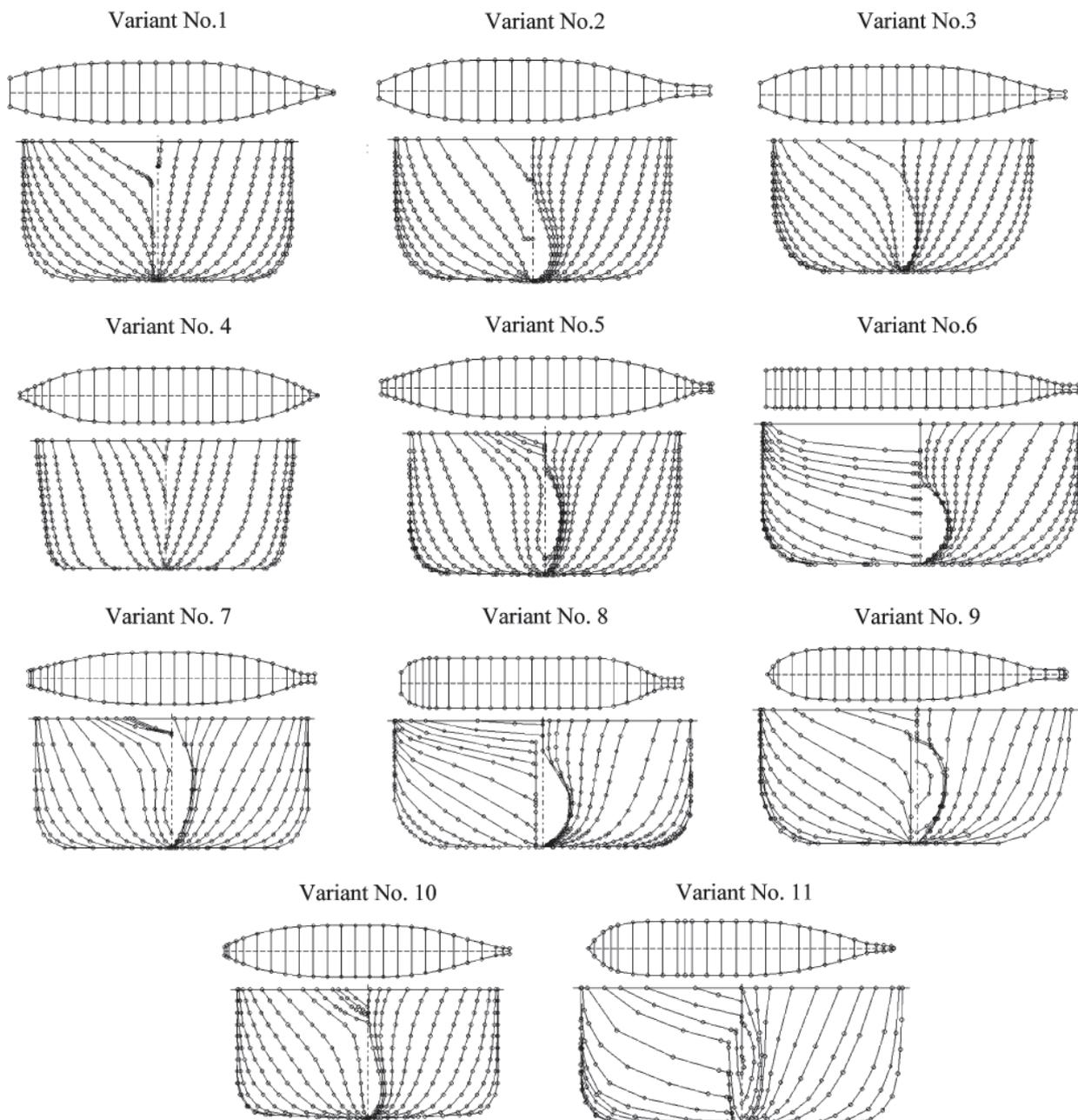


**Fig. 3.** *Assumed variants of underwater hull body lines of containerships*

## WAVE PARAMETERS

For calculations of model values of seakeeping qualities the wave parameters resulting from the assumed operational scenario, were taken as follows:
- the JONSWAP wave spectrum, the amplification factor $\gamma = 3.3$;
- the significant wave height $H_s = 4.5$; $6.5$; $8.5$ m;
- the characteristic wave period $T_0$ in the range from 2 to 20 s, every 0.5 s.

In the operational scenario it was assumed to take into account occurrence of maximum values of seakeeping qualities in function of the characteristic wave period. Therefore to approximate significant amplitudes and occurrence rates of the assumed seakeeping qualities, only the significant wave height $H_s$ was selected out the all wave parameters, in order to serve as the explaining variable.

## MODEL VALUES OF SEAKEEPING QUALITIES

The following to-be-explained variables were assumed in this research:
- the occurrence rates of green water ingress to holds, $n_{zmax}$, in the point A (acc. Fig. 3),
- the significant amplitudes of transverse accelerations, $a_{t1/3max}$, and vertical accelerations, $a_{v1/3max}$, in the point B of the hold,

In Fig. 4 and Tab. 3 is given location of the points in which the above mentioned seakeeping qualities have been calculated.
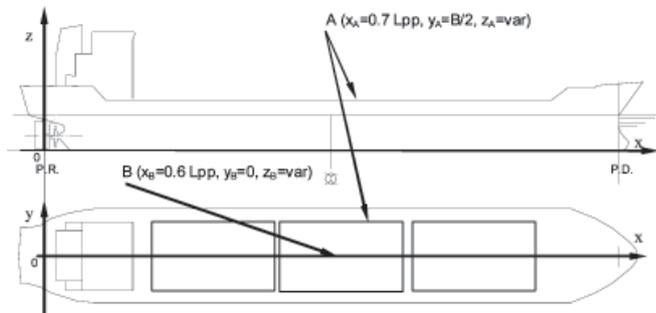


***Fig. 4.*** *Location of the points in which the occurrence rate of green water ingress to holds (the point **A**) and transverse and vertical accelerations (the point **B**) were calculated*

***Tab. 3.*** *Coordinates of the points for which model values of seakeeping qualities were calculated, where: **$L_{pp}$** – ship length b.p., **B** – ship breadth, **d** – ship draught*

| Seakeeping quality | Point | Coordinates of the point | | |
|---|---|---|---|---|
| Green water ingress to holds | A | $x_A = 0.7\, L_{pp}$ | $y_A = B/2$ | $z_A = (1{,}3 \div 3.3)d$ |
| Transverse and vertical accelerations | B | $x_B = 0.6\, L_{pp}$ | $y_B = 0$ m | $z_B = (0 \div 2)d$ |

Model values of the above mentioned seakeeping qualities were numerically determined by using SEAWAY software in the whole range of wave encounter angles. As resulted from the calculations the most unfavourable wave encounter angles from the point of view of the assumed seakeeping qualities

(and accounting for the most unfavourable wave periods) were the following:
- $\beta = 120°$ for green water ingress to holds,
- $\beta = 120°$ for transverse accelerations,
- $\beta = 90°$ for vertical accelerations.

## ELABORATION OF APPROXIMATION FUNCTIONS FOR THE ASSUMED SEAKEEPING QUALITIES OF OPEN-TOP CONTAINERSHIPS AND ASSESSMENT OF THEIR ACCURACY

At first approximations of model values of seakeeping qualities were searched for in the form of analytical functions by using linear regression in the domain of functions of the simplest forms. However it turned out that the elaborated approximations were not suffciently accurate. Just the use of the theory of artificial neural networks brought satsifactory results, that made it possible to elaborate the approximations:
- of maximum occurrence rates of green water ingress to holds:

$$n_{zmax} = H_S \cdot \dfrac{\dfrac{1}{1+e^{-[(X \times S + P) \times A - B]}} \times C - 0.876}{0.0037} \qquad (1)$$

where:
$n_{zmax}$ – maximum hourly occurrence rate of green water ingress to holds [Nr/h],
$H_S$ – significant wave height [m],
$X$ – vector of values of the design parameters:

$$X = [CBF, d, B, F_w, CB, F_b]$$

where:
CBF – underwater hull fore part block coefficient [-],
d – ship draught [m],
B – ship breadth [m],
CB – underwater hull block coefficient [-],
$F_b$ – freeboard height [m],
A – matrix of weighing walues:

$$\begin{bmatrix} 0.012 & -2.686 & -1.775 & -4.621 & -0.362 & -2.327 \\ 3.985 & -1.976 & -0.874 & 3.159 & 1.258 & 0.794 \\ -5.538 & 0.272 & 0.348 & -2.951 & 0.947 & -0.079 \\ -0.307 & -0.295 & -0.512 & -0.575 & 1.197 & -0.450 \\ 0.484 & -0.023 & -0.846 & -7.128 & -0.649 & -2.492 \\ -10.289 & -4.018 & -4.648 & -7.273 & 3.485 & -4.048 \end{bmatrix}$$

S – matrix of coeffcients:

$$\begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.118 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.046 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.57 \cdot 10^{-4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 6.219 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.167 \end{bmatrix}$$

B – vector of thershold values:

$$[3.091 \quad 0.712 \quad -0.115 \quad 1.240 \quad -2.197 \quad 0.834]$$

C – column vector of weighing values:

$$[4.715 \quad -1.716 \quad 1.412 \quad 5.511 \quad 0.879 \quad -1.277]$$

P – vector of displacement values:

$$[-2.250 \quad -0.511 \quad -0.749 \quad -0.196 \quad -3.338 \quad -0.667]$$

- of maximum significant amplitudes of transverse accelerations:

$$a_{t\,max} = H_S \cdot \dfrac{\dfrac{1}{1+e^{-[(X\times S+P)\times A-B]}} \times C + 2.667}{0.01723} \qquad (2)$$

where:

$a_{t\,max}$ – maximum significant amplitudes of transverse accelerations [m/s²],

X – vector of values of the design parameters:

$$X = [d,\ B,\ F_w,\ XF,\ V,\ L_{pp}/B,\ GM,\ z/d]$$

where:

XF – distance of waterplane geometric centre from aft perpendicular [m],
V – volume of underwater hull part [m³],
$L_{pp}$ – ship length b.p. [m],
GM – initial transverse metacentric height [m],
z – height of container gravity centre [m],
d – ship draught [m],
A – matrix of weighing values:

$$\begin{bmatrix}
1.726 & 0.466 & 1.452 & 1.314 & 0.341 \\
0.206 & -1.270 & -2.278 & 1.285 & 2.910 \\
0.317 & -0.830 & -1.312 & 0.895 & 1.734 \\
-0.834 & -0.638 & -0.703 & 0.027 & 1.733 \\
1.505 & -0.462 & -0.221 & -0.361 & -1.559 \\
0.325 & 0.746 & 1.012 & -0.203 & -0.762 \\
-1.638 & -1.101 & 2.993 & -1.194 & -2.260 \\
-0.843 & 1.748 & -2.496 & -1.105 & -1.683
\end{bmatrix}$$

S – matrix of coefficients:

$$\begin{bmatrix}
0.137 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.055 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2.96\,10^{-4} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.017 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 5.78\,10^{-5} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.333 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.714 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.135
\end{bmatrix}$$

B – vector of thershold values:

$$[-2.992\quad 1.744\quad 4.819\quad -1.105\quad -4.194]$$

C – column vector of weighing values:

$$[1.238\quad -0.486\quad -1.266\quad -1.011\quad -2.870]$$

P – vector of displacement values:

$$[-0.592\quad -0.892\quad -0.368\quad -0.792\quad -0.499\quad -1.765\quad -0.286\ 0]$$

- of maximum significant amplitudes of vertical accelerations:

$$a_{v\,max} = H_S \cdot \dfrac{\dfrac{1}{1+e^{-[(X\times S+P)\times A-B]}} \times C + 0.45}{0.0905} \qquad (3)$$

where:

$a_{v\,max}$ – maximum significant amplitudes of vertical accelerations [m/s²],

X – vector of values of the design parameters:

$$X = [CBF,\ F_w,\ XF,\ V,\ CB(V)]$$

where:

CB(V) – underwater hull prismatic coeffcient,
A – matrix of weighing values:

$$\begin{bmatrix}
1.645 & 0.065 & 0.000 & -0.048 \\
-0.919 & -1.420 & -3.301 & -7.784 \\
-0.571 & 1.225 & 1.681 & 0.724 \\
0.145 & -0.542 & 0.196 & -0.328 \\
-7.237 & -3.253 & 0.918 & 0.395
\end{bmatrix}$$

S – matrix of coefficients:

$$\begin{bmatrix}
5 & 0 & 0 & 0 & 0 \\
0 & 1.57\,10^{-4} & 0 & 0 & 0 \\
0 & 0 & 0.011 & 0 & 0 \\
0 & 0 & 0 & 2.96\,10^{-4} & 0 \\
0 & 0 & 0 & 0 & 6.01
\end{bmatrix}$$

B – vector of threshold values:

$$[3.122\quad 0.866\quad 1.659\quad 2.181]$$

C – column vector of weighing values:

$$[-2.722\quad 2.138\quad 1.506\quad 4.617]$$

P – vector of displacement values:

$$[-2.250\quad -0.196\quad -0.447\quad -0.238\quad -4.407]$$

Structures of the above described networks are presented in Fig. 8, and their statistical parameters – in Tab. 4. The parameters given in Tab. 4 concern the data from the set used to learn (amounting to 50% of cases) and that to validate the network (25% of cases), as well as the testing set (25 % of cases). The cases from the learning set were used to learn the network, and the cases from the validating set were used for tracking network operation errors, identifying the best network and stopping the learning process in the case of occurrence of over-learning symptoms. The testing set is not used at all during the learning process and is intended only for the independent determining of the network's effectiveness on completion of its designing process.
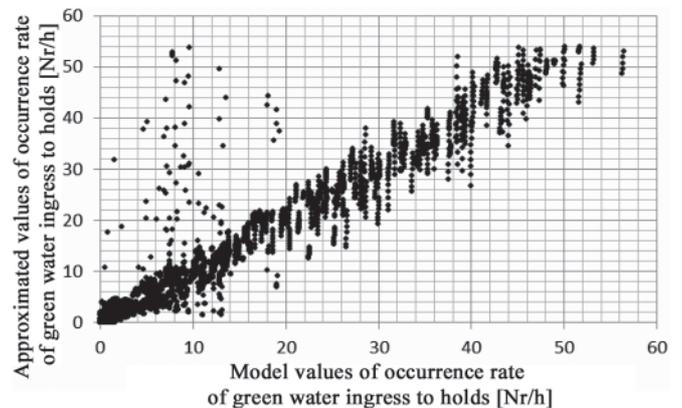


**Fig. 5.** *Comparison of approximations of maximum occurrence rate of green water ingress to holds with their model values*

In Fig. 5 through 7 are presented comparisons of approximated values with model ones for all the data contained

| | $n_{zmax}$ | | | $a_{tmax}$ | | | $a_{vmax}$ | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | U [Nr/h] | W [Nr/h] | T [Nr/h] | U [m/s²] | W [m/s²] | T [m/s²] | U [m/s²] | W [m/s²] | T [m/s²] |
| Standard deviation | 8.86 | 8.60 | 8.95 | 1.428 | 1.317 | 1.456 | 0.491 | 0.512 | 0.498 |
| Mean absolute error | 0.94 | 0.93 | 0.96 | 0 | 0.010 | -0.004 | -0.001 | 0 | -0.001 |
| Correlat-ion | 0.96 | 0.96 | 0.95 | 0.999 | 0.998 | 0.999 | 0.998 | 0.998 | 0.998 |

in the learning, testing and validating sets. As results from the comparisons in question, the elaborated approximations are characteristic of both simple structure and high accuracy.
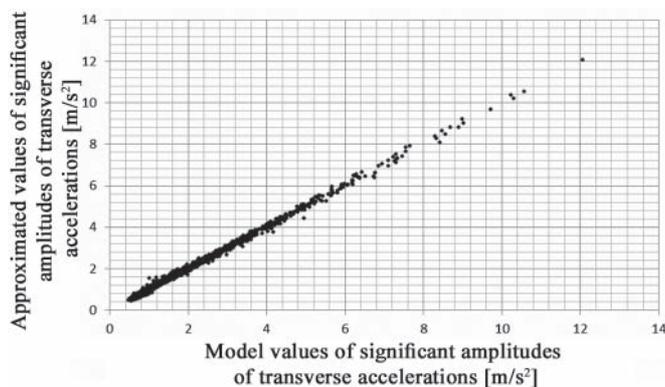


Fig. 6. Comparison of approximations of significant amplitudes of transverse accelerations with their model values
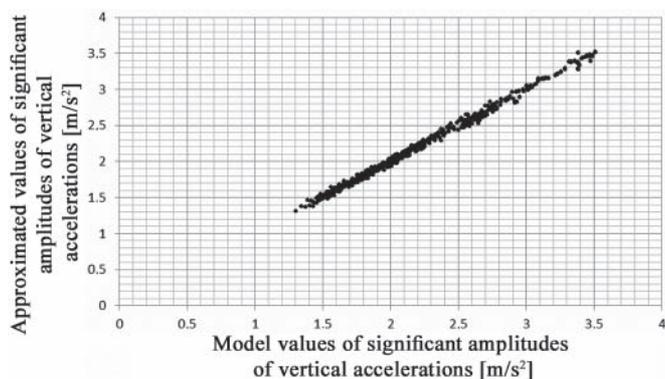


Fig. 7. Comparison of approximations of significant amplitudes of vertical accelerations with their model values
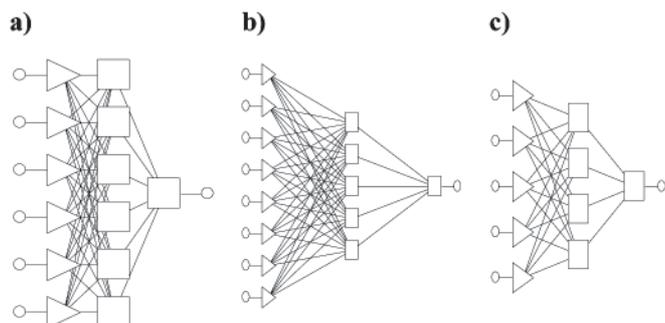


Fig. 8. Structures of the artificial neural networks which approximate:
a) maximum occurrence rate of green water ingress to holds,
b) maximum significant amplitudes of transverse accelerations,
c) maximum significant amplitudes of vertical accelerations

## SUMMARY

The elaborated approximations of seakeeping qualities of open-top containerships may be used in the preliminary design phase for:
- determination of crucial design parameters affecting green water ingress to holds, transverse and vertical accelerations,
- approximation of:
  - maximum occurrence rate of green water ingress to holds in the point A (acc. Fig. 4),
  - maximum significant amplitudes of transverse and vertical accelerations in the point B of the hold,
- assessment of ship seakeeping qualities,
- optimization of ship design parameters from the point of view of the assumed seakeeping qualities.

The elaborated approximations described by Eq. (1) through (3) appeared very accurate. In the research was assumed a wide range of values of ship design parameters as well as a limited range of the parameters which describe ship operational conditions according to ship operational scenario. Application of linear methods to predicting the assumed seakeeping qualities for the ships of the considered group has not led to reaching sufficient approximation accuracy. Hence the theory of artificial neural networks has been applied to this end.

The elaborated approximation functions are applicable to modelling seakeeping qualities in the following range:
- the assumed ranges of hull geometric parameters on the basis of which model values of seakeeping qualities were calculated, the following in particular:
  - the ship length b. p. / breadth ratio $L_{pp}/B$: from $5.3 \div 8.3$,
  - the ship breadth/draught ratio B/d: from $2.32 \div 4.72$,
  - the ship depth/draught ratio H/d: from $1.3 \div 3.33$,
  - the freeboard height Fb = H – d: from $3 \div 10$ m,
  - the underwater hull block coefficient CB: from $0.54 \div 0.70$,
  - the midship section coefficient CM: from $0.92 \div 0.98$,
  - the waterplane coeffcient CWL: from $0.67 \div 0.85$,
  - the blok coefficient of underwater hull fore part, CBF: from $0.45 \div 0.65$,
  - the distance from waterplane geometric centre to aft perpendicular, XF: from $41.6 \div 134.65$ m
  - the ship length b.p. $L_{pp}$: from $100 \div 280$ m,
  - the ship breadth B: from $16 \div 37.67$ m,
  - the ship draught d: from $4.3 \div 12.81$ m,
  - the underwater hull volume V: from $3780 \div 91\ 000$ m³,
  - the waterplane area Fw: from $1072 \div 9\ 000$ m²,

- the assumed wave conditions:
  - the significant wave height Hs: from 0 ÷ 8.5 m,
  - the JONSWAP wave spectrum,
  - the characteristic wave period $T_1$: from 2 ÷ 20 s,
- The assumed ship motion parameters:
  - the ship speed V = 10 m/s,
  - the wave encounter angles:
    - β = 120° for green water ingress to holds,
    - β = 120° for transverse accelerations,
    - β =   90° for vertical accelerations.

## BIBLIOGRAPHY

1. Baranowski A., Bech T., Lech R.: *Seakeeping qualities of open-top containerships*. Polish Maritime Research,vol. 4, No. 4(14), 1997
2. BOMEL Limited: *Analysis of green water susceptibility of FPSO/FSU's on the UKCS*, Offshore Technology Report 2001/005, 2001
3. Buchner B., Garcia J.L.C.: *Design aspects of green water loading on FPSOs.* The 22nd International Conference on Offshore Mechanics & Arctic Engineering Cancun, Mexico 8 ÷ 13 June, 2003
4. Cepowski T.: *Modelling of green water ingress to holds of an open-top containership in its preliminary design phase.* Polish Maritime Research, Vol. 16, No 1(59), 2009
5. Chądzynski W.: *Open top container vessel design problems.* Marine Technology Transactions, Technika Morska, Vol. 9, 1998
6. International Maritime Organization: *Interim Guidelines for Open-Top Container Ships.* MSC/Circ.608/Rev.1. London, 1994
7. *Report on the investigation of the collapse of cargo containers on Annabelle*, Baltic Sea, 2007
8. Watanabe, I., Ueno, M. and Sawada, H.: *Effect of Bow Flare Shape to the Wave Loads of a Container Ship*, JSNA Japan, Vol. 166, 1989
9. Wiśnicki B., *Vademecum of containerization. Forming the container load unit* (in Polish). LINK publishers, 2006.

---

**CONTACT WITH THE AUTHOR**
Tomasz Cepowski, Assoc. Prof.
Institute of Marine Navigation,
Maritime University of Szczecin
Wały Chrobrego 1/2
70-500 Szczecin, POLAND
e-mail: cepowski@am.szczecin.pl

# Fatigue strength determination of ship structural joints

## Part I
## Analytical methods for determining fatigue strength
## of ship structures

**Janusz Kozak**, Assoc. Prof.
**Zbigniew Górski**, M. Sc.
Gdansk University of Technology

## ABSTRACT

*Spectacular accidents at sea which have happened for a few last years show that hull structures of contemporary sea-going ships are not perfect and must be systematically improved. Fatigue strength is one of the groups of strength problems which affect design of contemporary ship's structures and greatly contribute in their improvement process. In this paper several approaches to estimation of fatigue life of hull structural elements are presented. In practice the some approaches based on nominal stresses, „hot spot" stresses or notch stresses which constitute the basis for fatigue life determination by using σ-N design curves are applied. In this paper the existing proposals have been critically analyzed and their drawbacks (often hidden) presented.*

Keywords: fatigue of ship structure; cumulative damage; fatigue life analysis

## INTRODUCTION

Contemporary merchant ships (especially VLCC and ULCC tankers and containerships (whose hulls reach ~400m in length, more than 50m in breadth and 30m in depth), as well as ocean engineering objects such as semi-submersible drilling rigs are ones of the largest technical objects worldwide. Their structures are to comply with many criteria considered crucial such as: of safety, reliability, mass, operational economy, producibility, ergonomy and esthetics. In many cases technical parameters which represent given criteria interact to each other and the interaction may be - on the one hand - of amplifying character and - on the other hand - of mutually opposing one.

From the point of view of operational conditions the structures are used in extreme conditions, i.e. they are subjected to loads generated by waves, sequential loading –unloading operations, temperature changes (day-night, crossing climatic zones during a given period of service), vibrations etc, as well as action of highly corrosive environment. It leads to determinate consequences. Firstly: occurrence of failures of hull structures usually leads to lowering their technical operational capability by limiting their functions (decreasing load-carrying capacity, occurrence of leaks etc). Secondly: in an extreme case it may lead to catastrophic collapse of the whole structure. As a rule the failures are of kind of corrosion wear or form of cracks.

The detected cracks are analyzed by classification societies as to their location and causes of occurrence [1]. As results from the reports, hull structure cracks constitute almost a half of total number of failures occurring in hull structure due to various causes. Most of the detected cracks have been qualified as fatigue cracks, that is in compliance with other data [2] where operational failures of ship structures have been divided into brittle fractures, fatigue cracks, immediate tensile as well as buckling failures, and ~75% of the cracks were considered to be generated by material fatigue.

In Fig. 1÷4 are presented examples of fatigue cracks detected on various ships and located in different zones of their structure [3, 4, 5, 6].

