

Todinov, M. (2017) 'Reducing Risk Through Inversion and Self-Strengthening', *International Journal of Risk and Contingency Management*, 6 (1), pp. 14-42.

DOI: <https://doi.org/10.4018/IJRCM.2017010102>

This document is the Version of Record.

License: <https://creativecommons.org/licenses/by-nc-nd/4.0>

Available from RADAR: <https://radar.brookes.ac.uk/radar/items/ae11a6f6-3744-46d8-9106-eba29ea30110/1/>

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners unless otherwise waved in a license stated or linked to above. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

Reducing Risk Through Inversion and Self-Strengthening

Michael Todinov, Department of Mechanical Engineering and Mathematical Sciences, Oxford Brookes University, UK

ABSTRACT

A number of new techniques for reliability improvement and risk reduction based on the inversion method, such as: 'inverting design variables,' 'inverting by maintaining an invariant,' 'inverting resulting in a reinforcing counter-force,' 'negating basic required functions' and 'moving backwards to general and specific contributing factors' have been introduced for the first time. By using detailed calculations, it has been demonstrated how the new technique 'repeated inversion maintaining an invariant' can be applied to reduce the risk of collision for multiple ships travelling at different times and with variable speeds. It has been demonstrated that for pressure vessels, an inversion of the geometric parameters by maintaining an invariant volume could result not only in an increased safety but also in a significantly reduced weight. The method of self-strengthening (self-reinforcement) has been introduced for the first time as a systematic method for improving reliability and reducing risk. The method of self-strengthening by capturing a proportional compensating factor and the method of self-strengthening by creating a positive feedback loop have been proposed for the first time as reliability improvement tools. Finally, classifications have been proposed of methods and techniques for risk reduction based on the methods of inversion and self-strengthening.

KEYWORDS

Inversion, Inverting, Reliability Improvement, Reverse Thinking, Risk Reduction, Self-Reinforcement, Self-Strengthening

1. INTRODUCTION

A systematic classification of generic methods for reducing technical risk is crucial to safe operation, engineering designs and software, yet this very important topic has been overlooked in the reliability and risk literature. Distilling approaches for reliability improvement and risk reduction and formulating new approaches, provides a valuable and much needed support for design-engineers.

The work on formulating principles and methods for improving the reliability of engineering systems was initiated in (Todinov, 2007) and continued in Todinov (2015).

The reverse strategy for reliability improvement and risk reduction, which is in the focus of this work, has always been an important technique in problem solving. It was already known to Pappus of Alexandria about 320 AD, who described it as 'beginning with the desired outcome and working backwards until something known is reached.' The famous algebraist Carl Gustav Jacobi, whose favourite maxim was 'invert, always invert,' used the reverse strategy with great success in solving many hard mathematical problems. An important example of the reverse strategy in problem solving is the recursion, widely used for solving difficult combinatorial problems (Wirt, 1976). A typical approach in recursion is reducing the initial problem of size n to two or more problems of size $n-1$,

which in turn are reduced to problems of size $n-2$ and so on, until problems with trivial solutions are reached. Starting from the trivial solutions and moving backwards helps to assemble the solution.

The struggle between the need of reducing the weight of components and systems and the need of high reliability is a constant source of technical and physical contradictions. The reverse strategy has been an important technique for resolving some of these contradictions. It is then no surprise that the reverse strategy features in texts related to creativity and inventive problem solving (Altshuller, 1