This paper initiates the series of publications devoted to the determining of fatigue strength of ship structure joints, in which problems of fatigue cracking and estimation of fatigue life of ship structure joints will be successively discussed both on local and zone level.

## METHODS FOR ANALYZING FATIGUE OF SHIP STRUCTURES

### Analysis of the modeling of fatigue phenomena in ship structures

Fatigue analysis of ship structures is a very complex problem. It results from many factors among which the following are most important:

a) the material fatigue phenomenon itself has been not sufficiently recognized so far due to complex influence, on run of fatigue changes, of such factors as: kind of
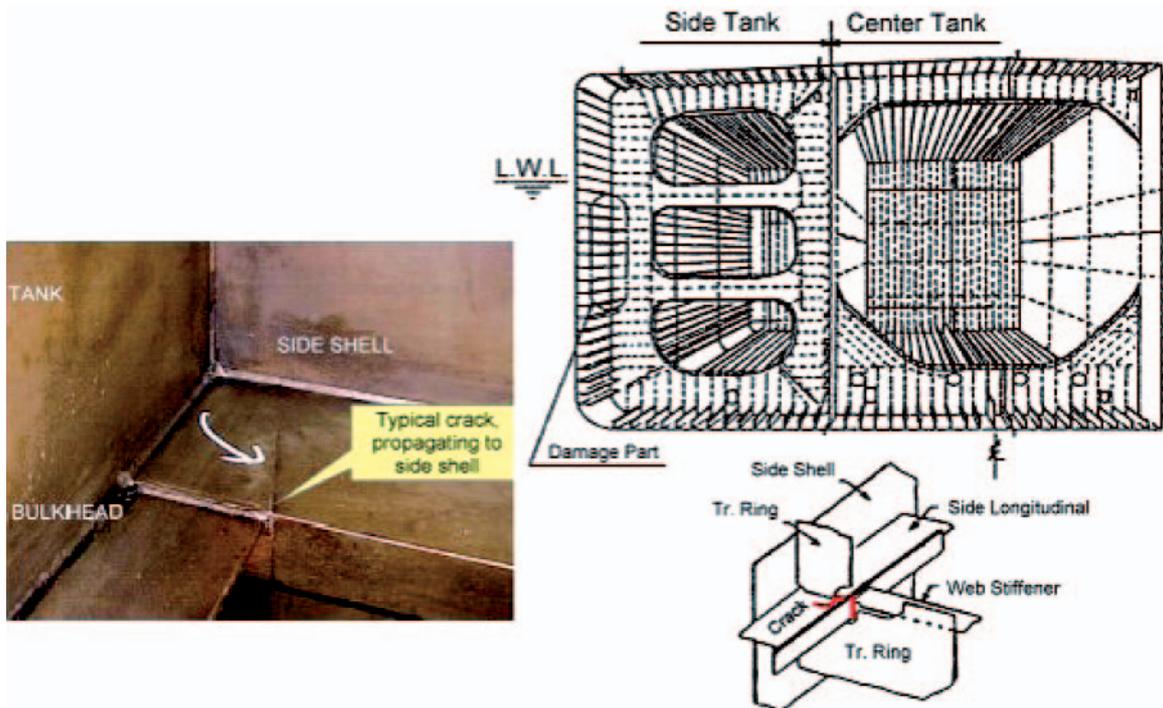
**Fig. 1.** *Example of fatigue cracks in tankers [3]*



**Fig. 2.** *Fatigue - cracked deck of the tanker Castor [4],*
*(Crack length of about 24 m)*
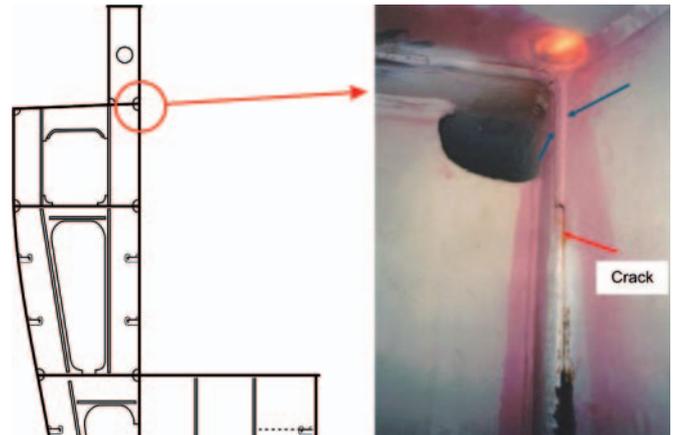


**Fig. 3.** *Example of fatigue crack in welded joint*
*of ship structure detail of a containership [3]*



**Fig. 4.** *Example of fatigue crack in base plate of ship structure detail [5]*



**Fig. 5.** *Example of fatigue crack in welded joint of ship structure detail [6]*
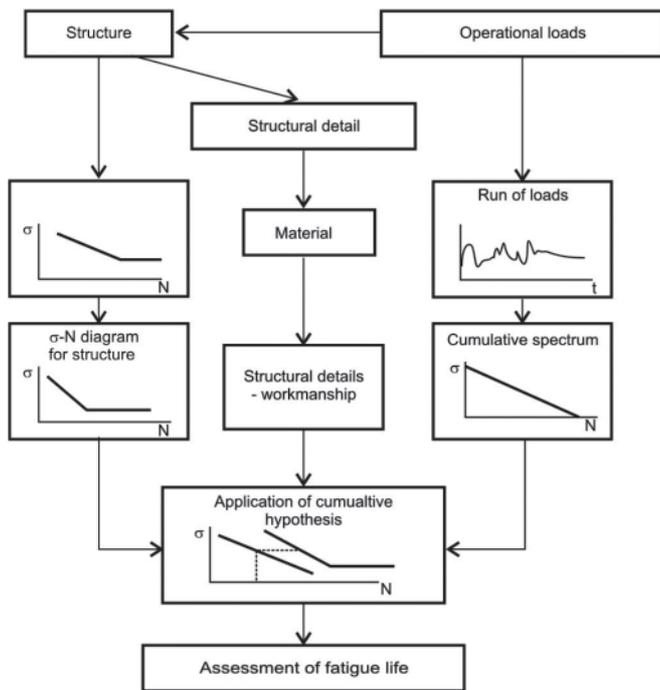
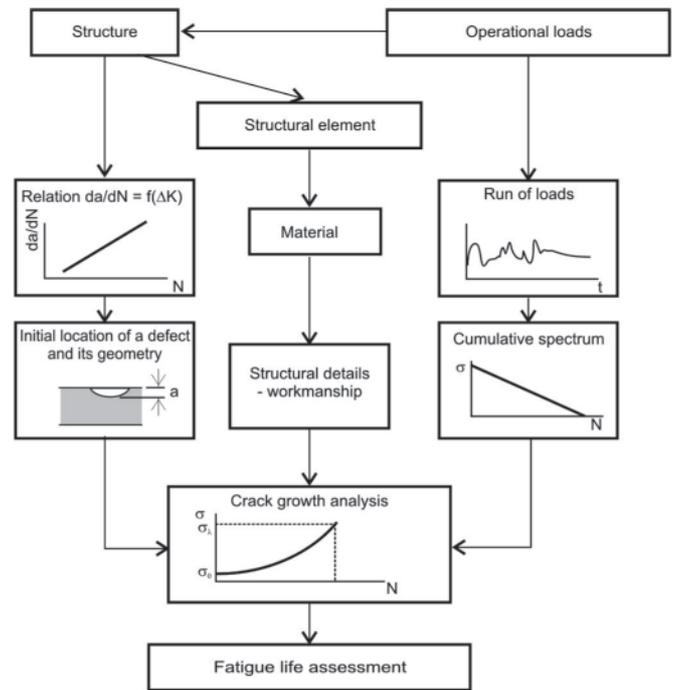**Fig. 6.** *„Safe-life" approach to fatigue analysis [8]*



**Fig. 7.** *„Fail safe" approach to fatigue analysis [8]*

material (its structure, mechanical properties, chemical composition), loading mode (its magnitude, effect of mean stresses), structural element geometry, impact of environmental conditions etc [7],

b) difficulties in unambiguous determining loads applied to the entire hull structure and its particular elements (zones, joints) as well, both as to load components, their directions and simultaneity of action,

c) different state of protective coating in particular ship zones, which may cause different run of phenomena because of a different degree of environmental influence,

d) interaction of geometrical, technological and material notches.

As a result of the above described situation, fatigue analyses of ship structures are performed under many simplifying assumptions, by using different procedures depending on aim and degree of accuracy of a given analysis.

As a result of action of variable loads in ship structures fatigue cracks appear in structural zones especially susceptible to the kind of loading, namely, in stress concentration areas to which belong regions of changeable shape of structural element (geometrical discontinuities – see Fig. 1, 3, 4 and 5), sometimes additionally amplified by presence of welds.

Fatigue failure of an analyzed ship structural element (like other elements of welded structures and/or machinery parts) and the total fatigue life (NC) corresponding to the process and measured by number of fatigue load cycles can be split into the initiation period (NI) and the propagation period (NP) of fatigue failure, as follows:

$$N_C = N_I + N_P \qquad (1)$$

Determination of fatigue life for a given load level (usually represented by stresses, and sometimes by strains) or inversely – determination of an allowable load level for a given fatigue life are the main tasks of fatigue analysis in the domain of the so called limited fatigue life. Such analysis may be conducted by applying two fundamental approaches to design (selection of scantlings) of structures, namely: according to „safe life" criterion, i.e. safe duration time (which does not allow to initiate

a fatigue crack in a considered structure or its chosen elements) – or „fail safe" criterion, i.e. safe cracking (which allows to initiate a fatigue crack in a considered structure or its chosen elements and to continue its stable growth within controlled limits up to the so called critical length (see e.g. Fig. 2).

Fig. 6 and 7 show the simplified schematic diagrams of the fatigue analyses carried out in accordance with the above mentioned criteria [8].

In the case of complex welded structures among which ship hull structures are numbered, analysis of their fatigue properties, by performing calculations of fatigue strength or fatigue life of analyzed structural joints, can be realized by means of one of the approaches making use of many possibilities of carrying out the calculations, presented in Fig. 8, [9], and Fig. 12, [10].

Choice of a given approach results from the aim attributed to the analysis to be performed. Basing on the relation (1) one is able to present complete fatigue characteristics of a given structural element in the form of sum of partial characteristics corresponding to two phases of fatigue failure, namely:

a) the initiation phase – period up to appearance of a fatigue crack,

b) the propagation phase – period of further stable growth of the crack up to the instant of triggering unstable crack propagation, that corresponds to the complete fatigue life.

Lack of an unambiguous criterion of the end of fatigue crack initiation phase (depending on aims of performed analysis there are distinguished [9]) the so called scientific and engineering criteria for the end of the initiation phase) results in arbitrary choice of an approach to fatigue analysis, which this way affects its final results.

In engineering practice of conducting fatigue analyses depending on a considered type and place of ship structure as well as consequences resulting from a mode of fatigue crack growth the notion of the „visible crack" of about 2 inches (50 mm) in length was adopted as the limit value for the crack initiation phase. The so defined „visible crack" length equal to 2 inches was confirmed as that practically distinguishable during surveys of real ship hull structures conducted by det Norske Veritas, as shown in Fig. 9 [1].
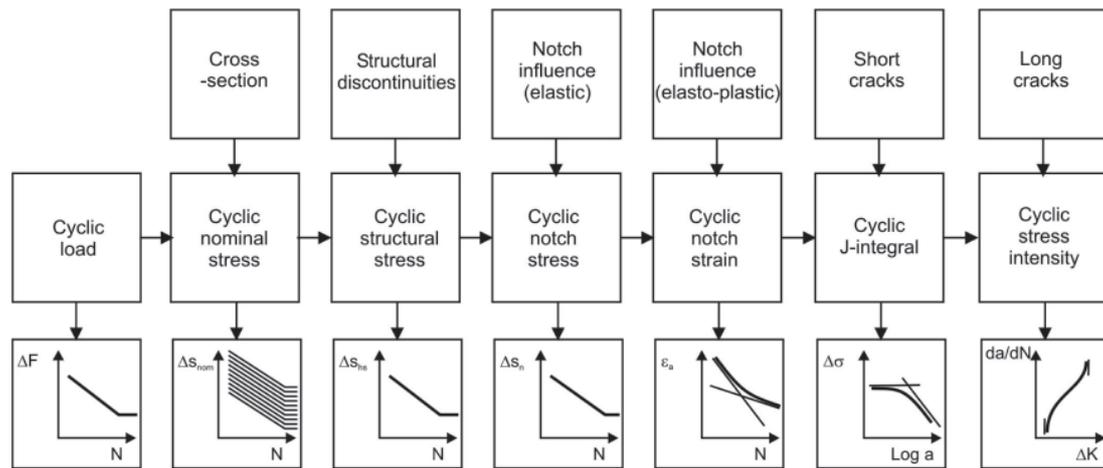
**Fig. 8.** *Overview of approaches to fatigue analyses for determining fatigue strength or fatigue life of welded structures [8]*
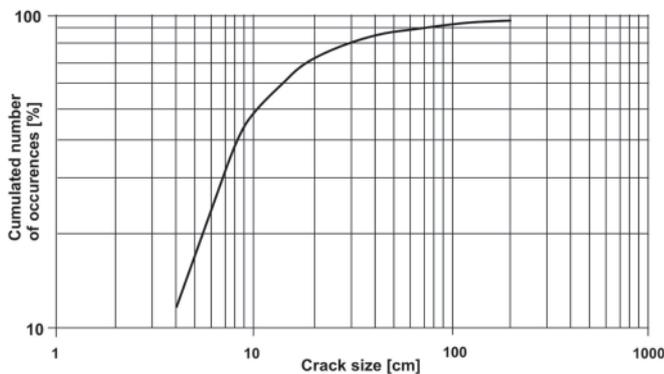


**Fig. 9.** *Cumulative length distribution of the cracks detected in ship structures by det Norske Veritas (DNV) [1]*

Depending on aims of fatigue analysis, in the case of calculation of fatigue life corresponding to the fatigue crack initiation period on the basis of the relations σ = f(N), are most often used Wőhler diagrams recommended in the rules of classification societies [11], IIW [12] and /or one's own research. In practical performing such analysis it is necessary to represent an analyzed structural fragment by means of the simplified models for which the relations σ = f(N), obtained from the same specimens as those analyzed, are known, see for instance [11] and [12].



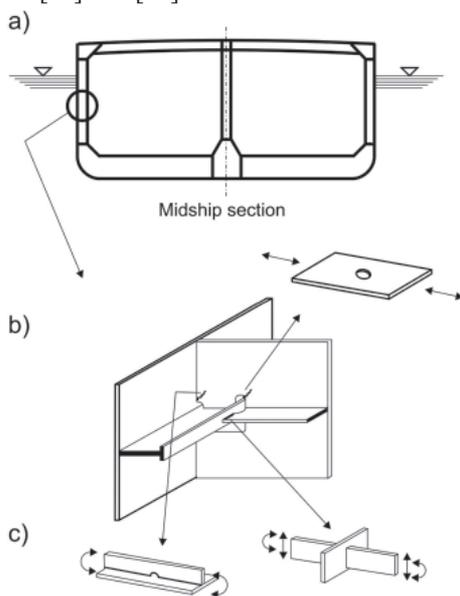**Fig. 10.** *Example of representing real ship hull structure joints by using models of elementary joints [13]:* **a)** *on global level,* **b)** *on local level,* **c)** *on elementary joint's level*

Fig.10 schematically illustrates the modeling process of elementary joints used in ship structures, and Fig. 11 [13] shows the example of the relation σ = f(N) recommended by IACS in their common structural rules [11]. Letters B,C,D… denotes here given class of curve related to type of structural element.
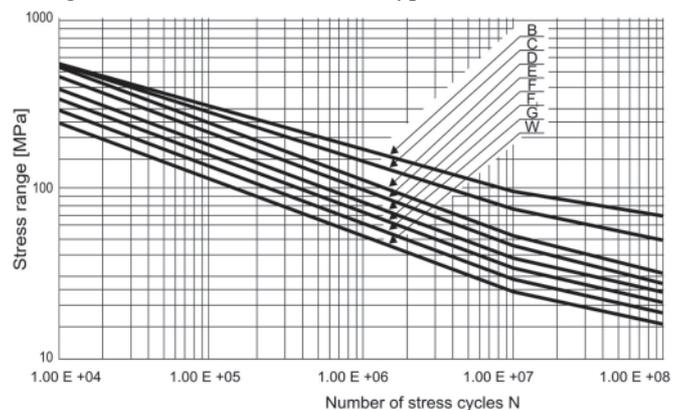


**Fig. 11.** *Wőhler diagrams for parent material and welded joints, recommended by IACS [11]*

In the case of fatigue analyses of ship structures when dimensioning their elements on the basis of the „safe life" approach, out of those presented in Fig. 8, are used the methods based on the nominal stresses σn, the „hot spot" stresses σhs, or the local notch stresses σk. In the case of determination of fatigue properties by applying the „fail safe" approach the principles resulting from fracture mechanics (FM) and dealing with fatigue crack growth (FCG) are used, Fig. 12 [14].
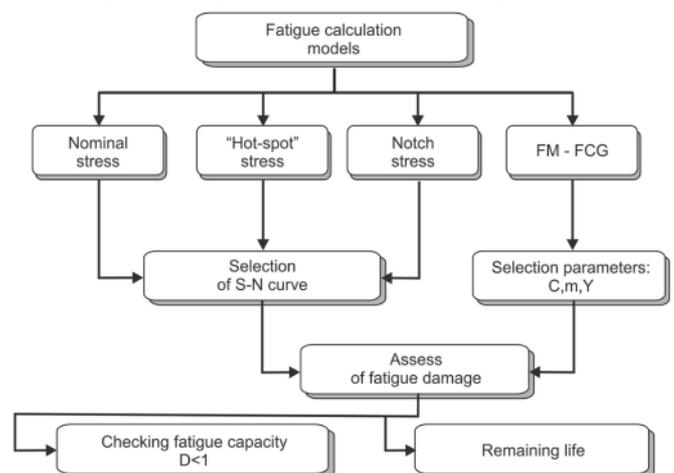


**Fig. 12.** *Schematic diagram of practical fatigue assessment procedure [14]*

## Fatigue analysis based on the σ-N fatigue diagrams

### a) Concept of the nominal stresses σ_n

The concept of performing strength calculations on the basis of nominal stresses is widely applied to welded structures. Analysis of such complex structural joint is conducted on the basis of the so called „elementary joints" distinguished in a considered structure (Fig. 10c). Loading state of such elementary joint is defined by its nominal stresses, i.e. stress values which can be determined in an elementary way on the basis of known internal tensile, compressive and shear forces and bending, twisting and other moments as well as known geometry of the considered structure [15]. If necessary, also influence of effective plate flanges are accounted for. But such effects as stress concentration resulting from geometrical discontinuities, e.g. brackets, cutouts, influence of local notches such as fusion penetration lines, initial deformations or manufacturing imperfections are not taken into account. In the approach in question the effects are accounted for in the σ -N curve representative for a given category of elements and mode of loading. The so determined stress values are considered an independent variable for fatigue life determining on the basis the relations σ = f(N) collected in various rules, standards, codes etc [11, 12, 16, 17, 18 ] and given in the form of diagrams or analytical functions containing the experimentally determined coefficients m and A:

$$N \cdot \sigma^m = A \qquad (2)$$

where:
N – fatigue life of element under the stress $\sigma_n$,
A – constant,
m – material constant.

Worth mentioning that the coefficient m is of a constant value for a given type of joint, and the coefficient A is randomly varying and described by means of logarithmic - normal distribution. After finding logarithms of both sides of the expression (2) the basic curve equation in the logarithmic reference system is obtained:
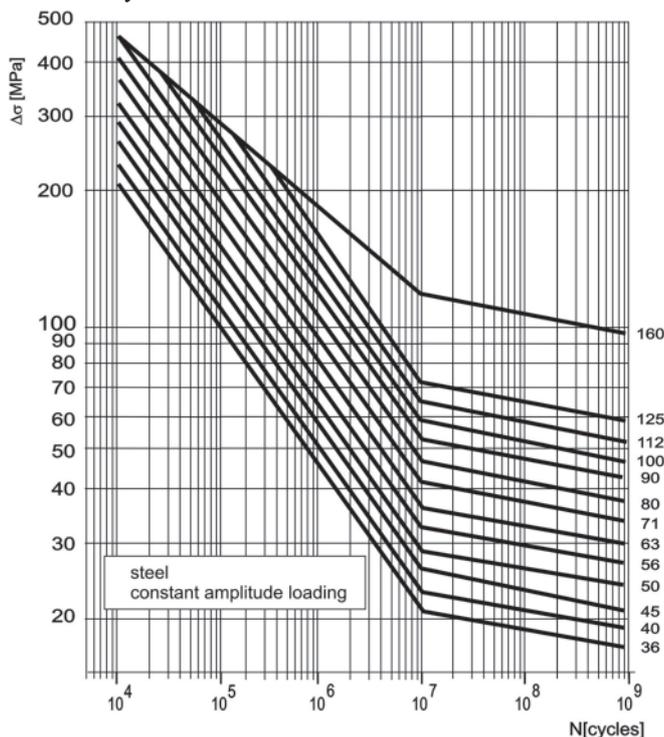
$$\log N = \log A - m \cdot \log \sigma \qquad (3)$$

The design curve Np – considered as a line which defines fatigue life up to a given phase of cracking (usually up to complete damage of a joint) - is defined as a line parallel to the basic one, shifted to the left by a multiple of the standard deviation d, namely:

$$\log N_p = \log A - m \cdot \log \sigma - k \cdot d \qquad (4)$$

where:
$N_p$ – fatigue life of element under the stress σ,
A – constant which describes the curve σ-N,
d – standard deviation of $\log(N_p)$ value,
m – material constant,
k – coefficient.

In shipbuilding practice k = 2 is usually assumed, that corresponds to 97,5% level of probability that no failure will occur [11,12 ]. Class of a given joint is determined, according to the curves, by a load value which usually corresponds to the fatigue life N = 2·106 cycles of load changes, Fig. 13. The curves presented in the figure are also determined on the probability level equal to 97,5%.

### b) Concept of the „hot spot" (structural, geometrical) stresses σ_hs

In contrast to the nominal stress the structural stress is that which contains impact of local stress concentrations resulting from changes in structure geometry except those introduced by welds. Therefore the effect of stress concentration due to weld notch is not taken into account in the approach in question. Only stress raise resulting from changes in geometry of structural elements and possible initial deformations are accounted for. Such stress determined for the point of potential fatigue crack initiation (usually taken as a border line between fusion penetration and parent material) is called the „hot - spot stress" or geometrical (structural) stress, Fig. 14 [18].
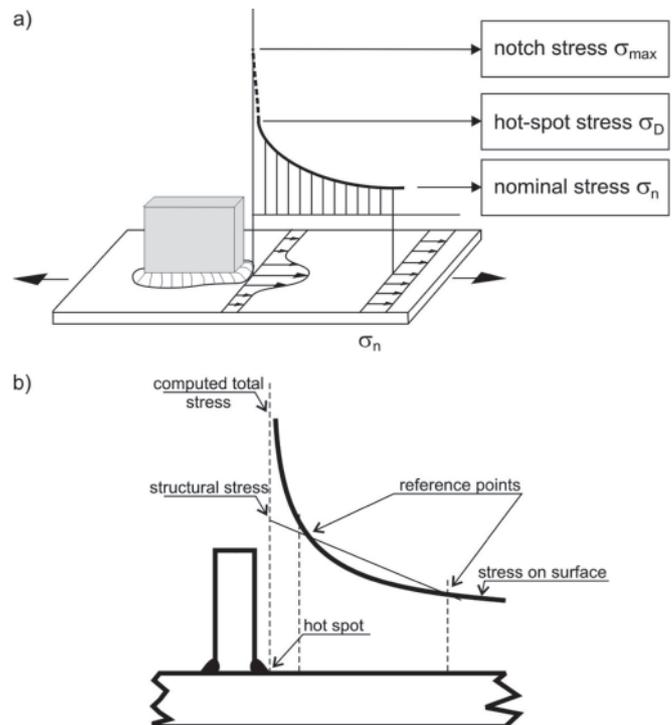


**Fig. 14.** Definition of structural hot- spot stress: **a)** stress representatives; **b)** determination of the hot-spot stress



**Fig. 13.** Example of the σ_n – N design curves (FAT), [12]

The fatigue calculation procedure based on the approach in question was described in detail in [15, 18]. In the approach it is assumed that it suffices, in fatigue analysis algorithms, to account for only the component of macro-geometrical stresses because of its variability and individual character depending on an analyzed geometry. The weld geometry effect is already contained in the design curves used for the analysis – as a rule the two: one for butt welds, the other for fillet welds – based on test results of elementary joints. It means that the structural stress in the critical point ("hot spot"), together with the curve σhs-N which is valid for a given kind of material and form of weld – make it possible to assess fatigue behaviour of an analyzed joint.

The stresses resulting from geometry of a joint ("hot spot") can be determined by:
- making use of appropriate coefficients of stress concentration associated with geometry of a considered element,
- applying the finite element method and processing the so obtained results to determine stresses in the point of their concentration [19].

In applying the second approach it is important to appropriately choose the reference points for determining the stress in the point of concentration: stress values used for extrapolation should be taken so much distant from weld as not to account for weld notch effect and so much close to it as to account for effect of geometry change, Fig. 14b. Certain considerations as to choice of reference points for plate elements are given in [20]. In present the selection procedure for the reference points is defined differently by various institutions, for instance [1]:
- International Institute of Welding recommends the linear extrapolation of stress values from the points located in the distance of 0,4·t and t from the point of weld penetration into parent material (where t – plate thickness) [21], on the other hand Yagi [22] shows that making use of the stresses in the points located respectively in the distance of $1.57\sqrt[4]{t^3}$ and $4.9\sqrt[4]{t^3}$ from the fusion penetration point, gives very good conformity with experimental data.
- Det Norske Veritas and Germanischer Lloyd recommend the linear extrapolation of stresses up to the section of weld fusion penetration into parent material on the basis of stress values taken for the points located in the distance of 0,5·t and 1,5·t from the fusion penetration point (where t – plate thickness), Fig. 15, [23, 24].
- Bureau Veritas recommends using, for geometrical peculiarity places, the linear extrapolation of two closest stress values calculated by using the finite element method [25].
- In some publications the stresses in the 0.3xt distant point is indicated as the reference values.

The stress value calculated in the stress concentration area is next used as the input data to the design curve σ-N to determine number of cycles to failure. In order to be able to consider the so obtained fatigue life values reliable it is important to ensure that values of the calculated stresses and reference stresses for the used curve σ-N correspond to each other as to geometry, material and loading conditions. In particular, the same procedure of determining stresses in the "hot spot" point should be used for the process of defining the reference stresses intended for elaboration of the design curve σ-N on the basis of laboratory test results, as well as for the calculations themselves.

The "hot spot" approach is often criticized. Especially the problem of ambiguity in selecting the reference points used for extrapolation triggers many discussions. There are several rules and guidelines which recommend different procedures for determining the "hot spot" stresses. As above presented, even for ship and offshore structures the procedures issued by some classification societies and other institutions differ from each other (Fig.15).
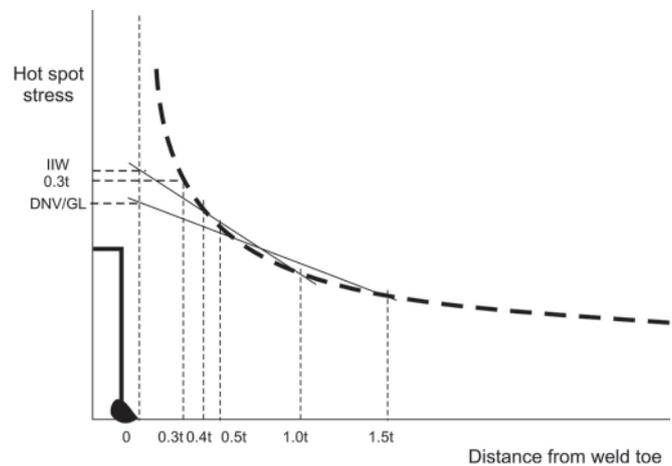


**Fig. 15.** *Differences in obtained "hot-spot" stress values depending on the points chosen for extrapolation*

Designers of welded structures consider the "hot spot" approach very useful for practical applications and providing better possibilities as compared with the approach based on the concept of nominal stresses, which leads, because of very different geometry of joints in ship and offshore structures, to inaccurate characteristics of stresses in the joints. The problem of itself is the proper modelling of weld and fusion penetration zone for calculations by using the finite element method. Wide possibilities of contemporary calculation programs and a great number of types of elements involve risk of obtaining different results for the same geometry only as an effect of different way of modelling a given task. Additionally it should be mentioned that the approach is not applicable in the case of fatigue cracks initiated at weld root and propagating across the weld. The range of application of the "hot spot" approach is limited to the cases (a) through (e) presented in Fig. 16. For the cases (f) through (j) it is not applicable at all.

c) Concept of the notch stresses σk

The notion of "notch stresses" defines the stresses which locally undergo concentration in a notch such as cutout edge or fusion penetration line.

As in the fusion penetration line sudden changes in geometry always occur, in consequence material yield point is as a rule exceeded locally under design load (which generates lower stresses beyond the weld). Simultaneously, the load is considered to be taken over by the surrounding material since the plastic zone is generally small and the surrounding structure still remains elastic that ensures blocking excessive raise of local deformations in the stress concentration zone. Hence the stress concentration coefficient $K_t$ as well as the maximum notch stress σk is not a representative parameter for assessing fatigue strength. Occurrence of the effect of load taking - over by the surrounding area is accounted for in the effective fatigue notch coefficient $K_f$ of a reduced value as compared with Kt. For sharp notches the value can be determined on the basis of effective notch stresses if only a suitably large radius of notch root is assumed, Fig. 17.

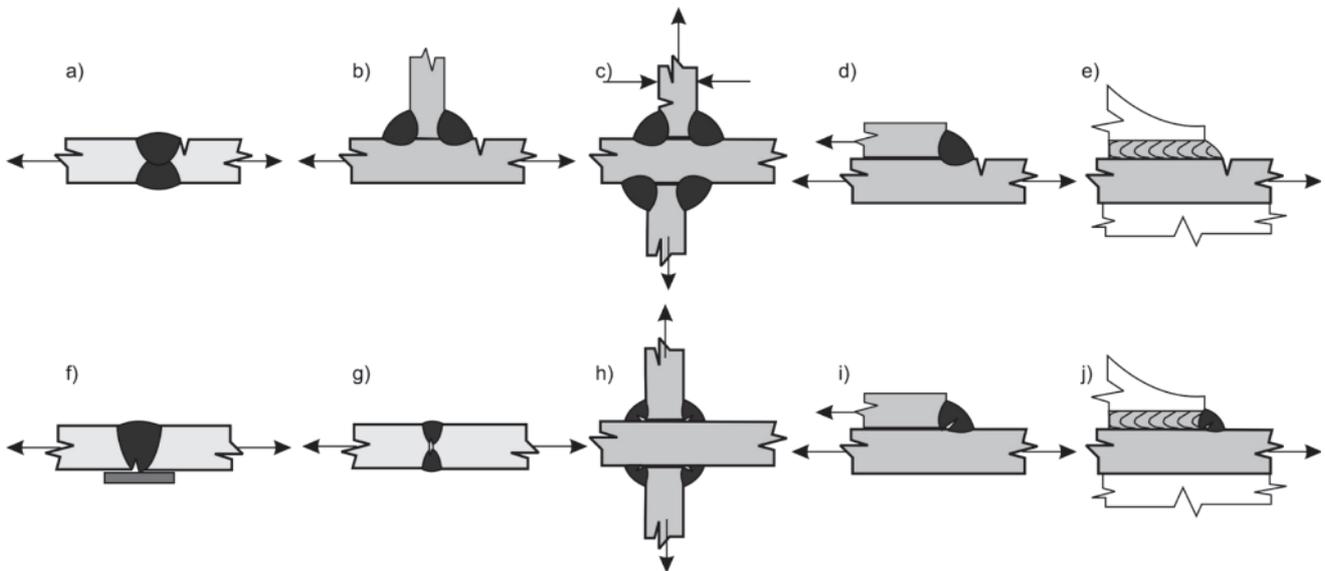Various approaches are proposed in this case. In [26] application of the radius r = 1 mm is suggested on the basis of

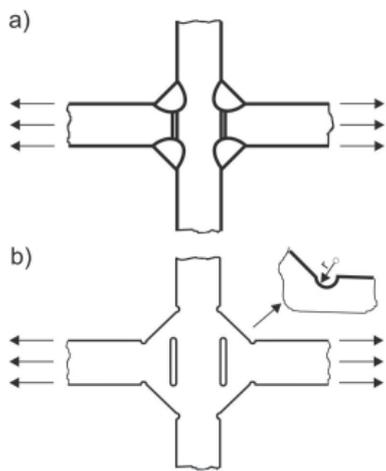**Fig. 16.** *Various locations of cracks in welded joints*



**Fig. 17. a)** *cruciform joint;*
**b)** *model of the joint for calculating effective notch stresses*

the effect of micro-support given to the intermediate structure between the weld and parent material. The similarly justified method was proposed in [13]; it leads to smaller values of the radius, depending on a given case, and requires additional transformation to be performed from $K_t$ to $K_f$. An additional factor which greatly affects stress value in the fusion penetration zone and the weld itself - and may raise it many times – are

various workmanship imperfections such as non-axiality and angular deformations of a joint or departures from its ideal geometry [26].

Disadvantages and merits of application of particular models of fatigue analysis based on the fatigue diagrams σ-N were discussed in [27] and [28]. Basing on [29] one can present general guidelines for application of a given fatigue assessment method (strategy of realization of a given approach), as shown in Tab. 1.

d) Application of fracture mechanics to analyzing fatigue of ship structures

As commonly considered, occurrence of geometrical notch in a structural element, which may cause significant stress concentration (e.g. a weld together with micro-defects contained in it), leads to shortening crack initiation phase and this is crack propagation phase which decides on total fatigue life of the element, i.e. Nc ≈ Np [7, 8, 29, 30, 31].

Therefore the information on fatigue crack growth rate in a given material as well as welded joint zones (HAZ, weld – deposited metal) corresponding to it becomes especially important as it makes it possible to classify materials as to its capability of resisting the so called final fatigue under given loading conditions [7] and to determine fatigue life of structural element especially in the limited fatigue life range.

**Tab. 1.** *Strategy for the fatigue assessment [29]*

| Type | Stress raisers | Stress determined | Assessment procedure |
|------|----------------|-------------------|----------------------|
| A | General analysis of sectional forces using general theories e.g. beam theory, no stress riser considered | Gross average stress from sectional forces | Not applicable for fatigue analysis, only for component testing |
| B | A + macrogeometrical effects due to design of the component, but excluding stress risers due to the welded joint itself | Range of nominal stress (also modified or local nominal stress) | Nominal stress approach |
| C | A + B + structural discontinuities due to the structural detail of the welded joint, but excluding the notch effect of the weld toe transition | Range of structural hot-spot stress | Structural hot-spot stress approach |
| D | A + B + C + notch stress concentration due to the weld bead notches a) actual notch stress b) effective notch stress | Range of elastic notch stress (total stress) | a) Fracture mechanics approach b) Effective notch stress approach |

The growth rate at a given crack length l, i.e. dl/dN, is expressed by the crack length increase dl per one cycle (in m/cycle or mm/cycle). The rate can be determined from slope of tangent line to fatigue crack growth curve in a considered point of the curve, which is schematically demonstrated by the point A in Fig. 18. The particular curves of the diagram l - f(N) are associated with different values of the stress amplitude $\sigma_a$ (Ni stands for number of cycles corresponding to the fatigue crack beginning).
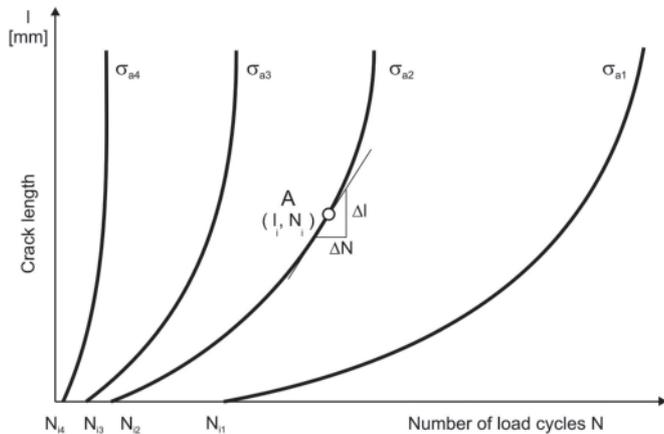


*Fig. 18. Schematic diagram of growth of fatigue cracks at various stress amplitudes: $\sigma_{a1} < \sigma_{a2} < \sigma_{a3} < \sigma_{a4}$*

The fatigue crack growth rate can be expressed in a very general form as follows:

$$\frac{dl}{dN} = f(\sigma, l, C, R,) \qquad (5)$$

where:
$\sigma$ – stress,
l – current crack length,
C – material constant,
R – stress ratio.

The functional relations (5) have been elaborated in various forms, however in practical applications those based on fracture mechanics have appeared most useful. It results from that the crack growth rate is controlled by the state of stress just ahead the fatigue crack root and the state can be unambiguously determined by the stress intensity factor K. Nonetheless the changeability range of stress intensity factor, $\Delta K = K_{max} - K_{min}$, is most often used for describing the fatigue crack growth rate. The diagram of the relation dl/dN = f($\Delta K$) has a characteristic signoidal form (Fig. 19);

From the diagram of Fig. 19 it can be observed that the fatigue crack growth can be split into three ranges:
- Range I - the range of fatigue crack initiation and its slow growth, and, as assumed, if $\Delta K < \Delta K_{th}$ the crack will not propagate,
- Range II - the range of stable fatigue crack propagation,
- Range III - the range of unstable, sudden propagation of fatigue crack preceding fatigue damage of structural element.

The range of stable fatigue crack propagation is as a rule the longest period, usually described by the so called Paris law:

$$\frac{dl}{dN} = C(\Delta K)^m \qquad (6)$$

where:
C and m – experimentally determined coefficients - the so called material constants, and the changeability range of stress intensity factor is determined in the following form:
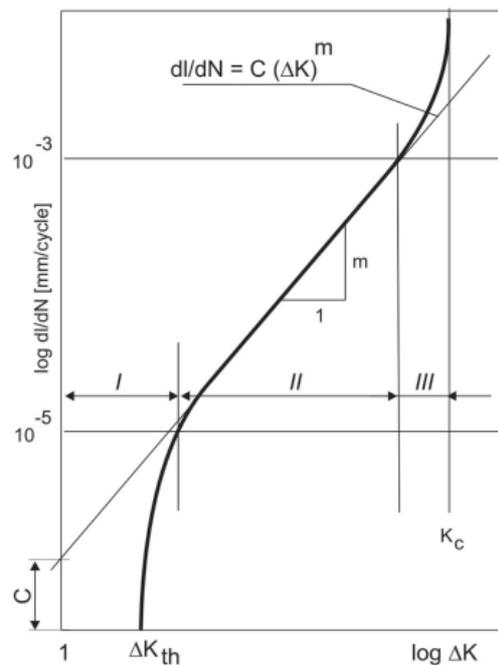


*Fig. 19. Diagram of the fatigue crack growth rate in function of the changeability range of stress intensity factor ΔK*

$$\Delta K = \Delta \sigma \cdot \sqrt{l} \cdot Y \qquad (7)$$

where:
$\Delta \sigma = \sigma_{max} - \sigma_{min;}$ $\sigma_{min} \geq 0$,
l – crack length,
Y – shape correcting function dependent on a crack length and type.

By making use of Paris relation between dl/dN and $\Delta K$, (Eq. 6), it is possible to determine fatigue life by integrating the relation (6), i.e.:

$$N = \int_{l_0}^{l_{kr}} \frac{dl}{C(\Delta K)^m} \qquad (8)$$

where:
$l_0$ – initial length of a crack, assumed – in the case of welded joints – equal to size of an existing defect (e.g. size of weld non-penetration) or the smallest dimension of a hidden defect which can be detected by using non-destructive test methods;
$l_{kr}$ – critical length of a crack, at which ultimate damage of an element will take place or - in accordance with other criterion - its unserviceability, e.g. leakage, will occur.

The application of fracture mechanics to fatigue analysis of ship structures may concern the following issues:
a) the determining of the fatigue curve S-N for an analyzed element of ship structure (e.g. a joint etc) by integrating the Paris equation (8),
b) the determining of the critical crack length $l_{kr}$ after reaching of which the following may occur e.g.: lack of tightness, drop of load-carrying capacity of a structural fragment,
c) the determining of permissible size of initial defect whose value can be taken as an initial fatigue crack length,
d) the estimating of the so called residual fatigue life for an existing crack of determined length.

## RECAPITULATION

Spectacular accidents at sea which have happened for a few last years show that hull structures of contemporary sea-going ships are imperfect and must be systematically improved. It can be done by improving the existing solutions or applying entirely novel design and technological ideas. One of the groups of strength problems which affect design of contemporary ship's structures and greatly contribute in their improvement process is that of fatigue strength. Rules of classification societies contain procedures for estimating fatigue strength of hull structural elements. They are based on the approach in which nominal stresses, „hot spot" stress or notch stress, being the base for determining fatigue life by using the design curves σ-N, are applied. For summation of effects of load action on various levels the Palmgren-Miner damage cummulation hypothesis is most often used. Though the proposed procedures have been intensively developed for a few last years, degree of compliance of results obtained with their use with real behaviour of structures remains still unsatisfactory. The reason is that many factors are not taken into account; these are among other such as: influence of load sequence both in the scale of entire structure life and in the instant of load level change, not always precisely defined reference stresses for using the design curves σ-N (because of different forms, mean values or stress ratio in test cycle, different specimen dimensions). This is also contributed by: unambiguous influence of life under small loads, imperfection of the hypothesis on linear cummulation of damages, very weakly accounted for influence of cycle mean stress, lack of a clear definition and differentiation between initiation phase and propagation phase in the scale of entire ship hull, unambiguous definitions of ways for determining stresses in „hot-spot" approach, influence of a way of modelling used in finite element method on resulting stress values, lack of differentiation of damage mechanism influence on choice of appropriate design curve, as well as many other factors. It means that the procedures have not been finally elaborated so far and hence they must be continuously improved.

### *Acknowledgements*

### BIBLIOGRAPHY

1. Kozak J. *at al.: Experimental methods in fatigue assessment of materials and structures* (in Polish) Vol. III, Wydawnictwo Naukowe Instytutu Technologii Eksploatacji (Publishing House of Operational Technology Institute, PIB, Radom, 2009
2. Rosochowicz K.: *Problems of fatigue cracks in ship hull structures* (in Polish). Wyd. Okrętownictwo i Żegluga (Publishing House Shipbuilding and Shipping), Gdańsk, 2000.
3. Fricke W.: *Local Approaches for Fatigue Assessment of Marine Structures*. Materials of Seminar, Germanischer Lloyd, Hamburg,21-22 January 2008
4. Mahmoud H.M., Dexter R.J.: *Propagation rate of large cracks in stiffened panels under tension loading*, Marine Structures 18, 2005
5. Yamamoto Satoshi: *Shipyard Experience in Structural Design of Double Hull Tankers- Past, Present and Future.*- Tanker Structure Cooperative Forum (TSCF), Shipbuilders Meeting 2007
6. Lodsberg I, Landet E.: *Fatigue capacity of side longitudinals in floating structures.* Marine Structure 18(2005)
7. Kocańda S.: *Fatigue fractures in metals* (in Polish), WNT (Scientific Technical Publishing House), Warsaw, 1985
8. Kubera S., Górski Z., Rosochowicz K., Jakubowski M.: *Influence of material properties and technological factors on fatigue strength of ship structures* (in Polish). Research reports, Ship Research Institute, Gdansk University of Technology, Gdansk, 1978
9. Radaj D., Sonsino C. M., Fricke W.: *Fatigue assessment of welded joints by local approaches.* Woodhead Publishing Limited, 2006
10. Górski Z., Kozak J.: *Estimation of development of fatigue damage in a bilge corner of Ro- Ro ship*. ASRANet Conference, Edinburgh, 2010.
11. IACS: *Common Structural Rules*, 2008.
12. Hobbacher A.F: *The new IIW recommendations for fatigue assessment of welded joints and components- A comprehensive code recently updated*. International Journal of Fatigue, Vol. 31, Is. 1, January 2009
13. Stambauch K., Leeson D., Lawrence C., Hou Y.: *Reduction of S-N Curves for Ship Structural Details.* SSC-369, Ship Structural Committee,1993
14. Presentation of Gdansk University of Technology. Materials from Technical Meeting "RISPECT Project" - PARIS November 2009, www.rispect.org.uk
15. Fricke W.: *Fatigue Strength of Ship Structures,* Part I. Germanischer Lloyd, Hamburg, 1997.
16. Germanischer Lloyd*: Rules for Classification and Construction*, Part 1: *Seagoing Ships*, Chapter 1: *Hull Structure*. Germanischer Lloyd, Edition 2002.
17. DNV: *Fatigue assessment of ship structures*. Classification Notes No 30.7, Det Norske Veritas, June 2010
18. IIW: *Recommendations for fatigue design of welded joints and components.* IIW document XIII-1965-03 / XV-1127-03
19. Lotsberg I., Larsen P.K.: *Developments in Fatigue Design for Offshore Structures*. Det Norske Veritas paper No. 2001-IL-44
20. Niemi E.: *Recommendations Concerning Stress Determination for Fatigue Analysis of Welded Components*. Abington Publishing, Cambridge UK, 1995. Stavanger, June 2001.
21. Hobbacher A.: *Fatigue Design of Welded Joints and Components*. IIW, Abington Publishing, Cambridge, 1996.
22. Yagi J., *at al.: Definition of Hot Spot Stress in Welded Plate Type Structure for Fatigue Assessment.* IIW-XIII-1414-91, 1991.
23. DNV: *Fatigue Strength Analysis for Mobile Offshore Units.* Classification Note No 30.2, Det Norske Veritas, 1984.
24. Germanischer Lloyd: *Rules and guidelines, V – analysis techniques, chapter2 – guidelines for fatigue strength analyses of ship structures.* Hamburg:. Germanischer Lloyd, Edition 2004.
25. Bureau Veritas: *Fatigue Strength of Welded Ship Structures.* Bureau Veritas, NI 393 DSM R00 E, September 1994.
26. Radaj J.: *Review of fatigue strength assessment of non-welded and welded structures based on local parameters*. Int. Journal of Fatigue, No3, 1996
27. Maddox S.J.: *Fitness for purpose assessment of misalignment in transverse butt welds subject to fatigue load*. IIW XIII-1180-85, 1985.
28. *Proceedings of the 16th International Ship and Offshore Structures Congress,* Vol. 1, Report of ISSC Committee III.2b- Fatigue and Fracture, 2006
29. *Proceedings of the 17th International Ship and Offshore Structures Congress*, Vol. 1, Report of ISSC Committee III.2b- Fatigue and Fracture, 2009
30. Crooker T.W.: *Fracture mechanics fatigue design.* Mechanical Engineering, June 1977
31. Maddox S.J.: *Calculating the fatigue strength design of welded joints using fracture mechanics.* Metal Construction and BWJ, No 8, August 1970
32. Nykanen T, Li X, Bjork T, Marquis G.: *A parametric fracture mechanics study of welded joints with toe cracks and lack of penetration.* Engineering Fracture Mechanics, 72, 2005.

---

### CONTACT WITH THE AUTHORS

Janusz Kozak, Assoc. Prof.
Zbigniew Górski, M. Sc.
Faculty of Ocean Engineering
and Ship Technology,
Gdansk University of Technology
Narutowicza 11/12
80-952 Gdańsk, POLAND
e-mail: kozak@pg.gda.pl

# Exhaust gas temperature measurements in diagnostic examination of naval gas turbine engines

## Part I
## Steady-state processes

**Zbigniew Korczewski**, Assoc. Prof.
Gdansk University of Technology

*The article presents a possible method of detecting failures in the flow section and supply system of a naval gas turbine engine based on the exhaust gas temperature measured behind the gas generator. This temperature is a basic diagnostic parameter use for evaluating the technical state of the turbine engine in operation, and is monitored during engine start-ups, accelerations and decelerations of rotor units, and steady-load states.*

## ABSTRACT

*The first part of the article is limited to the presentation of the results of diagnostic examination of a three-shaft engine with a separate power turbine, operating in steady-load states. The here presented diagnostic analyses and syntheses based on a simplified mathematical model of thermal and flow processes taking place in the combustion chamber, which was used for deriving the equation describing the average steady-state temperature of the exhaust gas flow at the outlet cross-section of the chamber.*

**Keywords**: technical diagnostics; naval turbine engines; exhaust gas temperature; steady-state processes

## INTRODUCTION

A basic control parameter of naval turbine engine operation is the temperature of the exhaust gas flow behind the combustion chamber, measured at engine steady-load states, during the start-up process, and during intermediate time intervals between consecutive steady states (acceleration and deceleration of rotor units). It results from the performed operational (diagnostic) examination of the engine that this parameter, which as a rule is the arithmetic average of the values shown by circumferentially situated thermocouples, provides opportunities for detecting places in which largest energy losses are observed in the operating processes realised by the engine. When the measured (static and dynamic) values of the exhaust gas temperature exceeds the set operating tolerance limits, it is a signal of inadmissible disturbances in the energy conversion processes realised in the engine, which threat with engine break-down (for instance unstable operation of the compressor) [2, 4, 8, 11]. It is also possible to identify and localise known and detectable failures of sub-assemblies and structural elements of the engine flow section and the system of fuel supply for the combustion chamber during engine operation [1, 7].

## EXHAUST GAS TEMPERATURE

The exhaust gas temperature is the output parameter of the combustion chamber, which in the reliability aspect is the most vulnerable element of the entire turbine engine [13]. High dynamics of the realised thermal and flow processes is the source of heavy thermal and gasodynamic loads acting on the components of its structure. No matter how complicated the structural form of the combustion chamber is, and how complicated is the aerodynamics of the primary and secondary air flow coming from the compressor [3], the temperature of the exhaust gas leaving KS can be calculated from the equations of mass and energy balance in the gas space accumulating the energy and substance of the flowing thermodynamic medium, at continuous inflow of the supply fuel – Fig.1 [2, 7, 8, 11, 12].

Inn the simplest zero-dimensional form of the model, the combustion chamber dynamics can be described by a set of differential equations of the thermal and flow processes realised with respect to time τ as the independent variable:

$$\frac{dm_{KS}}{d\tau} = \dot{m}_{pow} + \dot{m}_{pal} - \dot{m}_{spal} \tag{1}$$

$$\frac{dU_{KS}}{d\tau} = \dot{H}^*_{pow} + \dot{Q}_{pal} - \dot{H}_{spal} - \dot{Q}_M \tag{2}$$

$$\frac{dU_M}{d\tau} = \dot{Q}_M - \dot{Q}_{ot} \tag{3}$$

where:

$m_{KS}$ and $U_{KS}$ – substance and energy of the thermodynamic medium accumulated in the KS gas space,

| $U_M$ | – internal energy accumulated in the constructional material of the walls limiting the KS gas space, |
| $\dot{m}_{pow}, \dot{m}_{pal}, \dot{m}_{spal}$ | – mass flow rates of the air and fuel delivered to KS, and the exhaust gas discharged from KS, |
| $\dot{H}^*_{pow}$ | – enthalpy flux of the air delivered from the compressor, |
| $\dot{H}_{spal}$ | – enthalpy flux of the exhaust gas discharged from KS, |
| $\dot{Q}_{pal}$ | – heat flux delivered with the fuel, |
| $\dot{Q}_M, \dot{Q}_{ot}$ | – heat fluxes, respectively, passed from the thermodynamic medium to the wall limiting the KS gas space, and from the KS walls to the surroundings. |



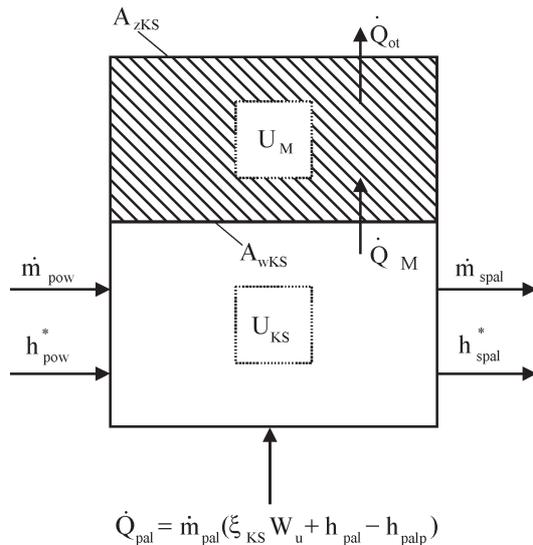$$\dot{Q}_{pal} = \dot{m}_{pal}(\xi_{KS} W_u + h_{pal} - h_{palp})$$

*Fig. 1. Physical model of thermal and flow processes in the turbine engine combustion chamber*

We can derive the following form of the equation describing changes of the exhaust gas temperature behind the combustion chamber:

$$T^*_{spal} = \frac{\dot{H}^*_{pow} + \dot{Q}_{pal} - \dfrac{dU_{KS}}{d\tau} - \dot{Q}_M}{\dot{m}_{spal} c_{pspal}(T^*_{spal})} \tag{4}$$

In steady states, in which the processes of energy and substance accumulation (alternatively dissipation) are neglected and the time derivatives of mass and internal energy of the medium in KS are equal zero, we can write the formula describing changes of the exhaust gas temperature behind KS in a much simpler form:

$$T^*_{spal} = \frac{\dot{m}_{pow} c_{ppow}(T^*_{pow}) T^*_{pow} + \dot{m}_{pal}(\xi_{KS} W_u + h_{pal} - h_{palp}) - \alpha_{wKS} A_{wKS}(T_{spalKS} - T_{MKS})}{\dot{m}_{spal} c_{pspal}(T^*_{spal})} \tag{5}$$

where:

| $\xi_{KS}$ | – coefficient of heat emission in KS, |
| $W_u$ | – net caloric value of the fuel, |
| $h_{pal}$ | – specific enthalpy of the fuel delivered to KS at the temperature equal to 293K, |
| $h_{palp}$ | – specific enthalpy of the fuel delivered to KS at the temperature $T_{spalKS}$ (heat lost for heating the fuel up to the temperature $T_{spalKS}$). |
| $\alpha_{wKS}$ | – coefficient of the convective heat-transfer from the exhaust gas to the walls limiting the KS space, |
| $A_{wKS}$ | – convective heat transfer area, |
| $T_{spalKS}, T_{MKS}$ | – temperature of, respectively, the exhaust gas and the inner KS wall surface, |
| $c_{pspal}(T^*_{spal})$ | – mean specific heat of the exhaust gas at constant pressure, within the gas temperature range characteristic for KS combustion. |

Comprehensive discussion of particular terms of the equation (5) in the context of designing control systems and controlling gas turbine engine load can be found in refs. [3, 4, 12], while their analysis and synthesis taking into account energy losses in thermal and flow processes, done for the purpose of diagnostic inference about the technical state of the engine, are given in refs. [6, 7, 9].

High requirements concerning outputs and economic indicators of the naval turbine engines force their permanent development. New, more complicated constructional forms appear, in which the combustion chamber is moved above the high-pressure compressor. Such a design leads to the reversion of the working medium flow direction, and considerable reduction in the engine length. A scheme of the working medium flow in this combustion chamber, complemented with approximate values of local flow velocities (in m/s) around the flame tube is shown in Fig. 2.
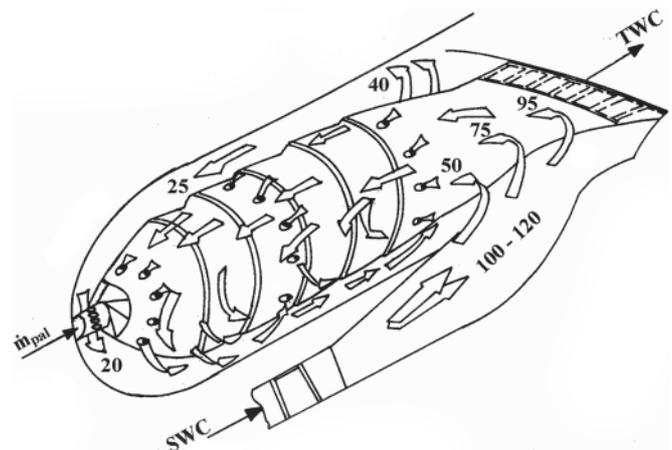


*Fig. 2. Scheme of the structural form of the combustion chamber in the reversal-flow turbine engine, wit marked local air flow velocities in m/s*

Even without any comment, the presented figure clearly shows that the thermal and flow processes realised inside the combustion chamber, and in particular inside each flame tube, are extremely complicated. In states of unsteady engine operation the course of these energy conversion processes is even more complicated, making it impossible to verify experimentally the methods of mathematical modelling (numerical simulations) conventionally used for evaluating the exhaust gas temperatures for diagnostic purposes. Therefore the basic and, most frequently, the only method of determining diagnostic relations of defect-symptom type for those engines, taking into account the exhaust gas temperature, are experimental investigations (passive and active) done on a real object [3, 7, 16].
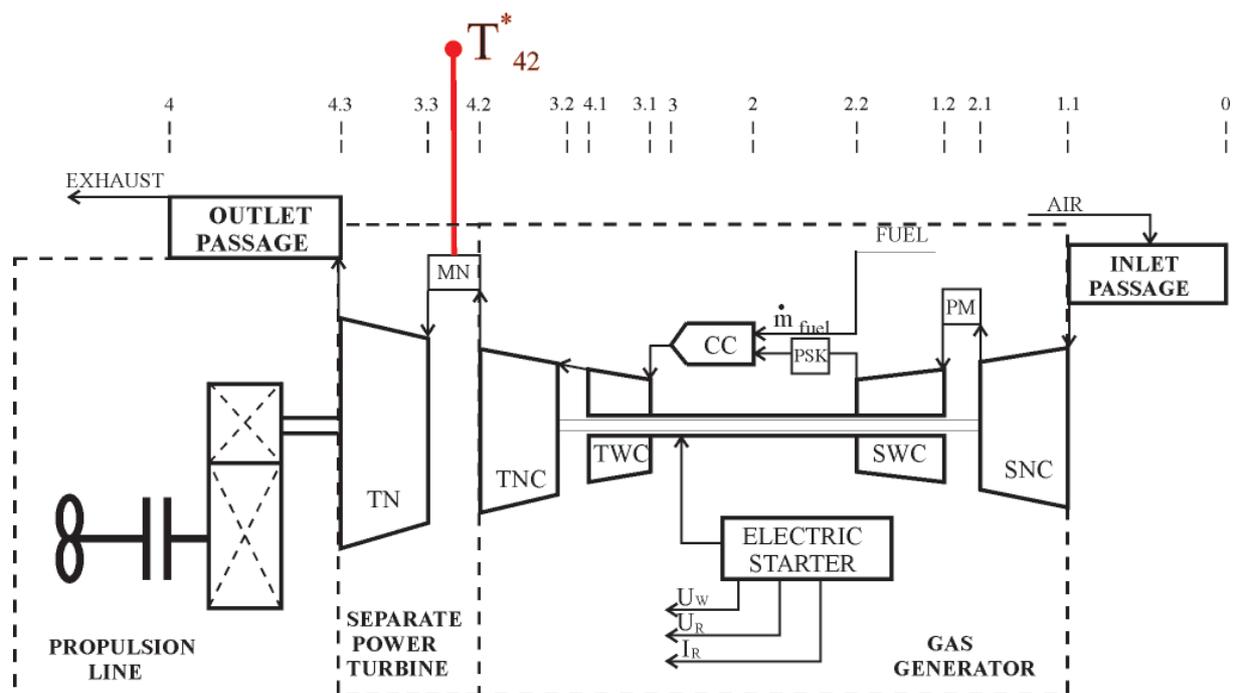
**Fig. 3.** *Schematic diagram of the naval gas turbine engine with the marked control section for measuring the exhaust gas flow temperature behind the gas generator $T^*_{42}$. SNC, SWC – low-pressure and high-pressure compressor; TWC, TNC, TN – high-pressure turbine, low-pressure turbine and power turbine; KS – combustion chamber; PM, PSK, MN – space, respectively, between SNC and SWC, between SWC and KS, and between TNC and TN; $U_w$, $U_R$, $I_R$ – excitation voltage, starting voltage and starting current*

## MEASUREMENT OF EXHAUST GAS FLOW TEMPERATURE

During gas turbine engine operation, the stagnation temperature of the exhaust gas flow is most frequently measured using thermocouples [2, 12]. The analysis of the recorded results should take into account that these measurements are burdened with errors resulting from the following reasons:

- temperature field disturbance caused by the thermocouple,
- thermal inertia of the thermocouple joint and cover,
- accelerated ageing and changes in the chemical composition and physical-and-chemical properties of the thermocouple wires due to their operation at high temperatures, (re-crystallisation, oxidation and diffusion in the surface layers of the wires close to the thermocouple's measuring joint may result in thermoelectric voltage changes exceeding ten percent, after 1000 hours of engine operation) [12,14,15].

Due to remarkable unsteadiness of thermal and flow processes in the combustion chamber and resultant pulsations of the exhaust gas flow, as well as the nonuniformity of the circumferential distribution of the exhaust gas temperature in the combustion chamber outlet section, its selective averaged measurement is done in serial engines in the flow passage control section situated at a distance from the combustion chamber [12,15] – most frequently behind the gas generator (at inlet to the separate power turbine) - Figs. 3 and 4.

Noteworthy is also the durability of the traditionally used thermocouples, which is higher when the measured exhaust gas temperatures are low [14].

Moving the thermocouples further from the combustion chamber also results from the need for maximal reduction of possible mechanical damage of the thermocouple, the broken parts of which usually cause vast damages in the blade system of the turbine part of the engine.
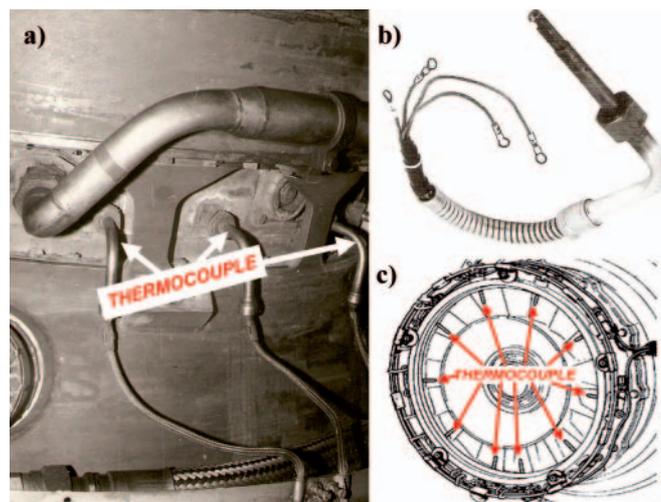


*Fig. 4. Installing thermocouples in the flow section behind the turbine engine gas generator. a) outside view of the engine, b) chromel-alumel thermocouple TXA-1368, c) engine cross-section*

### *Steady load states*

The relative increase of the exhaust gas temperature for a turbine engine in operation is the measure of decrease of efficiency of energy conversion processes realised in individual fluid-flow machines (compressors and turbines) and in the combustion chamber, as a result of flow passage fouling (or wear) (Fig. 5), or disturbances in functioning of the fuel supply system (rate of fuel delivered to the combustion chamber) – Fig. 6.

To evaluate precisely the technical state of the flow section of an engine in operation, it should be divided into smaller parts in characteristic cross-sections, and the energy balance for these parts is to be done [5, 6, 7, 16]. The obtained thermodynamic relations which capture basic gasodynamic parameters of the engine provide opportunities for calculating temperature characteristics of the flow section based on operating measurements of engine control parameters [2, 7] – Fig. 5.
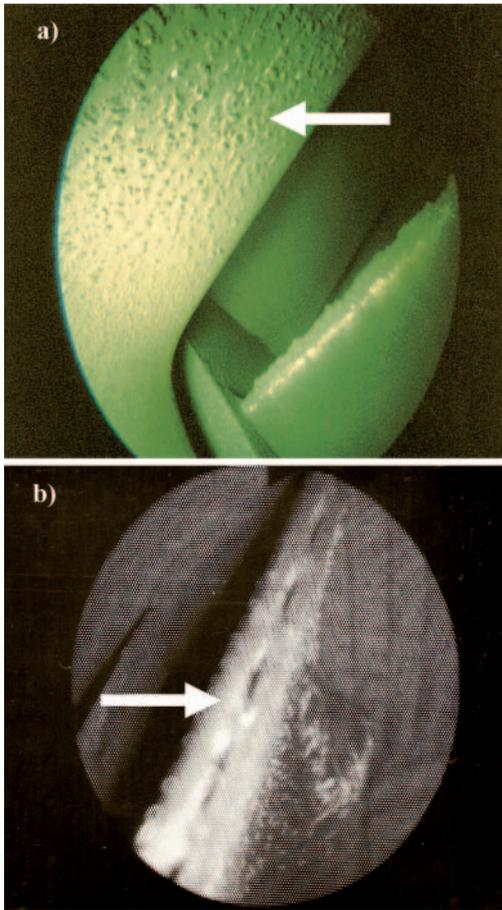
**Fig. 5.** *Deposits of impurities:* ***a)*** *on compressor stage 1 stator blades,* ***b)*** *in the area of cooling air outlet openings at the high-pressure turbine rotor trailing edge*
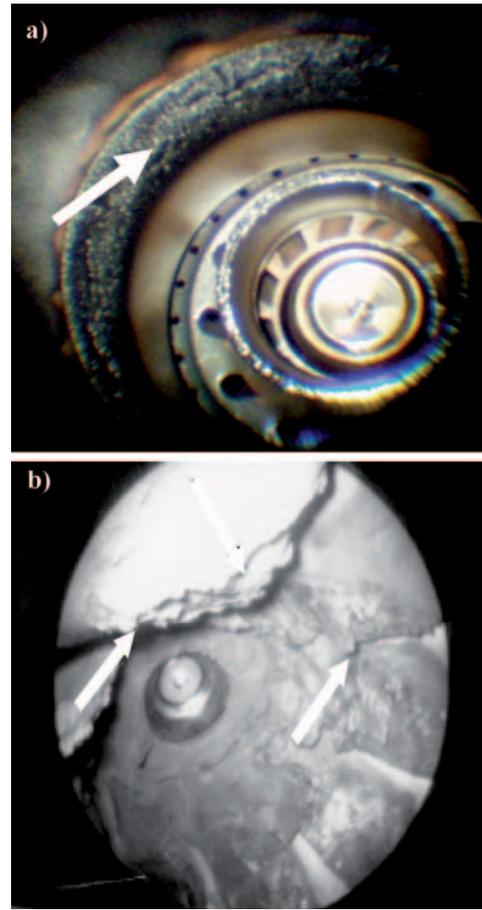


**Fig. 6.** *Carbon deposits resulting from incomplete fuel combustion in the combustion chamber:* ***a)*** *on the injector seat surface,* ***b)*** *on the flame tube swirl vane blades*

During engine operation, the state of the engine is systematically evaluated based on trends observed in the temperature characteristic of its flow section [6, 7, 8, 16]. It is represented by the distribution of the working medium stagnation temperature along the flow passage $T^*_{X-X} = f(L)$, worked out for a steady load range, most frequently equal to 1.0 $P_{nom}$. The so-called lamination rate of the temperature profile in particular engine control cross-sections x-x, determined using systematic experimental examination, is the measure of the generalised diagnostic parameter calculated from the relation [7]:

$$\Delta T^*_{X-X} = T^*_{X-X(pom)} - T^*_{X-X(o)} \qquad (6)$$

The distributions of changes of the average working medium stagnation temperature in particular control cross-sections of the flow passage, given in Fig. 7, refer to a new engine (solid line) and the same engine before and after washing the flow passage (dashed lines).

Figure 8 shows a sample diagnosis done based on the average exhaust gas temperature measured behind the gas generator of a three-shaft gas turbine engine with a separate power turbine. The engine control programme keeps the mass flow rate of the delivered fuel constant, $\dot{m}_{pal}$ = const, within the set load range related to the rotational speed of the high-pressure rotor $n_{WC}$. The relative increase of the exhaust gas temperature behind the engine's low-pressure turbine (behind the gas generator – Fig.3) is expressed as:

$$\Delta T^*_{42} = T^*_{42pom(zr)} - T^*_{42o(zr)} \qquad (7)$$

Figure 9 presents the courses of changes of the exhaust gas temperature behind the gas generator depending on the rotational speed of the high-pressure rotor, calculated from relation (7), complemented by the relative values of this increase calculated from the formula:

$$\delta(\Delta T^*_{42}) = \frac{\Delta T^*_{42\text{"AFTER WASHING"}} - \Delta T^*_{42\text{"BEFORE WASHING"}}}{\Delta T^*_{42\text{"BEFORE WASHING"}}}\Big|n_{WC} = idem \qquad (8)$$

Having analysed the numerical data shown in Fig. 9 we can conclude that washing the flow passages resulted in the decrease of the relative temperature $\delta T^*_{42}$ by 16.1%. This is a quantitative indicator of the efficiency of the performed preventive service. Comparing the temperatures for the new engine and before washing still reveals remarkable deflections from the calculated value (about 30 K). This fact confirms the need for washing the flow passage several times in order to reconstruct fully the technical state of the engine [7]. When the experimental investigations were performed, additional increase of the hourly fuel consumption accompanied by the increased power of gas generator turbines was observed, which was connected with the decreased average temperature in the control section 4.2. All this testifies to the increased mass flow rate of the air delivered to the combustion chamber, i.e. some improvement in the efficiency of the working medium compression and decompression processes in the engine.

A characteristic gasodynamic feature of the turbine engine is remarkable circumferential nonuniformity of the temperature distribution in the flow of the exhaust gas before the (hot) turbine part of the flow passage. Reliable engine diagnosing based on changes of the average exhaust gas temperatures needs complementing it by a so-called temperature field nonuniformity coefficient. This coefficient makes the basis

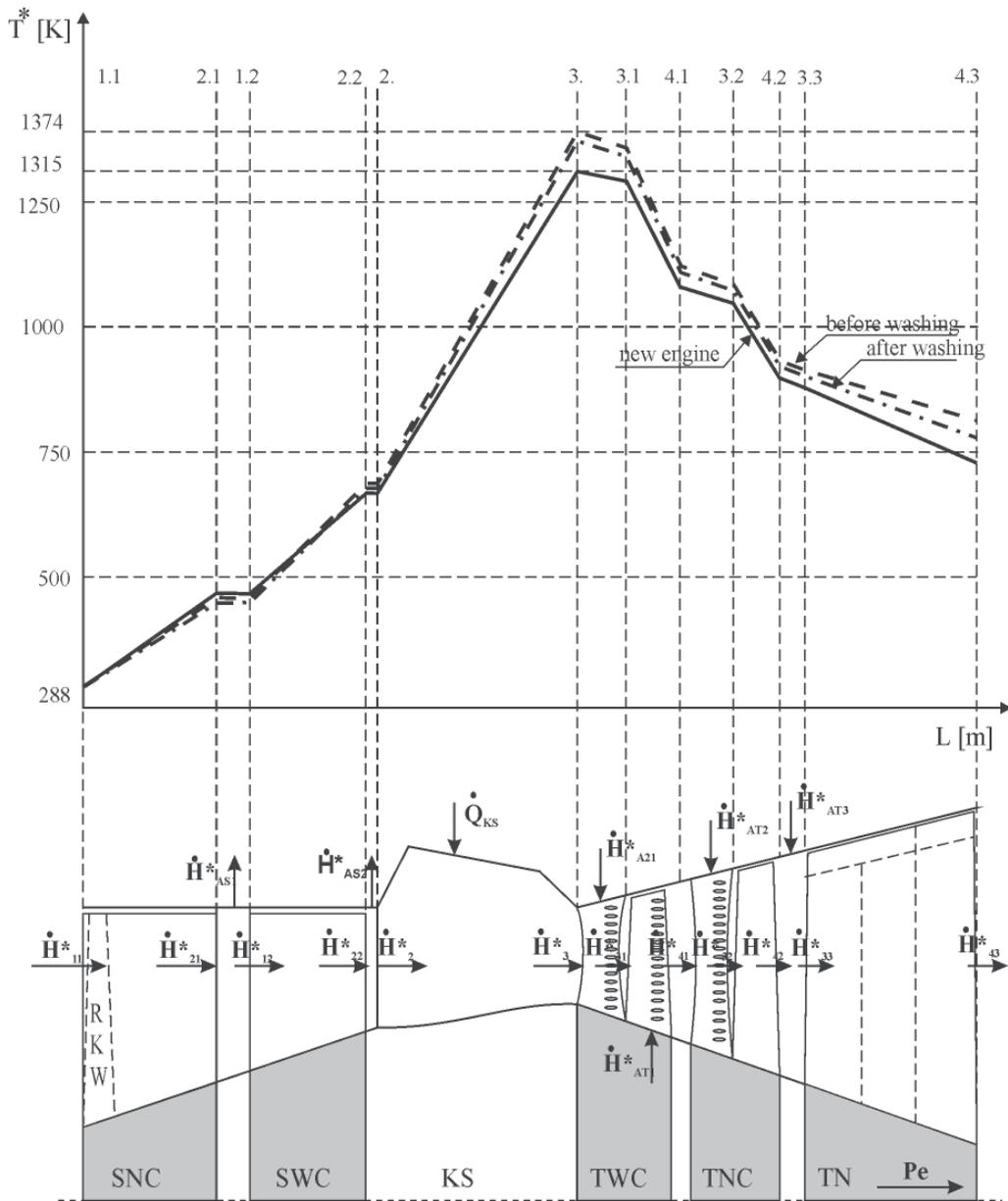**Fig. 7.** *Distribution of the working medium stagnation temperature in characteristic control cross-sections of the flow passage in a three-shaft engine with a separate power turbine for nominal load*
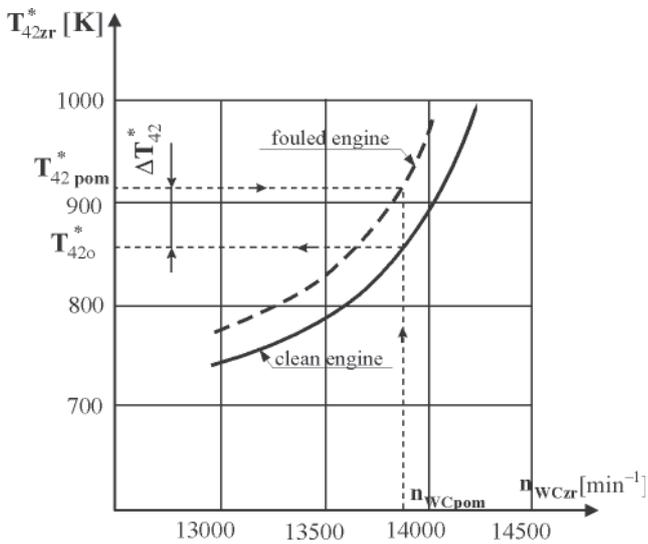


Fig. 8. Temperature characteristic $T^*_{42} = f(n_{WC})$ for a three-shaft engine with a separate power turbine
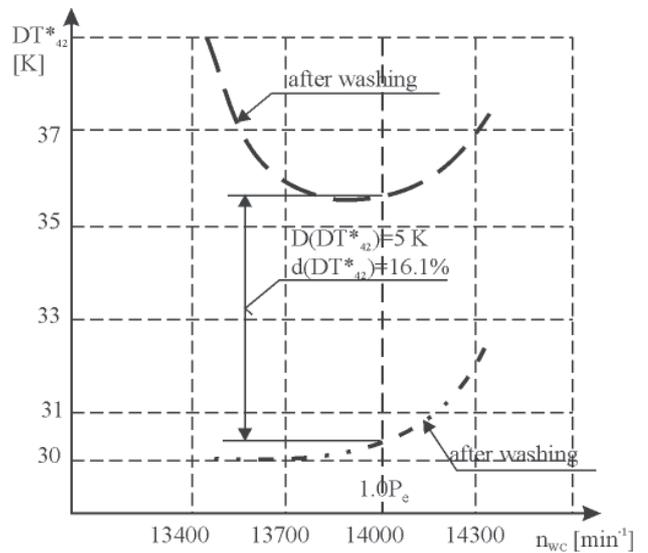


**Fig. 9.** *Temperature increment behind the gas generator vs. high-pressure rotor rotational speed*
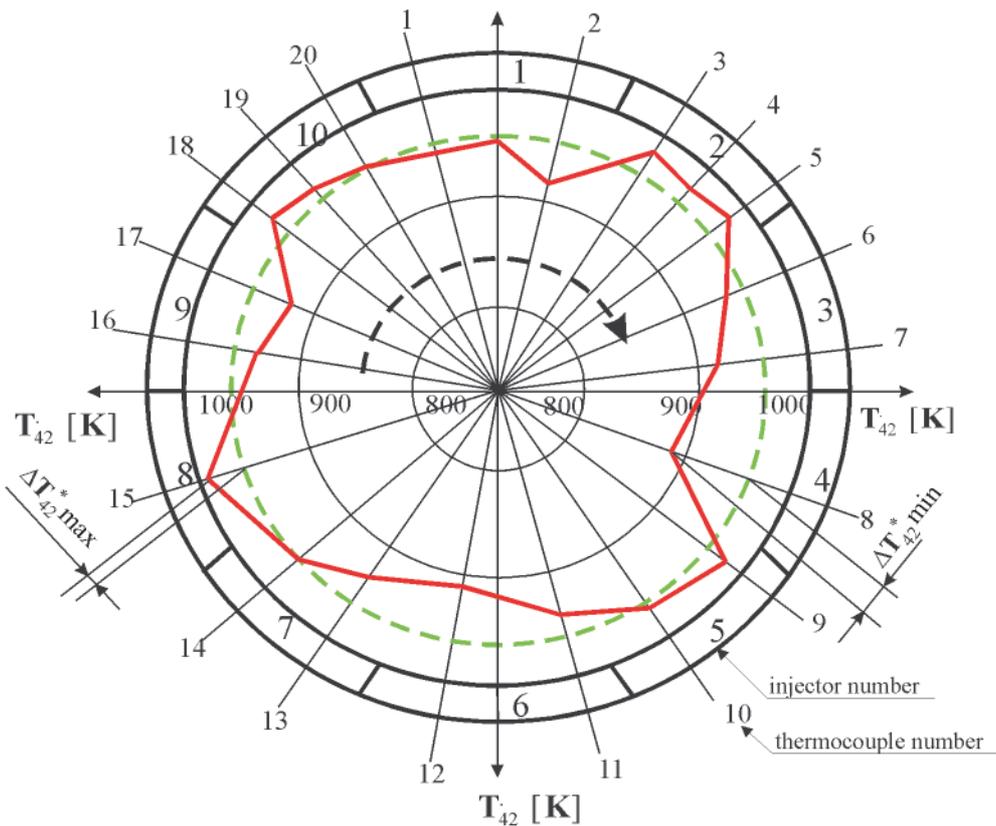
*Fig. 10. Distribution of the exhaust gas temperatures behind the gas generator
in the naval three-shaft internal combustion engine with a separate power turbine*

for assessing the quality of functioning of the engine fuel installation, with particular attention being paid to the injectors, and is defined as the maximal deflection from the average exhaust gas temperature, which is received from readings of the individual thermocouples:

$$\Delta T^*_{42\,min} = T^*_{42\acute{s}r} - T^*_{42\,min} \qquad (9)$$

$$\Delta T^*_{42\,max} = T^*_{42\,max} - T^*_{42\acute{s}r} \qquad (10)$$

According to the requirements of the producer of UGT ZORYA naval engines, the permissible dispersions of the temperature values in this control section of the flow passage must not exceed $\Delta T^*_{42} < 60$ and $\Delta T^*_{42} < 100$ at nominal load. A practical method of engine diagnosing based on the nonuniformity of the exhaust gas temperature distribution behind the gas generator is shown in Fig. 10. The diagram was worked out for the three-shaft engine equipped with 20 chromel-alumel thermocouples TXA-1368 installed in the control section 4.2 (behind the low-pressure turbine). The series-and-parallel system of connections of the thermocouples, circumferentially distributed on the same radius of the flow passage cross-section, secures high accuracy of the average exhaust gas temperature measurement (measuring error below 1%), and high reliability of operation of the engine temperature protection system [15].

When assessing the operation of the injectors based on changes in the nonuniformity of temperature distribution, we should take into account a circumferential measuring shift between particular injectors and thermocouples, which is caused by the swirl of the exhaust gas flow. For instance, the indications of the thermocouples 1 and 2 give the diagnostic information for evaluating the state of the injector which supplies the flame tube 10.

Figure 10 shows the circle diagram of the exhaust gas distribution around the perimeter of the gas generator outlet section. This diagram can be used for verifying the applicability of the average exhaust gas temperature changes in this engine control section as:

- diagnostic parameter for evaluating the technical state of the gas generator flow section,
- diagnostic parameter for evaluating the technical state of the engine fuel installation (injectors).

It is calculated using the formula:

$$\Delta T^*_{42u} = \frac{T^*_{42\,max} - T^*_{42\,min}}{T_{42sr}} \leq 10\% \qquad (11)$$

Based on the statistical analysis of the dispersion of this parameter for new engines revealing full technical capability (or those after repair done in a factory), the diagnostic tolerance of this parameter in operating conditions was assumed equal to 10% [7, 16].

## CONCLUSIONS

- Systematic measurements of the exhaust gas temperature in characteristic control cross-sections of a naval turbine engine are the basis for the diagnostic inference process about the technical state of the flow section, and the air and fuel supply system (injectors, in particular).
- The presented results of operational diagnostic investigations of three-shaft engines with a separate power turbine confirm the importance of the selected and averaged exhaust gas temperatures measured behind the gas generator in steady-load states as high-value diagnostic information.

**BIBLIOGRAPHY**

1.  Balicki W., Szczeciński S.: *Diagnosing aircraft turbine engines* (in Polish). Maszyny wirnikowe. Biblioteka Naukowa Instytutu Lotnictwa. Warszawa 2001
2.  Boliński B., Stelmaszczyk Z.: *Aircraft propulsion systems. Operation of turbine engines* (in Polish). WKiŁ, Warszawa 1981
3.  Cohen H., Rogers G.F.C., Saravanamuttor H.I.H.: *Gas turbine theory*. Longman Scientific & Technical, New York 1987
4.  Dzida M.: *Influence of gas turbine controller adjustment on ship propulsion system behavior in rough sea conditions Part 2. The simulation investigations*. Polish Maritime Research, No 1(39), Vol. 11, 2004
5.  Dzida M.: *Simulation of ship propulsion gas turbine dynamics - an educational laboratory model*. Polish Maritime Research, No 4(22), Vol. 6, 1999
6.  Hardin J.R. and others: *A gas turbine condition - monitoring system*. Naval Engineers Journal, November, USA 1995
7.  Korczewski Z.: *Methods of diagnosing the flow section of a naval turbine engine in operation* (in Polish). AMW (Ph. D. thesis), Gdynia 1992
8.  Korczewski Z.: *Identifying gasodynamic processes in the compressor system of a naval turbine engine for diagnostic purposes* (in Polish). AMW Gdynia 1999
9.  Korczewski Z.: *Endoscopy of naval engines* (in Polish). AMW Gdynia 2008
10. Korzyński M.: *Methodology of experiment* (in Polish). WNT Warszawa 2006
11. Orkisz M.: *Selected problems of the theory of jet turbine engines* (in Polish). ITE, Radom 1995
12. Pawlak W., Wiklik K., Morawski J.M.: *Synthesis and investigations of aircraft turbine engine control systems using computer simulation methods* (in Polish). Instytut Lotnictwa, Warszawa 1996
13. Szczeciński S.: *Aircraft turbine engines* (in Polish). Wydawnictwo MON, Warszawa 1965
14. Wiśniewski S.: *Temperature measurements in investigations of thermal engines and systems* (in Polish). WNT, Warszawa 1983
15. *Technical and operating documentation of naval turbine engines GTU6a, DE59, ZORYA type UGT, General Electric LM2500* (in Polish)
16. *Report on diagnostic investigations of piston and turbine engines in operation on Polish Navy vessels* (in Polish). Prace badawcze AMW, Gdynia 1992 ÷ 2008.

---

**CONTACT WITH THE AUTHOR**

Zbigniew Korczewski, Assoc. Prof.
Faculty of Ocean Engineering
and Ship Technology
Gdansk University of Technology
Narutowicza 11/12
80-233 Gdansk, POLAND
e-mail: z.korczewski@gmail.com

# A new method for detection of rolling bearing faults based on the Local Curve Roughness approach

**Mehdi Behzad**, Prof.
**Abbas Rohani Bastami**, M. Sc.
Sharif University of Technology

## ABSTRACT

*Detection of rolling bearing faults by vibration analysis is an important part of condition monitoring programs. In this paper a new method for detection of bearing defects based on a new concept of local surface roughness, is proposed. When a defect in the bearing grows then roughness of the defective surface increases and measurement of the roughness can be a good indicator of the bearing defect. In this paper a method of indirectly measuring surface roughness by using vibration signal is introduced. Several attached examples including both numerically simulated signals and actual experimental data show the effectiveness of the new, easy-to-implement method.*

**Keywords**: rolling bearing; diagnosis; vibration; local curve roughness

## INTRODUCTION

Prediction of rolling bearing life has been a big challenge since implementation of the widely used component. Early work of Stribeck and Goodman, based on Hertz stress theory, was not very successful [1]. The most important improvement in bearing life prediction was the Lundberg-Palmgren (L-P) formula based on Weibull theory of strength of materials. The theory assumes probabilistic life for bearings and finds the life with 90% survival chance, $L_{10}$, based on equivalent load and dynamic capacity of the bearings [1]. The L-P theory created a base to calculate bearing replacement intervals. In spite of widely acceptance of L-P theory by industry, the final answer to this issue was monitoring of the bearing condition instead of a fixed replacement interval.

Vibration monitoring is one of the most successful techniques of the fault detection in the rolling element bearings. There are various literature sources on the bearing fault detection by vibration analysis [2, 3]. The main techniques include using frequency spectrum [4, 5], envelope analysis [6], kurtosis [7-10], shock pulse [11, 12], synchronous averaging [13], wavelet analysis [14-20], higher spectral analysis [21], cyclic spectrum [22] and empirical mode decomposition [23, 24]. Ball pass frequencies are the main symptoms of defective bearing in the frequency spectrum. Ball pass frequencies may be also used to determine a defective part of the bearing. The major difficulty in the detection of the rolling bearing defects is the masking of weak bearing fault signature by more strong background vibration. The ball pass frequencies are usually hidden between larger peaks in the frequency spectrum. Envelope analysis is a powerful technique in detection of ball pass frequencies, but the resonance frequency must be also known. Other techniques have either some limitations or they are too complicated for practical applications.

In this paper a new method for detection of bearing faults is presented. The method called the Local Curve Roughness (LCR) has been introduced first time by the authors [25]. The LCR method uses a quantitative roughness measure as a bearing defect indicator. Effectiveness of the method is demonstrated by using both numerically-simulated and real vibration data.

## FAULT SIGNATURE OF ROLLING BEARINGS

Rolling bearings act as a source of noise and vibration in the machinery. Radially loaded bearings, even if they are healthy, generate vibrations due to a phenomenon called varying compliance [26]. This results from using finite number of rolling elements and change in effective stiffness of the bearing. Presence of a local defect in contact surfaces causes a significant increase in vibration level of the bearing. Local defects including crack, pit and spall in the bearing parts are known to produce a train of impacts which propagate through the bearing and housing to the vibration transducer which is located on the bearing housing. The train of short - duration impulses excites structural resonance and produces a train of

damped natural vibrations [27]. Thus the vibration signature of a faulty bearing can be expressed as [22]:

$$x(t) = \sum_{i=-\infty}^{\infty} q(t) \cdot h(t - iT) \qquad (1)$$

where:
the index i denotes $i^{th}$ impact,
h – the impulse response of the structure,
T – the time interval between impacts,
q(t) – periodic function synchronized with shaft rotation, which account for load variation due to bearing defect.

If the defect is located on a fixed part of the bearing then q(t) has a constant value. In a simple model, as follows from [28], h(t) can be considered the impulse response of a single DOF system and characterized by two parameters, namely the natural frequency $\omega_n$ and the damping ratio $\zeta$:

$$h(t) = \frac{1}{m\omega_d} e^{-\zeta\omega_n t} \sin(\omega_d t) \qquad (2)$$

where:
m – mass,
$\omega_d$ – damped natural frequency.

To model real conditions in the equation (1) some vibrations related to the shaft rotation, external sources and random noises are to be added. A more complete model of vibration generation in the rolling bearings is introduced in [29].
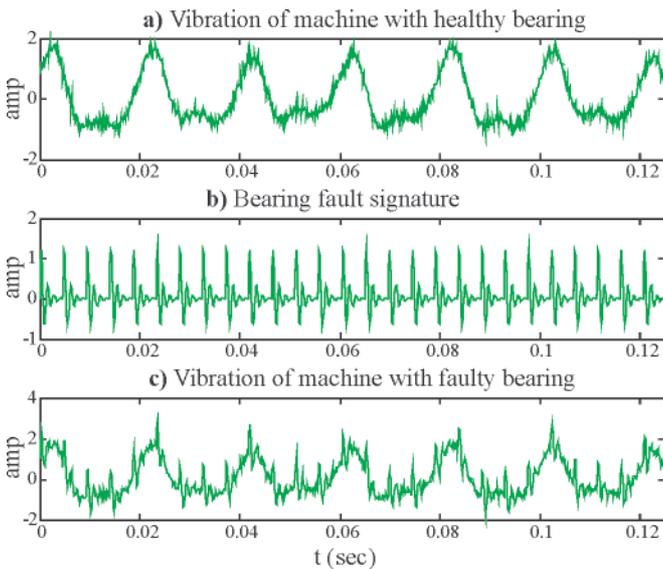


**Fig. 1.** *Numerically-generated vibration time-waveform curves for:*
*a) a machine with healthy bearing; b) bearing fault signature;*
*c) a machine with faulty bearing, (amp - vibration amplitude, m/s²)*

A numerically- generated vibration signal of a 50 Hz frequency machine with healthy and faulty bearings is shown in Fig. 1. The signal is consisted of the following parts:
1. Vibration due to shaft rotation including the shaft speeds of the first, second and third order with the amplitudes of 1, 0.5 and 0.1 m/s², respectively.
2. Vibration resulting from an external source with the frequency of 37 Hz and amplitude of 0.2 m/s²
3. Vibration due to outer race defect with the ball pass frequency of 4.29x50 = 214.5 Hz , natural frequency of 800 Hz, $\zeta$=0.1 and amplitude of 1.2 m/s²
4. Gaussian white noise with zero - mean value and standard deviation of 0.1 m/s².

5. The frequency spectrum of the signal shown in Fig. 1 is presented in Fig. 2. It is notable that in spite of the high amplitude of bearing fault signature in the time domain, the amplitude of the characteristic frequency in the frequency spectrum is low (of only 0.11 m/s²) and can be easily masked by other vibration sources. The problem arises from using sine bases in Fourier transform to decompose the impulsive vibration of defective bearing. Obviously in the decomposition the first harmonic does not possess the majority of the signal energy and higher harmonics are usually predominant in the spectrum.

To overcome the difficulty, a nonlinear transform is introduced which filters out the harmonic content of the signal and amplifies the train of impulsive vibration. Therefore the Fourier transform of the resultant signal will contain the frequency of impacts and bearing fault signature can be easily detected.
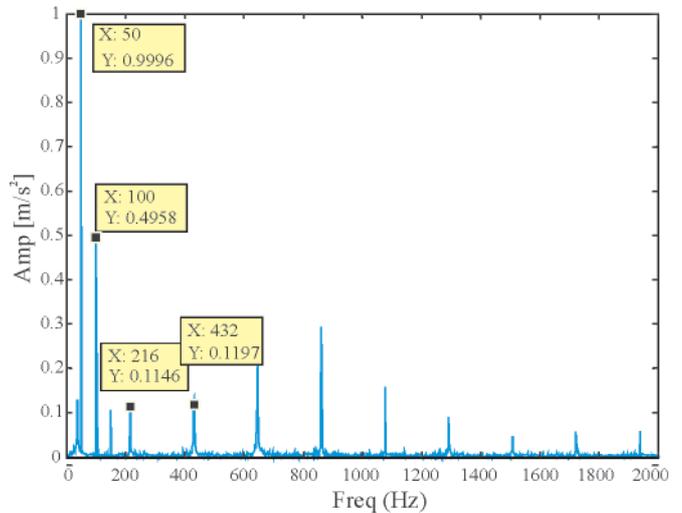


**Fig. 2.** *Frequency spectrum of the vibration signal shown in Fig (1-c), the graph shows vibration amplitude (m/s²) versus frequency (Hz)*

## DEFINITION OF VIBRATION ROUGHNESS

The defective area in the rolling bearing differs, as to the surface roughness property, from the healthy area. In other words, defective area has rougher surface than healthy area. Surface roughness affects the contact force in the rolling process and subsequently changes the generated vibration. The main idea of the proposed method is that rougher surface produces "rougher" vibration [in the sense of a more distinct irregularity of vibration signal]. If one has a quantitative measure of the so understood roughness of vibration signal the defective bearing can be identified. Various parameters may be selected to quantize roughness of a signal. Actually, the vibration roughness shall not only depend on the value of the signal but also on the rate of change of the signal. The definition used herein for the signal roughness x(t) is the actual curve length of the continuous differentiable signal x(t) per unit time, given by:

$$LCR_x(t_0, d) = \frac{1}{d} \int_{t_0}^{t_0+d} \sqrt{1 + \left(\frac{dx}{d\theta}\right)^2} \, dt \qquad (3)$$

where:
LCR – stands for local curve roughness,
$t_0$, d – a moving window position and width, respectively,
x(t) – continuous differentiable signal assumed dimensionless,
$\theta$ – the rotation angle equal to $\theta = \Omega t$.

If x has a dimension it must be made dimensionless first. If dimension of x is that of acceleration, velocity or displacement, the following dimensionless parameters are proposed:

*Acceleration*:

$$x^* = \frac{x}{R_m \omega_c^2}$$

*Velocity*:

$$x^* = \frac{x}{R_m \omega_c}$$

*Displacement*:

$$x^* = \frac{x}{R_m}$$

where:
x – used for vibration signal in all three cases of acceleration, velocity and displacement,
$R_m$ – pitch diameter of the bearing,
$\omega_c$ – cage rotation frequency.

Some properties of the LCR are as follows:
*Nonlinearity*: It is evident that LCR is a nonlinear transform.
*Value of LCR is always greater than one*: only the LCR of a constant function x(t) = c equals to 1. LCR of any other signal is greater than one. The LCR can be divided into a constant mean value and oscillating part.
*Averaging*:

$$d_1 \cdot LCR_x(t_1, d_1) + d_2 \cdot LCR_x(t_1 + d_1, d_2) =$$
$$= (d_1 + d_2) \cdot LCR_x(t_1, d_1 + d_2)$$

## FREQUENCY RESPONSE OF LCR

In this section local curve roughness of a sine signal will be studied in more detail. If to assume x(t) = Asin(ωt) the LCR will be:

$$LCR_x(t,d) = \frac{1}{d} \int_t^{t+d} \sqrt{1 + (A^* \omega / \Omega)^2 \cos^2 \omega\tau}\ d\tau =$$
$$= \frac{1}{d} \int_t^{t+d} \sqrt{1 + \omega^{*2} \cos^2 \omega\tau}\ d\tau$$

(4)

where:
$A^*$ – non-dimensional amplitude;
$\omega^*$ – non-dimensional frequency defined as $\omega^* = A^*\omega/\Omega$.

To show LCR frequency response the low and high frequencies will be treated separately. For the low frequencies, $\omega^* \ll 1$, the following approximation can be used:

$$\sqrt{1 + \omega^{*2} \cos^2 \omega\tau} \approx 1 + \frac{\omega^{*2}}{2} \cos^2 \omega\tau$$

(5)

Therefore LCR can be approximated by substituting Eq. (5) into Eq. (4):

$$LCR_x(t,d) \approx 1 + \frac{\omega^{*2}}{44} +$$
$$+ \frac{\omega^{*2}}{\omega d} \sin(\omega d) \cos(2\omega t + \omega d)$$

(6)

Eq. (6) shows that LCR is a combination of a constant value term of $1 + \omega^2/4$ and an oscillating part with a frequency that is

twice the original signal frequency. It can be observed that the frequency increasing amplifies both the mean and oscillating part of LCR.

For the high frequency signals of ω ≫ 1 we use the Fourier series of the integrand of Eq. (4) which is a periodic function of the period π/ω. Since $\sqrt{1 + \omega^{*2} \cos^2 \omega\tau}$ is an even function of τ, only cosine terms remain in the Fourier series which can be written as:

$$\sqrt{1 + \omega^{*2} \cos^2 \omega t} = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2n\omega t)$$

(7)

$$a_n = \frac{2\omega}{\pi} \int_0^{\pi/\omega} \sqrt{1 + \omega^{*2}\cos^2\omega t} \cdot \cos(2n\omega t)dt, \ n = 0,1,2,K$$

The integral in Eq. (7) can be evaluated by using the change of variable ωt = ø. Then:

$$a_n = \frac{2\omega}{\pi} \int_0^{\pi} \sqrt{\frac{1}{\omega^{*2}} + \cos^2\phi} \cdot \cos(2n\phi)d\phi, \ n = 0,1,2,K$$ (8)

Under the assumption that ω ≫ 1, the term $1/\omega^{*2}$ can be ignored and we have:

$$a_n \approx \frac{2\omega}{\pi} \int_0^{\pi} |\cos\phi| \cdot \cos(2n\phi)d\phi, \ n = 0,1,2,K$$

(9)

The integral in Eq. (9) can be evaluated as follows:

$$a_n = \frac{(-1)^{n+1} 4\omega}{(4n^2 - 1)\pi}, \ n = 0,1,2,K$$

(10)

From Eq. (10), the first five coefficients of the Fourier series will be:

| $a_0/2$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| 2ω/π | 4/3 · ω/π | −4/15 · ω/π | 4/35 · ω/π | −4/63 · ω/π |

By keeping the two first terms of the Fourier series, the LCR of a sine curve will be:

$$\sqrt{1 + \omega^{*2} \cos^2 \omega t} \approx \frac{2\omega^*}{\pi} + \frac{4\omega^*}{3\pi} \cos(2\omega t)$$ (11.a)

$$LCR_x(t,d) \approx \frac{1}{d} \int_t^{t+d} \left( \frac{2\omega^*}{\pi} + \frac{4\omega^*}{3\pi} \cos(2\omega t) \right) d\tau$$ (11.b)

$$LCR_x(t,d) \approx \frac{2\omega^*}{\pi} +$$
$$+ \frac{2\omega^*}{3\pi\omega d} [\sin(2\omega t + 2\omega d) - \sin(2\omega t)]$$ (11.c)

$$LCR_x(t,d) \approx \frac{2\omega^*}{\pi} + \frac{4\omega^*}{3\pi\omega d} \sin(\omega d) \cos(2\omega t + \omega d)$$ (11.d)

Again the LCR is consisted of two parts: a mean value and an oscillating part with a frequency that is twice the original signal frequency. The increase of ω increases the mean value of the LCR monotonically, but the oscillating part amplitude shows an oscillating trend. The window width has no effect on the mean value but is inversely proportional to the oscillating part.

The actual LCR of a sine curve calculated by using numerical integration of Eq. (4) with low and high frequency approximations, is plotted in Fig. (3-a) and (3-b) where Ω = 1 and $A^* = 1$ are used. The figures verify the low frequency approximation in Eq. (6) and high frequency approximation in Eq. (11-d).

## IMPULSE RESPONSE OF CURVE ROUGHNESS

In this section the LCR of the impulse signal $x(t) = \delta(t)$ is studied. The impulse with time duration b is defined as follows:

$$\delta(t) = \begin{cases} 1/2b, & |t| < b \\ 0, & \text{else} \end{cases} \quad (12)$$

The LCR of impulse function defined by Eq. (12), if $d > 2b$ is assumed, is:

$$LCR_\delta(t,d) = \begin{cases} 1, & t + d < -b \\ 1 + 1/2bd, & t < -b, |t+d| < b \\ 1 + 1/2bd, & t < -b, t + d > b \\ 1 + 1/2bd, & |t| < b, t + d > b \\ 1, & t > b \end{cases} \quad (13)$$

Eq. (13) shows that the impulse width 2b is extended to d + 2b and its height is changed to 1/bd. The increasing of the LCR window length d makes the LCR maximum value decreasing because of the averaging property. Therefore narrow impulses can be amplified by using LCR if only a narrow window will be chosen. LCR of an impulse of b = 0.01 sec calculated by applying a window of the length d = 0.04 sec is shown in Fig. 4.
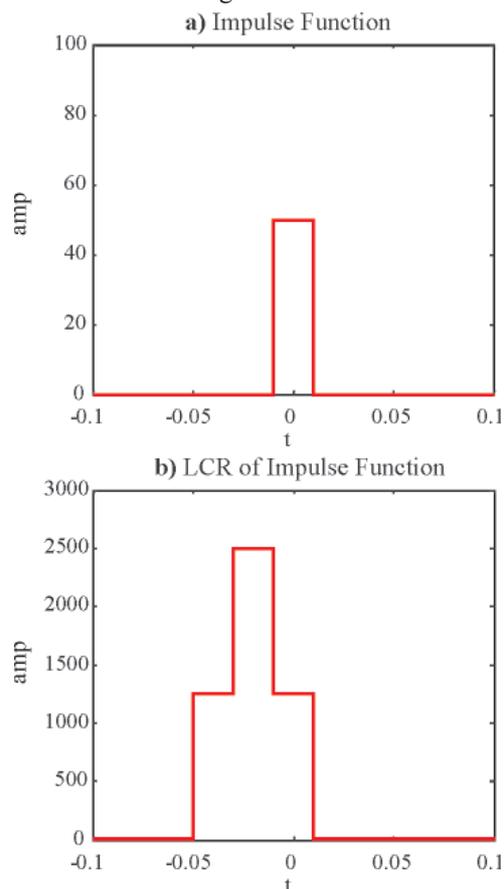


***Fig. 4. a)*** *An impulse of the width b = 0.01 sec, **b)** LCR of the impulse calculated by applying a window of the length d = 0.04 sec, amplitude (m/s²) versus time (sec)*

## DISCRETE LCR

In the case of the discrete signal, x[n], the LCR can be calculated by means of the following formula:

$$LCR_X(n_0, d) = \frac{1}{d} \sum_{i=n_0}^{n_0+d-1} \sqrt{1 + \frac{f_s^2}{\Omega^2} [x(i+1) - x(i)]^2} \quad (14)$$

where:
$f_s$ – sampling frequency,
$d$ – window length.

It is necessary to study effect of sampling frequency on the LCR. In Fig. (5) mean value and oscillating part of LCR of a sine curve is plotted against sampling frequency. It is evident that oscillating part is more sensitive to sampling frequency. In order to avoid sampling problems it is necessary to use a sampling frequency at least 20 times greater than the frequency of the signal.

## APPLICATION OF LCR TO BEARING FAULT DETECTION

The main idea of LCR for bearing fault detection is that periodic impacts inside the bearing produce local disturbances in the vibration signal, which have large local roughness in terms of its above defined notion. To illustrate this idea, the normal vibration signal of a machine as well as compound normal vibration - bearing fault signature are shown in Fig. (6). LCR decreases the magnitude of oscillating part of harmonic vibration associated with normal shaft vibration, while it amplifies the amplitude of impulsive vibration. This makes LCR an ideal transformation to perceive bearing faults. Since impulsive events are better observed in acceleration signal, its use is more beneficial than that of velocity and displacement.
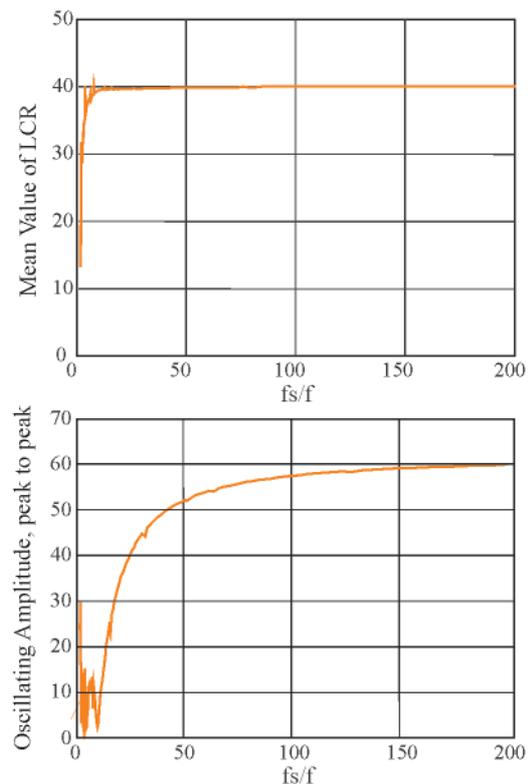


***Fig. 5.*** *Mean value and peak-to-peak amplitude of oscillating part of LCR of a sine curve versus **relative** sampling frequency (fs/f)*

To interpret the LCR, one should take Fourier Transform of the oscillating part of LCR. The frequency of consecutive impacts will be clear in the resulting spectrum. By comparing this frequency with ball pass frequencies of the bearing, the presence of a defect in the bearing and its location can be recognized as usual.
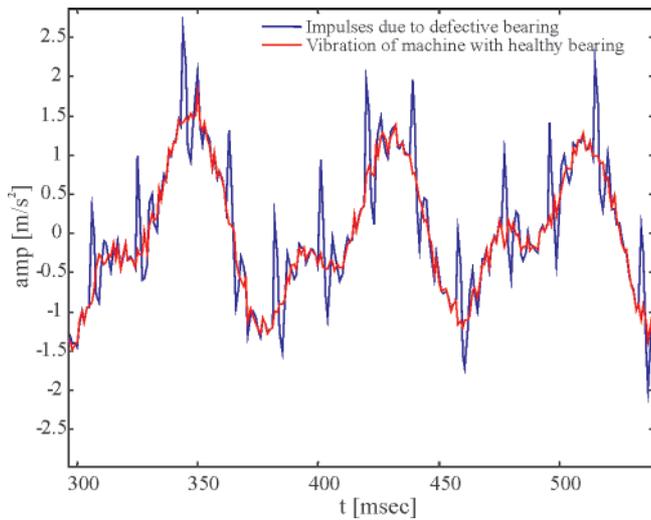
**Fig. 6.** *Comparison of normal vibration signal of a machine and compound normal vibration - bearing fault signature*

In the following, several examples will be presented to show the capabilities of the proposed method in diagnostics of rolling bearings. The first two examples deal with numerically simulated signals and the third example presents actual experimental data from a test stand.

**Example 1**

The vibration signal consists of several components indicated in Tab. (1). The vibration spectrum and LCR spectrum is shown in Fig. (7). While the predominant peak in the vibration spectrum is of 50 Hz frequency, it is totally absent in the LCR spectrum and bearing fault peak is dominant instead. Therefore the bearing fault can be detected in the LCR spectrum, easily and with no confusion.

**Tab.1.** *Frequency content of vibration signal of example 1*

| Component | Amplitude [m/s²] |
|---|---|
| 1st Harmonic: 50 Hz | 1 |
| 2nd Harmonic: 100 Hz | 0.5 |
| 3rd Harmonic: 150 Hz | 0.1 |
| External Source: 37 Hz | 0.2 |
| Bearing outer race frequency: 216 Hz | 1.2 |
| Background Gaussian noise | 0.1 |



**Fig. 7.** *Vibration and LCR frequency spectrum of vibration signal of example 1(bearing outer race fault)*

**Example 2**

In this example the vibration signal consists of several components indicated in Tab. (2). The vibration spectrum and LCR spectrum are shown in Fig.(8). A twenty - point window is used for computation of LCR. While the predominant peak in the vibration spectrum is of 50 Hz frequency and the bearing fault is not distinguishable, the 50 Hz peak is reduced in the LCR spectrum and the bearing fault peak of 108 Hz frequency is dominant instead. The harmonic vibration of 110 Hz frequency close to the bearing fault frequency is also omitted in the LCR.

**Tab. 2.** *Frequency content of vibration signal of example 2*

| Component | Amplitude [m/s²] |
|---|---|
| 1st Harmonic: 25 Hz | 3 |
| 2nd Harmonic: 50 Hz | 4 |
| External Source: 110 Hz | 1 |
| Bearing outer race: 108 Hz | 2 |
| Background noise | 0.5 |



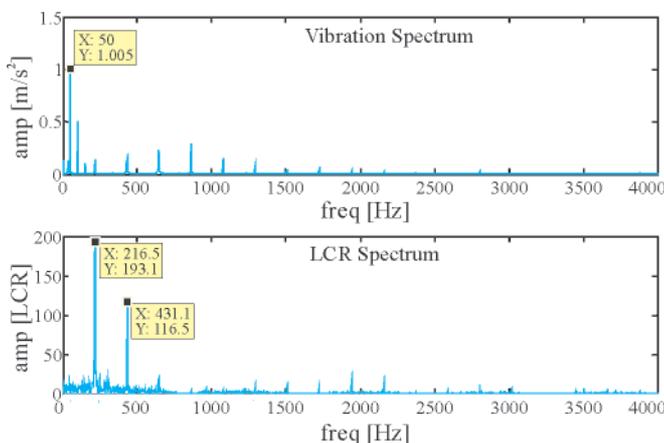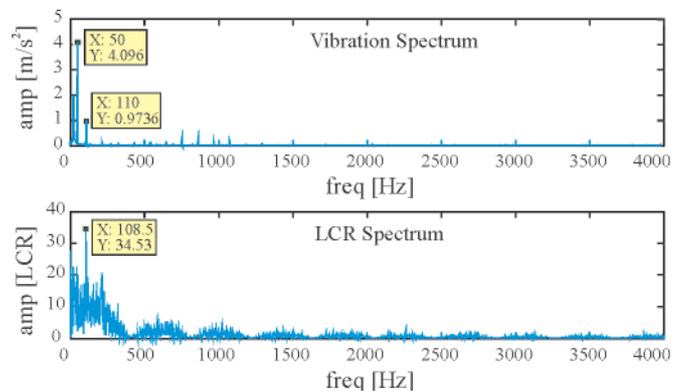**Fig. 8.** *Comparison of vibration and LCR frequency spectrum of vibration signal of example 2(bearing outer race fault)*

**Example 3**

In this example vibration of a self-aligning ball bearing of 1206 type with an inner race defect is investigated. The test stand is shown in Fig (9). The load is applied to the tested bearing by four screws and measured by a load cell. The shaft is driven by an electrical motor through a variable - speed gear box and belt. The speed can be adjusted in the range from 450 to 4000 rpm. Vibration measurements were made with the use of an Endevco 2235 accelerometer with 20 kHz sampling rate. The inner race defect was introduced artificially by means of EDM[1].

The shaft speed of 25 Hz and the inner race defect frequency of 206 Hz is applied in this test. The inner race defect produces an amplitude - modulated signal which shows sidebands of rotation speed around harmonics of inner race frequency. The vibration spectrum, envelope spectrum and LCR spectrum is shown in Fig. (10-a) through Fig. (10-c). In the vibration spectrum, 206 Hz frequency response is almost hidden and cannot be detected from the background noise. However the envelope analysis which is a powerful method, reveals the response peak at 206 Hz and sidebands spaced at +/-25 Hz around it and its harmonics. It is notable that the source of the peak at 25 Hz in the envelope spectrum is amplitude modulation

---

[1] Electrical Discharge Machining

of bearing fault symptom. The LCR produces a spectrum similar to envelope with a more distinct peak at 206 Hz. This example validates usage of LCR in rolling bearing diagnosis.
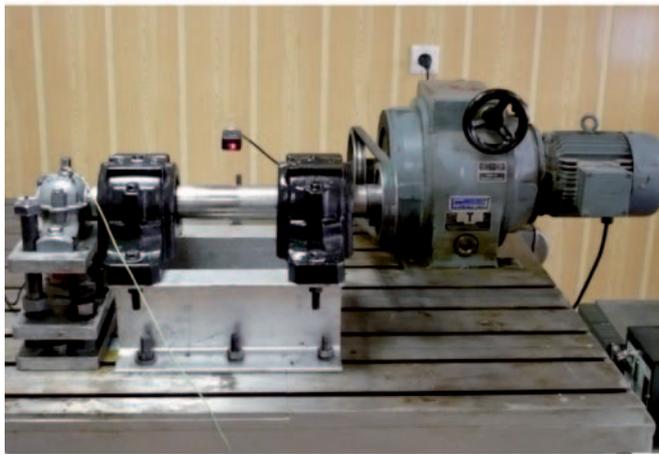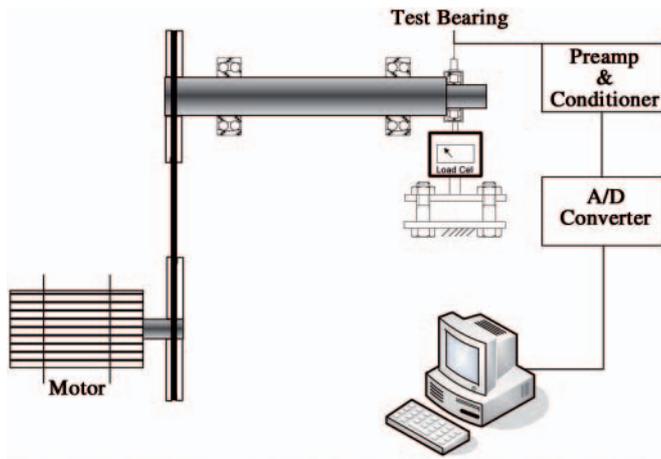


**Fig. 9.** *Experimental test stand for testing the bearings*

## CONCLUSIONS

In this paper a new method for detection of bearing faults by using vibration analysis is described. The method called Local Curve Roughness (LCR) employs the concept of a curve roughness in a moving window to distinguish between sine waveform and impulsive waveform. The latter one generated by bearing defects is amplified as a result of the above mentioned transformation and this way it can be easily detected in the LCR spectrum. The attached examples showed the advantage of the method as compared with normal vibration spectrum analysis. Several examples given in the paper demonstrated effectiveness of the method. The examples of outer race and inner race fault are included in numerically simulated and actual test data. The simplicity of the proposed method is a key advantage for industrial applications as compared with more advanced techniques for detecting bearing faults.

### NOMENCLATURE

b       – Impulse width [sec]
d       - Window width [sec]
f       - Frequency [Hz]
$f_s$   - Sampling frequency [Hz]
h       - Impulse response function
LCR  - Local Curve Roughness [-]
q       - Load [N]
$R_m$  - Bearing pitch radius [m]
t        - Time [sec]
T       - Time period [sec]
x       - Vibration signal
θ       - Rotation angle [rad]
$\omega_c$  - Cage rotation frequency [rad/sec]
$\omega_d$  - Damped natural frequency [rad/sec]
$\omega_c$  - Natural frequency [rad/sec]
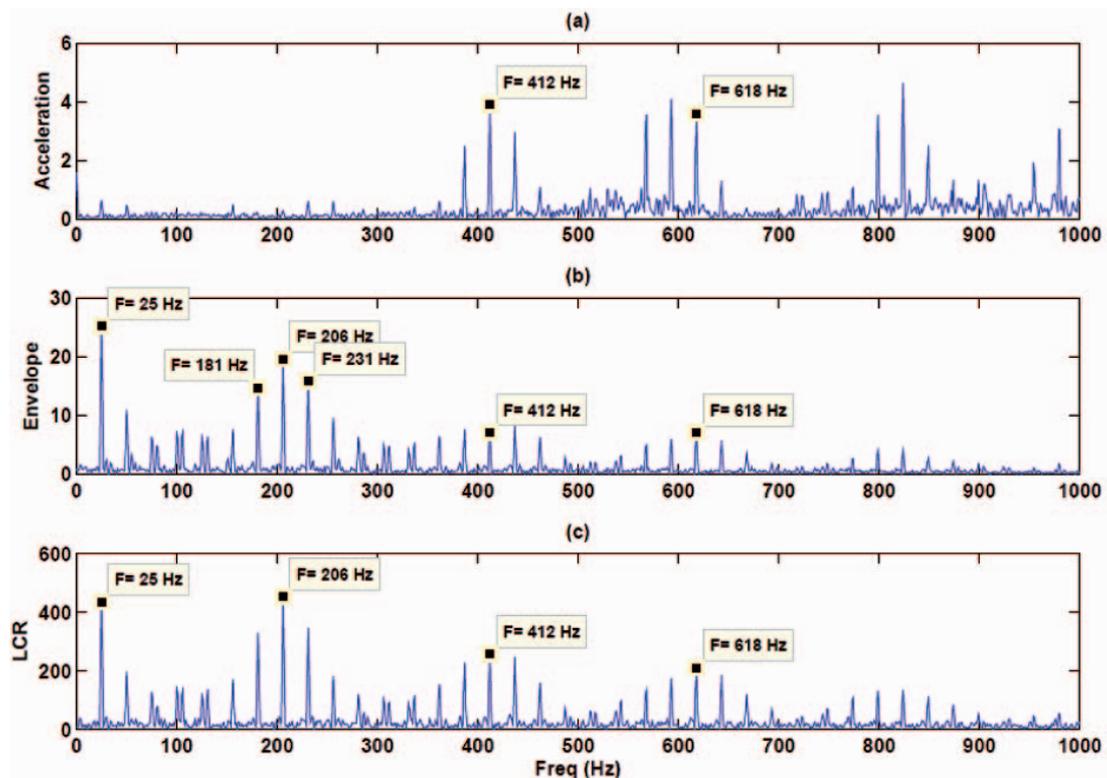Ω      - Shaft rotation speed [rad/sec]
ζ       - Damping ratio [-]



**Fig. 10.** *Comparison of:* **a)** *Vibration spectrum,* **b)** *Envelope spectrum and* **c)** *LCR spectrum of vibration signal of example 3 (bearing with inner race fault)*

## BIBLIOGRAPHY

1. Zaretski E. V., Poplawski J. V., and Miller C. R.: *Rolling Bearing Life Prediction - Past, Present, and Future*. NASA, NASA/TM-2000-210529, 2000

2. Tandon N. and Choudhury A.: *A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings*. Tribology International, Vol. 32, No. 32, 1999, pp. 469-480.

3. Tandon N. and Nakra B. C.: *Vibration and acoustic monitoring techniques for the detection of defects in rolling element bearings - a review*. Shock and Vibration Digest, Vol. 24, No. 3, 1992, pp. 3-11.

4. Taylor J. I.: *Identification of bearing defects by spectral analysis*. Trans. ASME, Journal of Mechanical Design, Vol. 102, 1980, pp. 199-204.

5. Berry J. E.: *How to track rolling element bearing health with vibration signature analysis*. S V, Sound and Vibration, Vol. 25, No. 11, 1991, pp. 24-35.

6. McFadden P. D. and Smith J. D.: *Vibration monitoring of rolling element bearings by the high frequency resonance technique-a review*. Tribology International, Vol. 17, No. 1, 1984, pp. 3-10.

7. Dyer D. and Stewart R. M.: *Detection of rolling element bearing damage by statistical vibration analysis*. Trans. ASME, Journal of Mechanical Design, Vol. 100, No. 2, 1978, pp. 229-235.

8. Antoni J.: *The spectral kurtosis: a useful tool for characterizing non-stationary signals*. Mechanical Systems and Signal Processing, Vol. 20, No. 2, 2006, pp. 282-307.

9. Bolaers F., Cousinard O., Marconnet P., and Rasolofondraibe L.: *Advanced detection of rolling bearing spalling from de-noising vibratory signals*. Control Engineering Practice, Vol. 12, No. 2, 2004, pp. 181-190.

10. Dron J. P., Bolaers F., and Rasolofondraibe L.: *Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: application to the detection of defects in ball bearings*. Journal of Sound and Vibration, Vol. 270, No. 1-2, 2004, pp. 61-73.

11. *The shock pulse method for determining the condition of antifriction bearings*. SPM Instruments AB, 1980

12. Butler D. E.: *The shock pulse method for the detection of damaged rolling bearings*. NDT International, Vol. 6, No. 2, 1973, pp. 92-95.

13. McFadden P. D. and Toozhy M. M.: *Application of synchronous averaging to vibration monitoring of rolling element bearings*. Mechanical Systems and Signal Processing, Vol. 14, No. 6, 2000, pp. 891-906.

14. Junsheng C., Dejie Y., and Yu Y.: *Application of an impulse response wavelet to fault diagnosis of rolling bearings*. Mechanical Systems and Signal Processing, Vol. 21, 2007, pp. 920-929.

15. Li C. J. and Ma J.: *Wavelet Decomposition of Vibrations for Detection of Bearing-Localized Defects*. NDT & E International, Vol. 30, 1997, pp. 143-149.

16. Mori K., Kasashima N., Yoshika T., and Ueno Y.: *Prediction of Spalling on Ball Bearing by Applying the Discrete Wavelet Transform to Vibration Signals*. Wear, Vol. 195, 1996, pp. 162-168.

17. Ocak H., Loparo K. A., and Discenzo F. M.: *Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics*. Journal of Sound and Vibration, Vol. 302, 2007, pp. 951-961.

18. Purushotham V., Narayanan S., Suryanarayana and Prasad A. N.: *Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition*. NDT & E International, Vol. 38, 2005, pp. 654-664.

19. Rubini R. and Meneghetti U.: *Application of the Envelope and Wavelet Transform Analyses for the Diagnosis of Incipient Faults in Ball Bearings*. Mechanical Systems and Signal Processing, Vol. 15, No. 2, 2001, pp. 287-302.

20. Tse P. W., Peng Y. H., and Yam R.: *Wavelet Analysis and Envelope Detection For Rolling Element Bearing Fault Diagnosis - Their Effectiveness and Flexibilities*. Trans ASME, Journal of Vibration and Acoustics, Vol. 123, 2001, pp. 303-310.

21. Yang D. M., Stronach A. F., and McConnel P.: *Third-Order Spectral Techniques for Diagnosis of Motor Bearing Condition Using Artificial Neural Networks*. Mechanical Systems and Signal Processing, Vol. 16, No. 2-3, 2002, pp. 391-411.

22. Antoni J.: *Cyclic spectral analysis of rolling-element bearing signals: Facts and fictions*. Journal of Sound and Vibration, Vol. 304, 2007, pp. 497-529.

23. Du Q. and Yang S: *Application of the EMD method in the vibration analysis of ball bearings*. Mechanical Systems and Signal Processing, Vol. 21, No. 6, 2007, pp. 2634-2644.

24. Junsheng C., Dejie Y., and Yu Y.: *A fault diagnosis approach for roller bearings based on EMD method and AR model*. Mechanical Systems and Signal Processing, Vol. 20, No. 2, 2006, pp. 350-362.

25. Behzad M., Bastami A.R.: *Concept of Roughness of Vibration in Rolling Bearings Diagnosis*. Proceedings of ISMA Conference, Lueven, 2008

26. Sunnersjo C. S.: *Varying compliance vibrations of rolling bearings*. Journal of Sound and Vibration, Vol. 58, No. 3, 1978, pp. 363-373.

27. McFadden P. D. and Smith J. D.: *Model for the vibration produced by a single point defect in a rolling element bearing*. Journal of Sound and Vibration, Vol. 96, No. 1, 1984, pp. 69-82.

28. C. M. Harris and A. G. Piersol: *Shock and Vibration Handbook*. McGraw-Hill, 5th Ed. 2002

29. M. Behzad, A. R. Bastami, D. Mba: *A New Model for Estimating Vibrations Generated in the Defective Rolling Element Bearings*, ASME Journal of vib & Acoustics, Vol. 133, Aug 2011, pp. 041011-1 to 8

---

## CONTACT WITH THE AUTHORS

Mehdi Behzad, Prof.
Abbas Rohani Bastami, M. Sc.
Mechanical Engineering Department,
Sharif University of Technology
11155-9567, Azadi Avenue,
Tehran, IRAN
email: m_behzad@sharif.edu

# Life tests of a rotary single-stage magnetic-fluid seal for shipbuilding applications

**Leszek Matuszewski**, Ph.D.
Gdansk University of Technology
**Zbigniew Szydło**, Ph.D.
AGH University of Science and Technology

## ABSTRACT

*Use of the magnetic fluid seal technology in water is much more difficult technological problem in comparison to gas or vacuum environment. Some seals constructed up to time been are designed as hybrid, two stage sealing structures with shields or mechanical seals used as protection measures to the magnetic fluid seal. Anyway, there is always problem with immediate contact between magnetic fluid and the environmental liquid. In the paper are presented results of experiments carried out with one stage rotating magnetic fluid seal operating in immediate contact with utility water. The special test procedures have been elaborated and practically used, with the aim to define main characteristics of magnetic fluid seals and to simplify experiments. These characteristics were: critical pressure, critical motion velocity and working life of the seal. Four grades of commercial, silicon based magnetic fluids of hydrophobic properties were tested in a single stage sealing system. The results of tests are presented in the form of tables and diagrams. It is concluded that hydrophobic commercial magnetic fluids could be efficiently used in rotating shaft seals however in a limited range of motion velocity.*

**Keywords**: shaft sealing; magnetic fluid; utility water; seal durability; critical pressure; critical motion velocity

## INTRODUCTION

The sealing technique of shafts, covers, pipe penetrations and many other mechanical units with the use of ferro-fluids has been successfully applied to machinery operating in rotational or reciprocating motion. Many experimental and theoretical investigations have been devoted to problems of design, choice of materials and ferro-fluids for seals intended for operating in gas environment or vacuum, owing to which very effective and simple solutions of sealing problems in many technical devices, have been obtained [1].

Theoretical elaboration and design solution of ferro-fluid seals intended for operating in sea water and other liquids is much more difficult than in the case of gas environment or vacuum. Sealing a liquid requires many additional factors which appears in direct interphase contact of two environments being in the same physical state, e.g. a ferro-fluid and sealed liquid, to be taken into account. An essential condition for correct operation of a seal in such circumstances is complete lack of susceptibility to mixing together sealing fluid and sealed one. It means that in the seal in question a ferro-fluid of properties ensuring high interphase surface tension on the phase separation surface between sealing (magnetic) fluid and sealed (non-magnetic) one, should be used.

In static conditions, when seal's elements do not mutually displace, fulfillment of the conditions, at correct design of the unit, ensures a sufficient life of seals of the kind; hence they may be successfully used in shipbuilding. Lack of susceptibility

to mixing and high interphase surface tension may appear not sufficient factors if a magnetic-fluid (MF) seal is applied in its motion conditions or in case of sealed liquid flow, e. g. around ship propeller shaft. Relative displacement of two non-mixing together liquids results in generating surface instability at the interphase boundary [2, 3], whose intensity fast increases over a limiting difference of velocity, that causes magnetic-fluid sweeping and loss of system sealing. Research work on operation of MF seals in liquid environment has been conducted by some research centres, however only a few relevant reports are available. In the publications [4, 5, 6] was described research on possible operation of MF seals in sea water depending on pressure and shaft peripheral speed. Anyway these studies were mainly focused on defined technological goals without more general considerations.

In [7] was described a hybrid sealing system for liquid pump, where MF seal is introduced as the second sealing stage (two - stage system), placed in series behind a mechanical one. However the author did not attach any operational test results of the solution.

One of the particular examples of MF seal applications in interaction with liquid is the seal of precise rotor of a pump used in medicine for blood pumping during operations. The pump including the seal was elaborated by a team under supervision of Mitamura, Y. [8].

The authors of the elaboration carried out experimental work on MF seals operating in utility water environment. The research was conducted on a special test stand adjusted to MF

seal operation in liquids [9]. In view of utilitarian character of the research, MF multi-stage seals being models of real ferro-fluid seals were used, and their operational conditions were limited according to the assigned task. Despite the simplifications, the obtained results demonstrated that it is possible to apply some commercial magnetic fluids to MF seals working in direct contact with water in operational conditions occurring on ships and many other marine technical objects [10, 11, 12].

Loss of tightness of MF seal operating in water results from complex physical processes and depends also on many geometrical and motional factors. Diagnosis of the processes and assessment of possible operation of MF seals in water require to perform a wide range research with the use of various magnetic fluids of hydrophobic properties in seals of different geometry and in various operational conditions. And, basic tests should be performed with the use of a single-stage seal.

MF single-stage seals are very seldom used in sealing systems of rotary shafts of propellers or similar technical devices. In practice multi-stage seals are used and number of their stages is usually even [1] that results from magnetic circuit physical features. However investigation of a single sealing stage is a very essential experimental task as it can make it possible to recognize phenomena which occur in the seal, in a more transparent way than in the case of a multi-stage seal. This concerns especially MF seals operating in water (or other liquids) as in such systems to determine the instant of loss of tightness is more difficult than in the case of testing the seals in gas environment or vacuum. Investigations of MF single-stage seal operating in gas environment are presented in the publication [13] devoted mainly to problems of pressure to puncture and torque of seal resistance to motion. In the investigations the test stand called shortly MASTL1 – V2 was used. The stand is composed of a test head especially adjusted for performing assigned research tasks, driving system consisted of an electric motor and its controller, as well as a measuring system consisted of sensors and transducers for measuring values of pressure, motor torque and operational temperature. The stand is equipped with a multi-channel measuring system for acquiring, processing and recording measurement data.

Four kinds of magnetic fluids of hydrophobic properties were tested. These are magnetic fluids produced by FERROLABS Inc., a Russian firm, located in Sterling, VA, USA [17]. The investigations were conducted in the conditions of equal magnetic field (equal number of magnets and their arrangement in seal magnetic system).

**TEST STAND**

The MAST L1–V2 test stand is that used for testing magnetic fluid seals operating in contact with water or other liquids, located in the Laboratory of Seals and Application of Magnetic Fluids, AGH University of Science and Technology in Krakow. Dimensions of the head make it possible to place within it a magnetic fluid seal of 50 mm diameter equivalent to that of propeller shafts of fish cutters and smaller boats. In future tests to increase the diameter to at least 200 mm corresponding to that of shafts of medium ships is recommended. It should be checked whether and to which extent scale effect may occur. Only a small amount of magnetic fluid, namely about 50-200μl, is necessary to fill gaps of the stages sealing the 50 mm shaft.

The MAST L1-V2 test stand makes it possible to determine static puncture pressure of seal, limiting operational speed of seal, and its life for a given rotational speed up to 120 s$^{-1}$ and pressure up to 0.1 MPa.

### *Design of the test head used in MAST L1-V2 test stand*

In Fig. 1 is presented the head of the MAST L1-V2 test stand for research on magnetic - fluid seals operating in liquid environment. In the testing part of the head is placed the testing sleeve of sealing stages, 1, located on the hollow rotary shaft 18 seated in the slide bearings 19 mounted at the motionless casing 20. On the casing 20 the test head pivoting body 6 is seated in the slide bearings 21. In the body 6 is placed: the set of permanent magnets, 3, with pole shoes 2, and the sleeve 1, which form, together with magnetic fluid, a closed magnetic circuit. In the body 6 is fixed the casing of the sealed fluid chamber 5 closed with the transparent plate 7.
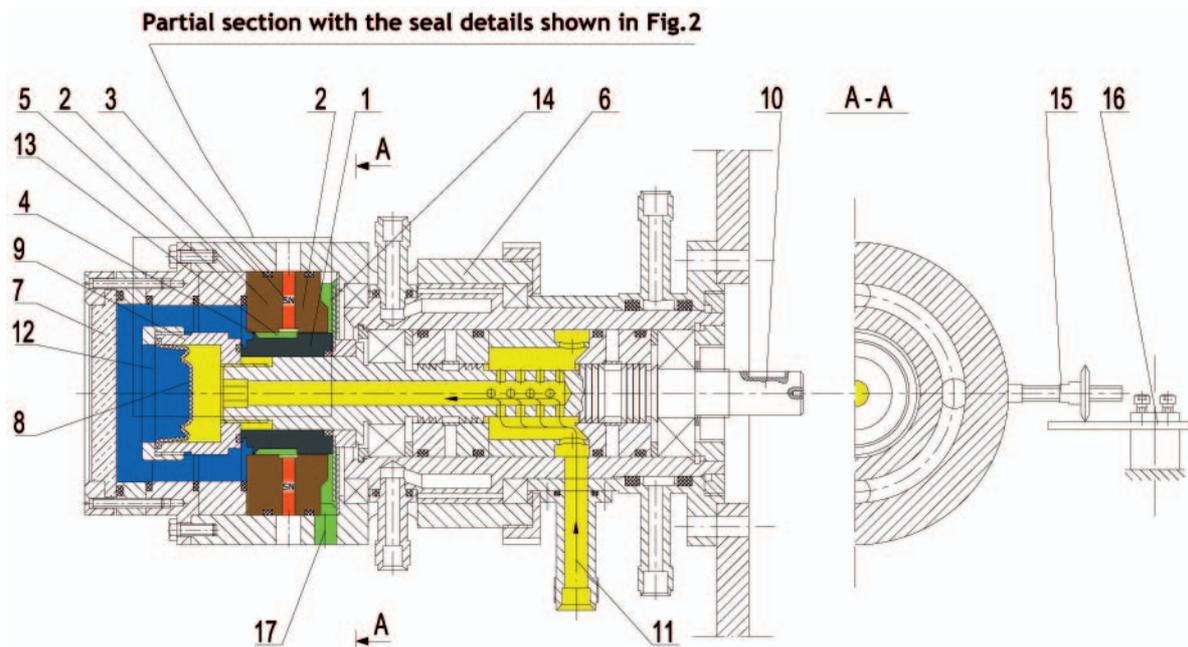


*Fig. 1. The structure of MAST L1-V2 head intended for testing magnetic fluid seal operating in liquid environment: 1 – sigle-stage sleeve, 2 – pole shoes, 3 – permanent magnets, 4 – magnetic fluid, 5 – test chamber casing, 6 – test head pivoting body, 7 – test chamber transparent cover, 8 – rubber membrane, 9 – rubber membrane holder, 10 – rotary shaft, 11 – compressed air passage, 12 – set pressure zone of fluid-filled chamber, 13 – fluid penetration zone, 14 – fluid outflow zone after leakage from tested seal, 15 – thrust arm of torque meter, 16 – extensometric beam, 17 – leakage draining hole*

*Fig. 2. Partial section to show design solution details of the tested single-stage sealing system*

Inside the chamber 5, is placed the rubber membrane holder 9, mounted on the threaded end of the shaft 10. In the holder 9, is fixed the membrane 8, which separates the chamber zone filled with the sealed fluid, 12, from the zone of the compressed air supplied through the passage 11 and the hole in the shaft 10. The 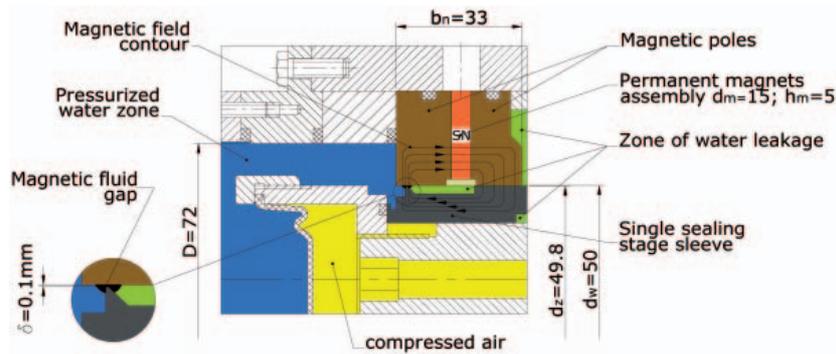compressed air is delivered to the elastic membrane to rise pressure in the sealed fluid. Balance obtained between fluidal medium pressure and gaseous medium pressure means that the seal prevents the fluid from leaking. If the magnetic fluid seal loses its tightness the sealed fluid penetrates to the zone 13, next finds its way to the fluid outflow zone 14, and further, through the leakage draining hole 17, flows out to the leakage tank (Fig. 4, item 8). The sealed fluid loss in the zone 12, makes the compressed gas pressure dropping, which is recorded by the pressure sensor. The pressure indications will serve as auxiliary diagnostic signals if direct observation of dropping leakage is not possible.

In Fig. 2 the design solution details of the tested single-stage sealing system, are presented. There are demonstrated: path of magnetic field lines, location of the zones of compressed air and water, as well as leakage zone and details of the seal with indicated location of the magnetic fluid and height of the gaps tested in the seal.

### *Drive and control &measurement systems of the MAST L1-V2 test stand*

In Fig. 3 the schematic diagram of the drive and control & measurement systems of the MAST L1-V1 test stand, are presented. The electric motor 1 drives the test head shaft 3 through the elastic coupling 2. Basically, the test stand's casing is composed of two parts: the motionless part fixed in the base 7 through a special fixing device, and the moving part set through bearing system in the motionless part where the pole shoes and the permanent magnets are placed. During testing the seal, air is under pressure delivered from the tank 17, by using the air compressor 16, to the inlet stub pipe 10. Pressure of the air delivered to the inlet stub pipe 10 is measured by using the Danfoss MBS 32 pressure sensor, 21, of the measuring range from 0 to1 MPa, as well as by means of an indication manometer of the measuring range from 0 to 0.1 MPa. Rotational speed of the driving motor 1 is controlled by means of the frequency converter PCZ which makes it possible to measure driving shaft rotational speed and power absorbed by the motor.

Signals of pressure, temperature, torque, rotational speed and absorbed power are sent through the measurement channels 24, 25, 26, 27 and 28 to the control & supply panel PSZ, and next to the NI USB-6009 eight-channel measurement card KP (of National Instruments), which cooperates with the computer PC by making use of the LabVIEW 8.5 software, also of National Instruments origin.
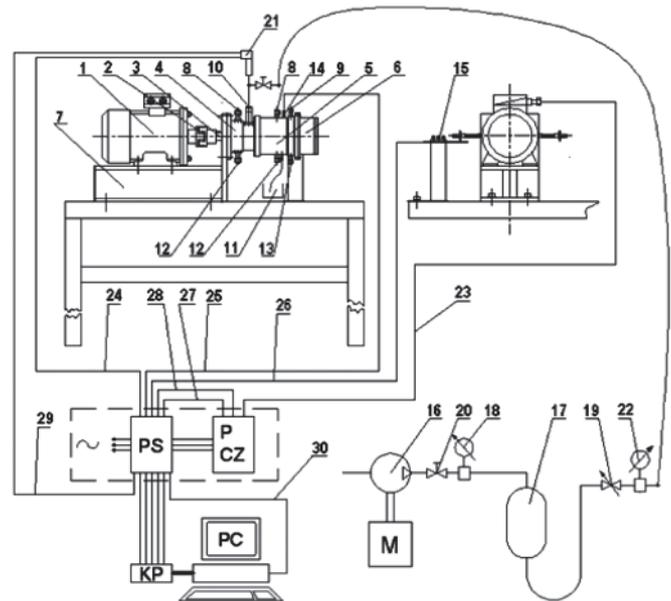


*Fig. 3. Schematic diagram of the drive and control &measurement systems of the MAST L1-V1 test stand intended for the testing of magnetic-fluid seals: 1 – electric motor; 2 – coupling; 3 – rotary shaft; 4 – casing of shaft bearing; 5 – pivoting casing of head's bearings; 6 – test chamber's casing; 7 – base; 8 – inlet stub pipes of cooling system of shaft sealing; 9 – inlet stub pipe of tested seal cooling; 10 – stub pipe of compressed air delivery system; 11 – leakage tank; 12 – outlet stub pipes of cooling system of shaft sealing; 13 – outlet stub pipe of tested seal cooling; 14 – temperature sensor; 15 – extensometric beam for torque measuring; 16 – air compressor; 17 – compressed air tank; 18 – pressure reducing valve; 19 – driving motor supply control valve; 20 – cut-off valve; 21 – pressure measuring transducer; 22 – manometer; 23 – driving motor supply; 24, 25, 26, 27, 28 – measurement channels of: pressure, temperature, torque, rotational speed and power, respectively; 29, 30 – supply of pressure transducer and computer measurement system; PS – control & supply panel; PCZ – frequency converter; KP – measurement card; PC – PC computer; M – air compressor motor*

In Fig. 4 is presented a photo of the MAST L1 test stand together with the control & supply panel and computer measurement system intended for aiding and recording run of the testing.

### CONCEPT AND SCOPE OF THE TESTS

Many factors affect a seal material and technological characteristics when water and particularly sea water is operating environment. These are: environmental pressure, motion velocity of the seal combined with the flow intensity and velocity of water, temperature and chemical reactivity of water against the magnetic fluid and other components of the seal. Some of the factors e.g. environmental pressure and the mixing ability of water and magnetic fluid may result in instantaneous leakage when others like motion velocity of the
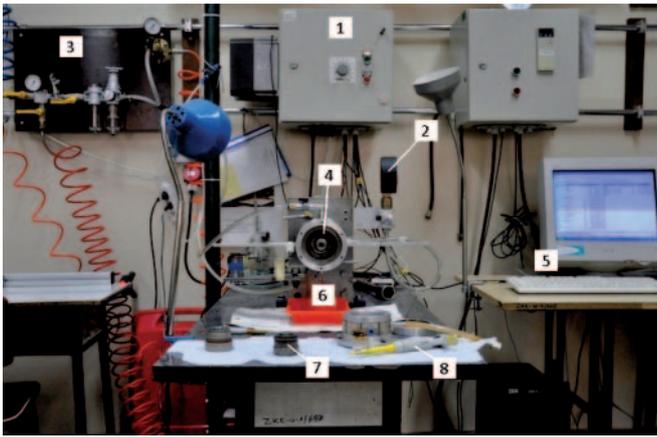
*Fig. 4. MAST L1 stand shown with open chamber for preparation a test run. control & supply panel and of the testing: **1** – control & supply panel, **2** – manometer to measure pressure in test chamber, **3** – compressed air preparation supply system, **4** – test head open to show magnetic poles, **5** – computer aided measurement and record system, **6** – leakage tank, **7** – single stage seal sleeve, **8** – precise magnetic fluid metering device*

seal, temperature and chemical activity are hazardous in view of long life effectiveness of the seal.

In tests were used magnetic fluids with high mixing resistance to water, and test pressures were selected far below critical pressure of the seal to avoid the risk of instantaneous break of the seal.

Owing to long time tests necessary to define time durability of the seal it was decided to carry the single factor tests based on the assumption that the motion velocity is the factor deciding on the seal durability in water environment with outer pressure maintained below the critical pressure and temperature within 20°C - 25°C. This assumption has been supported by results of earlier research [12].

The research concept has been based on the preliminary assumption that the factor decisive of life of seal which operates in water environment is speed of its motion.

To realize the research concept a scope of work comprising both short-term and long-term tests, was elaborated. The short-term tests were aimed at determining critical conditions for seal operation, i.e. seal critical pressure and critical speed of its motion. The long-term tests were focused on determining life of the seal depending on its motion speed.

The critical pressure was measured during the static test (motionless seal) at water pressure continuously increasing in the test chamber till occurrence of water leakage.

Pressure value at which the leakage takes place has been assumed critical. The pressure was used to determine values of water pressure to be set in the test chamber both during critical motion measurements and long-term life tests of the seal. In

the research use was made of the notion of the relative pressure $p/p_{kr}$, i.e. the ratio of the set pressure and the critical pressure, characteristic for a given type of tested seal.

The seal motion critical speed was measured at the relative water pressure $p/p_{kr} = 0.5$ set in the test chamber. The measurement consisted in continuous increasing seal rotational speed till occurrence of water leakage. The equal rate of rotational speed increasing (rotational acceleration) amounted to $1s^{-2}$ (the rotational speed increasing from 0 to $120\ s^{-1}$ within the time of 120 s) was applied. Speed value at which leakage occurs was taken as the seal motion critical speed. The measurement of the critical pressure and seal motion speed was used for preliminary comparative assessment of the applied magnetic fluids.

Life of the seal in function of its motion speed was measured during the long-term test consisting in the seal operation under constant rotational speed and stationary pressure conditions in the test chamber (the relative pressure $p/p_{kr} = 0.5$) and seal temperature, continued up to occurrence of water leakage. Total number of seal's rotations up to leakage was assumed equivalent to the seal's life in the set working conditions.

Although occurrence of leakage was taken as the main criterion for correct operation of seal, during the tests also friction torque, seal temperature and pressure, and shaft rotational speed was measured. The quantities served to control the tests running.

The tests in question were conducted according to the following programme:
- The static critical pressure test of the seal.
- The critical operational speed test of the seal at the relative pressure $p/p_{kr} = 0.5$.

The life test of the seal at the relative pressure - $p/p_{kr} = 0.5$ under various motion speed values beginning from the maximum seal motion speed close to critical one up to the minimum speed under which the seal's life was equal at least to 1 mln cycles.

The tests were performed with the use of four kinds of magnetic fluid applied to the single-stage seal with the magnetic-fluid-filled gap height $\delta = 0.1mm$.

## MATERIALS USED IN THE TESTS

### *Magnetic fluids*

In the tests four kinds of magnetic fluids produced by Ferrolabs Inc. [14] were used.

In Tab.1 are presented characteristics of the delivered magnetic fluids, based on the data available from their producer.

*Tab.1. Specification of the magnetic fluids used in the tests*

| No. | Magnetic fluid | Bearing liquid | Magnetic particles | Magnetization of saturation, $M_s$ [kA/m] | Plastic viscosity in 20°C, $\eta_{pl}$ [Pa·s] | Freezing point [°C] | Critical temperature [°C] | Operational temperature [°C] |
|---|---|---|---|---|---|---|---|---|
| 1 | **FLS 040.040** | Silicon liquid | Magnetite $Fe_3O_4$ | 45.00 | 0.550 | -100 | +200 | -70…+150 |
| 2 | **FLS 250.020** | Silicon liquid | Magnetite $Fe_3O_4$ | 25.00 | 1.120 | -90 | +200 | -60…+100 |
| 3 | **FLA 003.45** | Siloxane | Magnetite $Fe_3O_4$ | 45.00 | 0.450 | – | +200 | -70…+150 |
| 4 | **FLA 002.25** | Siloxane | Magnetite $Fe_3O_4$ | 25.00 | 0.725 | – | +250 | -70…+150 |

| Fluid amount [µl] | Gap height [mm] | Critical pressure $p_{kr}$ [MPa] | Relative pressure $p/p_{kr}$ | Set testing pressure P [MPa] | Critical speed [$s^{-1}$] | Testing rotational speed [ $s^{-1}$] | Number of revolutions to leakage | Time to leakage [s] |
|---|---|---|---|---|---|---|---|---|
| 100 | 0.1 | 0.0776 | 0.5 | 0.039 | 73 | 73 | 730 | 20 |
| | | | | | | 60 | 30000 | 500 |
| | | | | | | 40 | 200000 | 5000 |
| | | | | | | 20 | 9800000 | 49000 |

The magnetic fluids used in the tests were produced on the basis of different kinds of silicon fluids of strong hydrophobic properties, that is the necessary condition for effective operation of MF seal in water. The fluids marked FLS 250.020 and FLS 040.040 are standard ones intended for the applying to seals, and the fluids marked FLA 002.25 and FLA 003.45 are those for using in acoustic devices, mainly loudspeakers.

The fluids for acoustic devices are characteristic of good heat abstraction, and satisfactory results as to their range of tightness keeping, obtained from their tests in seals.

In selecting the fluids also their saturation magnetization and viscosity values were taken into account. The two fluids of lower saturation magnetization but higher viscosity (FLS 250.020 and FLA 002.25) and the two fluids of higher saturation magnetization but lower viscosity (FLS 040.040 and FLA 003.45) were selected. This way it was possible to preliminarily assess impact of the parameters on sealing effectiveness in conditions of operation in water.

### *Permanent magnets*

Magnetic field in the seal was generated by means of the set of permanent disk magnets circumferentially placed between the pole shoes of the seal's magnetic system. The commercial Neodymium 38 high-energy permanent magnets of the coercivity $H_c = 912$ kA/m and the magnetic energy density $(BH)_{max} = 294$ kJ/m³ [15] were used. The magnetic system was consisted of 14 disks of $\Phi$15x5mm in size, circumferentially evenly distributed in the space between the pole shoes.

### *Pole shoes and sleeves of single sealing stage (lip)*

The pole shoes and single stage sealing sleeves were made of a low-carbon steel of the relative magnetic permeability $\mu_r > 2000$.

## CHARACTERISTICS OF THE TESTED SEALS

The crucial geometrical and operational data of the tested seal are specified below.

**Geometrical data** (Fig.2):
- internal diameter of pole shoes (nominal diameter of seal): $d_w = 50$ mm
- internal diameter of test chamber: D = 85 mm
- external diameter of sealing stage sleeve: $d_z = 49.8$ mm
- gap height: $\delta = 0.1$ mm
- number of sealing stages: z = 1
- breadth of the magnetic unit (pole shoes + permanent magnets): $b_n = 33$ mm
- permanent magnet diameter: $d_m = 15$ mm
- permanent magnet thickness: $h_m = 5$ mm
- number of permanent magnets: L = 14

**Operational data**:
- magnetic fluid amount delivered to the gap: 100 µl
- working medium pressure: p = (0.02÷0.1) MPa
- rotational speed range: n = (0 – 120) $s^{-1}$
- working medium: utility water

## RESULTS OF THE TESTS

Results of the tests are given in the form of the tables containing both values of test parameters and results obtained from the tests, as well as in the form of the diagrams where values resulting from the tests are presented in function of the parameters of the tested seals or test duration time.

Results of the tests are given for each of the tested magnetic fluid, separately. In every test was applied the same sealing gap of 0.1 mm in height, to which 100 µl of magnetic fluid was delivered. In the diagrams are also given trend- line equations of the best conformity with the obtained results.

### *Results of the tests with the use of FLS 040.040 magnetic fluid*

In Tab. 2 are collected results of the tests with the use of FLS 040.040 magnetic fluid. The tests were performed at the relative pressure **$p/p_{kr}$ = 0.5**, consequently the pressure of 0.039 MPa was set in the water chamber.

In Fig. 5 are presented the diagrams of number of seal revolutions to the first leakage in function of seal rotational speed for the FLS 040.040 magnetic fluid. The first measurement point depicted in the diagram concerns the test performed at the seal critical speed equal to 73[$s^{-1}$].

The trend-line equation is of the form:

$$y = -7.14\ln(x) + 124.83$$

### *Results of the tests with the use of FLS 250.020 magnetic fluid*

In Tab. 3 are collected results of the tests with the use of **FLS 250.020** magnetic fluid. The tests were performed at the relative pressure **$p/p_{kr}$ = 0.5**, consequently the pressure of 0.031 MPa was set in the water chamber.

In Fig. 6 is presented the relation between number of seal revolutions (cycles) to the first leakage and seal rotational speed for the FLS 250.020 magnetic fluid. The first measurement point depicted in the diagram concerns the test performed at the seal critical speed equal to 22 [$s^{-1}$].

The trend-line equation is of the form:

$$y = -1.54\ln(x) + 31.08$$

### *Results of the tests with the use of FLA 003.45 magnetic fluid*

In Tab. 4 are collected results of the tests with the use of FLA 003.45 magnetic fluid. The tests were performed at the relative pressure **$p/p_{kr}$ = 0.5**, consequently the pressure of 0.043 MPa was set in the water chamber.
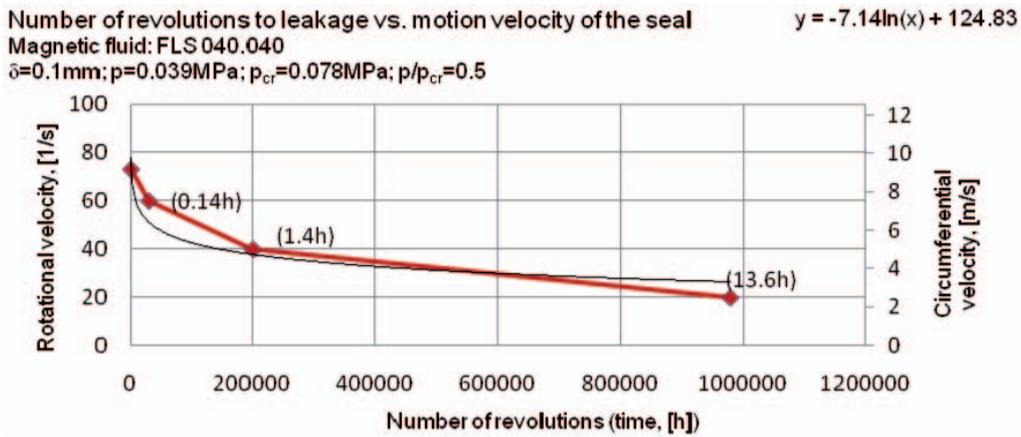
**Fig. 5.** *Life of the seal with the applied FLS 040.040 magnetic fluid for the gap height of 0.1 mm and the set pressure of 0.039 MPa*

**Tab. 3.** *Results of the tests with the use of FLS 250.020 magnetic fluid*

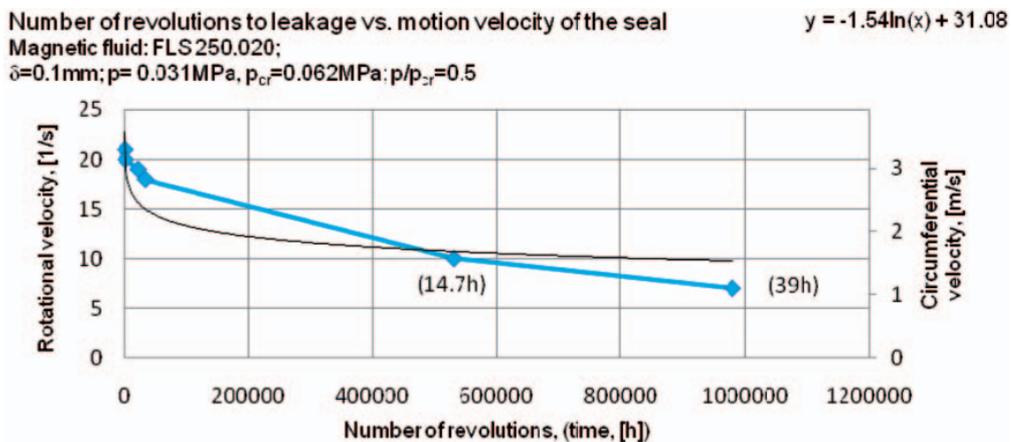| Fluid amount [μl] | Gap height [mm] | Critical pressure $p_{kr}$ [MPa] | Relative pressure $p/p_{kr}$ | Set testing pressure P [MPa] | Critical speed [s⁻¹] | Testing rotational speed [ s⁻¹] | Number of revolutions to leakage | Time to leakage [s] |
|---|---|---|---|---|---|---|---|---|
| 100 | 0.1 | 0.0616 | 0.5 | 0.031 | 22 | 22 | 250 | 12 |
| | | | | | | 20 | 500 | 25 |
| | | | | | | 19 | 20900 | 1100 |
| | | | | | | 18 | 32400 | 1800 |
| | | | | | | 10 | 530000 | 53000 |
| | | | | | | 7 | 980000 | 140000 |



**Fig. 6.** *Life of the seal with the applied FLS 250.020 magnetic fluid for the gap height of 0.1 mm and the set pressure of 0.031 MPa*

**Tab. 4.** *Results of the tests with the use of FLA 003.45 magnetic fluid*

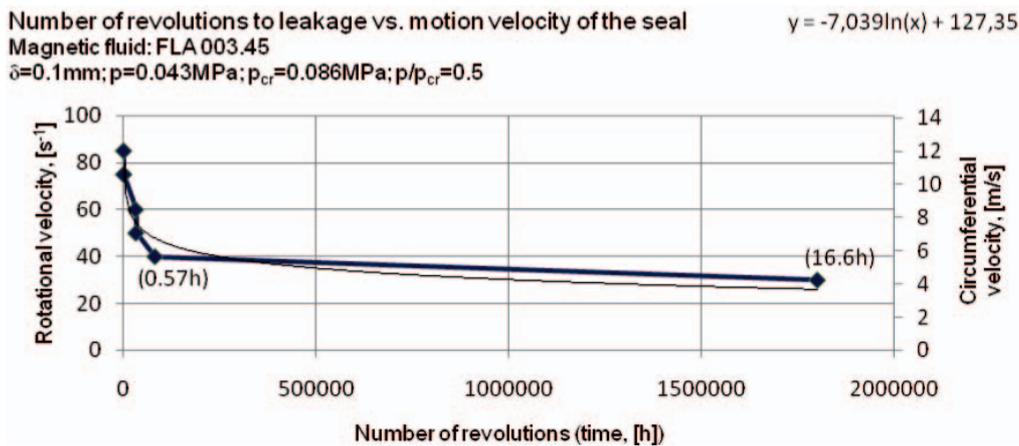| Fluid amount [μl] | Gap height [mm] | Critical pressure $p_{kr}$ [MPa] | Relative pressure $p/p_{kr}$ | Set testing pressure P [MPa] | Critical speed [s⁻¹] | Testing rotational speed [ s⁻¹] | Number of revolutions to leakage | Time to leakage [s] |
|---|---|---|---|---|---|---|---|---|
| 100 | 0.1 | 0.0868 | 0.5 | 0.043 | 82 | 82 | 850 | 22 |
| | | | | | | 75 | 1375 | 28 |
| | | | | | | 60 | 31200 | 520 |
| | | | | | | 50 | 32000 | 640 |
| | | | | | | 40 | 82000 | 2050 |
| | | | | | | 30 | 1800000 | 60000 |

*Fig. 7. Life of the seal with the applied FLA 003.45 magnetic fluid for the gap height of 0.1 mm and the set pressure of 0.043 MPa*

In Fig. 7 is presented the relation between number of seal revolutions (cycles) to the first leakage and seal rotational speed for the FLA 003.45 magnetic fluid. The first measurement point depicted in the diagram concerns the test performed at the seal critical speed equal to 82 [s$^{-1}$].

The trend-line equation is of the form:

$$y = -7.03\ln(x) + 127.35$$

### Results of the tests with the use of FLA 002.25 magnetic fluid

In Tab. 5 are collected results of the tests with the use of FLA 002.25 magnetic fluid. The tests were performed at the relative pressure $p/p_{kr} = 0.5$, consequently the pressure of 0.022 MPa was set in the water chamber.

In Fig. 8 is presented the relation between number of seal revolutions (cycles) to the first leakage and seal rotational speed for the FLA 002.25 magnetic fluid. No test was performed at the rotational speed equal to critical one for the fluid in question.

The first measurement point depicted in the diagram concerns the test performed at the seal critical speed equal to 50 [s$^{-1}$].

The trend-line equation is of the form:

$$y = -3.85\ln(x) + 81.07$$

## COMPARISON AND DISCUSSION OF THE TEST RESULTS

In Tab. 6 are collected the magnetic and viscosity properties of the tested fluids and values of statical critical pressures and critical speeds obtained from the tests of the seal.

Critical speed is one of the crucial parameters which characterize applicability of a magnetic fluid to operate

*Tab. 5. Results of the tests with the use of FLA 002.25 magnetic fluid*

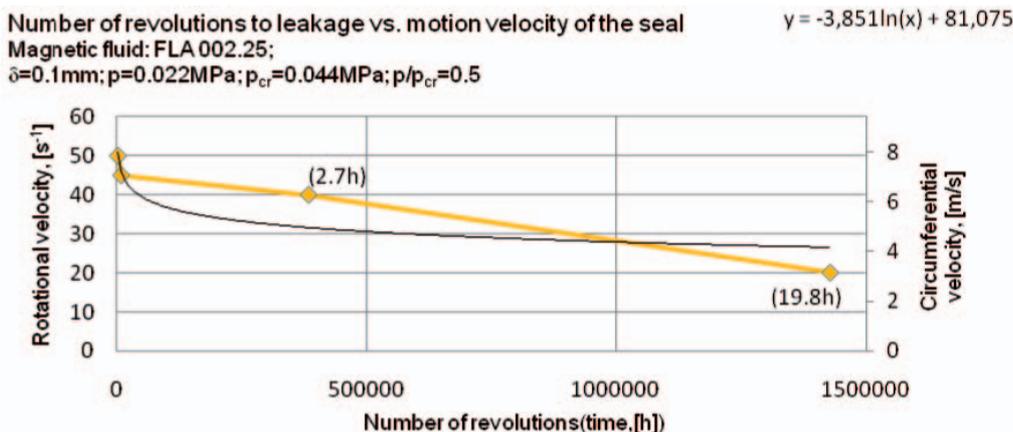| Fluid amount [µl] | Gap height [mm] | Critical pressure $p_{kr}$ [MPa] | Relative pressure $p/p_{kr}$ | Set testing pressure P [MPa] | Critical speed [s$^{-1}$] | Testing rotational speed [ s$^{-1}$] | Number of revolutions to leakage | Time to leakage [s] |
|---|---|---|---|---|---|---|---|---|
| 100 | 0.1 | 0.0443 | 0.5 | 0.022 | 55 | 55 | Not measured | |
| | | | | | | 50 | 2000 | 40 |
| | | | | | | 45 | 9000 | 200 |
| | | | | | | 40 | 384000 | 9600 |
| | | | | | | 20 | 1428000 | 71400 |



*Fig. 8. Life of the seal with the applied FLA 002.25 magnetic fluid for the gap height of 0.1 mm and the set pressure of 0.022 MPa*

**Tab. 6.** *Magnetic and viscosity properties of the tested magnetic fluids, as well as values of static critical pressure and critical speed obtained from the tests*

| Magnetic fluid | Saturation magnetization $M_s$ [kA/m] | Plastic viscosity at 20°C, $\eta_{pl}$ [Pa·s] | Static critical pressure $p_{kr}$ [Pa] | Set testing pressure p [MPa] | Critical speed | |
|---|---|---|---|---|---|---|
| | | | | | [s⁻¹] | [m/s] |
| FLS 040.040 | 45.00 | 0.550 | 0.0776 | 0.038 | 73 | 11.4683 |
| FLS 250.020 | 25.00 | 1.120 | 0.0616 | 0.031 | 22 | 3.4562 |
| FLA 003.45 | 45.00 | 0.450 | 0.0868 | 0.043 | 82 | 12.8822 |
| FLA 002.25 | 25.00 | 0.725 | 0.0443 | 0.022 | 55 | 8.6405 |

in water environment conditions. The greatest value of critical speed equal to 12.8822 m/s was reached for FLA 003.45 magnetic fluid produced on the basis of the siloxane bearing liquid, intended for the using in acoustic devices. The smallest value of critical speed equal to 3.4562 m/s was obtained for FLS 250.020 magnetic fluid produced on the basis of the silicon bearing liquid, intended for the using in seals.

Comparison of the obtained critical speed values show that the both FLA fluids intended for the using in acoustic devices demonstrated, during the tests, greater values of the speed than the FLS fluids to be used in seals.

It should be mentioned that the applied relative pressure **$p/p_{kr}$ = 0.5** (defined as the ratio of the pressure set in the test and the static critical pressure**)** was kept equal in all the tests. Since the critical pressure values for particular fluids are different the tests were performed at different values of pressure in water chamber that also could influence the obtained results.

For instance, in the test with the use of FLA 002.25 fluid, conducted at the smallest value of pressure applied in the tests, much greater value of the critical speed was obtained than in the test with FLS 250.020 fluid, in spite of the much greater static critical pressure for this fluid.

In Fig. 9, are presented the life diagrams for the seals with the gap of 0.1mm in height, based on the tests with the use of FLS 250.020, FLS 040.040, FLA 002.2 and FLA 003.45 magnetic fluids at the relative pressure **$p/p_{kr}$ = 0.5**.

Comparison of the number of seal revolutions done till loss of tightness at different motion speeds shows that within the scope of the performed tests greater life values were obtained for the seals with applied FLA 003.45 and FLA 002.25 magnetic fluids than for those with FLS 250.020 and FLS 040.040 ones.

The results, like in the case of critical speed measurements, show how important is a kind of bearing liquid for life of magnetic-fluid seal operating in water. The smallest life was obtained from the tests with the use of FLS 250.020 magnetic fluid. The result, like in the case of critical speed measurements, may be associated with the relatively high pressure set in the tests with the use of the fluid in question.

Run of the life diagrams in function of seal motion speed shows a common characteristic feature, namely: in the range of higher motion speed values the speed decreasing only slightly influences the life increasing, however in the range of lower motion speed values a relatively small decrease of the speed results in a significant increase of seal life. The feature is clearly observed in the tests with the use of FLA 003.45 and FLS 040.040 magnetic fluids characterized by higher values of saturation magnetization. For instance, in the tests with FLA 003.45 fluid, after decreasing the seal rotational speed value from 82 s⁻¹ to 40 s⁻¹, i.e. by 50%, the number of revolutions to leakage increased by $8,151 \times 10^4$ cycles, and the subsequent decreasing the seal rotational speed value from 40 s⁻¹ to 30 s⁻¹, i.e. by 10%, resulted in the life increase by $1.718 \times 10^6$ cycles only.
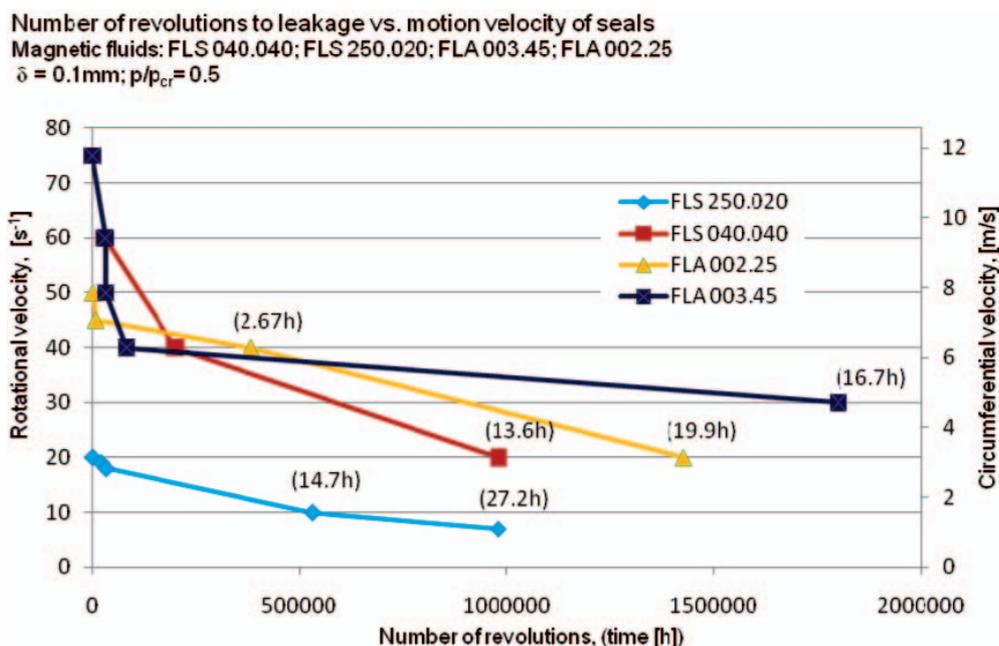


**Fig. 9.** *Comparison of the seal life values for various magnetic fluids used in the tests at the relative pressure of 0.5 and the gap height of 0.1mm*

## SUMMARY AND CONCLUSIONS

The performed tests of the single-stage magnetic-fluid seals operating in water environment confirmed correctness of the assumptions and procedure of the tests realized by means of the MAST L1-V2 test stand.

On the basis of the obtained results and observations, despite the limited range of the applied geometric and operational parameters, the following final conclusions and recommendations can be suggested:

1. The magnetic fluids made with the use of silicon liquids as a carrier liquid may be effectively applied to seals intended to operate in water, but in a limited range of seal motion speeds.

2. The tests in question should be supplemented by similar tests in sea water conditions aimed at checking chemical resistance of the magnetic fluids.

3. Life of magnetic-fluid seals operating in water greatly depends on seal motion speed. In the small motion velocities range the considerable growth of the lifetime with decrease of the motion velocity is observed. In the range of greater motion velocities the seal lifetime is very low.

4. The tests demonstrated that the pressure applied to magnetic –fluid seal, despite relatively low ratio of operational pressure/critical pressure was of significant effect on the obtained results. An additional research on influence of the pressure on life of magnetic-fluid seal operating in sea water of different salinity should be performed.

5. The MAST L1-V2 test stand for the testing of magnetic-fluid seals ensures correct run of the test. However its construction should be improved as to the detection of leakage because in its present state to precisely determine the instant of leakage is difficult.

## BIBLIOGRAPHY

1. Szydło, Z., Zachara, B., Ochoński, W.: *Ferro-magnetic fluids and their application in machine building* (in Polish). 1st Conference on automation of machines, devices and processes, Krynica, 1999

2. Rosensweig, R.E.: *Ferrohydrodynamics*. Dover Publications, Inc. Mineola, New York, 1997

3. Rinaldi, C., Chaves,A , Elborai, S.,He, X., Zahn, M.: *Magnetic fluid rheology and flows*. Current Opinion in Colloid & Interface Science Vol.10, Issue 3÷4, October, 2005

4. Orlov, D. B., at al..: *Eksperimentalnoe issledovanie resursa magnitozidkostnych uplotnenij pri germetazicii zidkich sred* (in Russian). Magnitnaya Gidrodinamika, No 4, 1989.

5. Kurfess,J., Muller, H.K.: *Sealing liquids with magnetic fluids*. Journal of Magnetism and Magnetic Materials,85,(1990)

6. Vihersalo, J., at al.: *Sealing of Liquids with Magnetic Fluid Seals*. Proc. of 6th Nordic Symposium on Tribology, NORDTRIB'94, Uppsala, Sweden, 12÷15 June 1994.

7. Chorney, Alvan F.; Mraz, Will.: *Hermetic Sealing with Magnetic Fluids.* Machine Design; ProQuest Science Journals May 64, 9, 7, 1992

8. Mitamura, Y., Arioka, S., Sakota, K., Azegami, M.: *Application of a magnetic fluid seal to rotary blood pumps*. J. Phys.: Condens. Matter 20 (2008) 204145

9. Szydło Z., Ochoński W., Zachara B. (inventors): *Measurement head for the testing of parameters of ferro-magnetic fluid seals* (in Polish). Stanisław Staszic Mining and Metallurgy Academy (AGH), Cracow, Patent application specification No. PL 377598 A1, dated 2005-10-12; published on 2007-04-16. Bulletin of Polish State Patent Office No. 8, 2007

10. Szydło, Z., Matuszewski, L.: *Ring thruster – a preliminary optimisation study of ferro-fluid seal and propeller*. Polish Maritime Research, Special issue No. S1, 2007

11. Szydło, Z., Matuszewski, L.: *The application of magnetic fluids in sealing nodes designed for operation in difficult conditions and in machines used in sea environment*. Polish Maritime Research, No. 3, Vol. 15, 2008

12. Szydło, Z.: *Tests of a seal with applied magnetic fluids (water – air) –preliminary determination of its life and permissible linear speed* (in Polish). Research project No. 6.6.130.261, AGH, 2008, (unpublished)

13. Bonvouloir, J.: *Experimental Study of High Speed Sealing Capability of Single stage Ferro fluidic Seal.* ASME Transactions, Journal of Tribology, Vol.119, July, 1997

14. Ferrolabs, Inc., Sterling, VA, United States: http://www.ferrolabs.com/ 2010

15. "ENES", Zientek P., Warsaw: www.magnesy.pl/ 2010

---

## CONTACT WITH THE AUTHORS

Leszek Matuszewski, Ph. D.
Faculty of Ocean Engineering
and Ship Technology
Gdansk University of Technology
Narutowicza 11/12
80-233 Gdansk, POLAND
e-mail: leszekma@pg.gda.pl

Zbigniew Szydło, Ph.D.
AGH University of Science and Technology
Faculty of Mechanical Engineering and Robotics
Al. Mickiewicza 30
31-069 Kraków, POLAND
e-mail: zbszydlo@agh.edu.pl

The Ship Handling Research and Training Centre at Ilawa is owned by the Foundation for Safety of Navigation and Environment Protection, which is a joint venture between the Gdynia Maritime University, the Gdansk University of Technology and the City of Ilawa.

### Two main fields of activity of the Foundation are:

⮑ Training on ship handling. Since 1980 more than 2500 ship masters and pilots from 35 countries were trained at Iława Centre. The Foundation for Safety of Navigation and Environment Protection, being non-profit organisation is reinvesting all spare funds in new facilities and each year to the existing facilities new models and new training areas were added. Existing training models each year are also modernised, that's why at present the Centre represents a modern facility perfectly capable to perform training on ship handling of shipmasters, pilots and tug masters.

⮑ Research on ship's manoeuvrability. Many experimental and theoretical research programmes covering different problems of manoeuvrability (including human effect, harbour and waterway design) are successfully realised at the Centre.

The Foundation possesses ISO 9001 quality certificate.

### Why training on ship handling?

The safe handling of ships depends on many factors - on ship's manoeuvring characteristics, human factor (operator experience and skill, his behaviour in stressed situation, etc.), actual environmental conditions, and degree of water area restriction.

Results of analysis of CRG (collisions, rammings and groundings) casualties show that in one third of all the human error is involved, and the same amount of CRG casualties is attributed to the poor controllability of ships. Training on ship handling is largely recommended by IMO as one of the most effective method for improving the safety at sea. The goal of the above training is to gain theoretical and practical knowledge on ship handling in a wide number of different situations met in practice at sea.

For further information please contact:
**The Foundation for Safety of Navigation and Environment Protection**

Head office:
36, Chrzanowskiego Street
80-278 GDAŃSK, POLAND
tel./fax: +48 (0) 58 341 59 19

Ship Handling Centre:
14-200 IŁAWA-KAMIONKA, POLAND
tel./fax: +48 (0) 89 648 74 90
e-mail: office@ilawashiphandling.com.pl
e-mail: office@portilawa.com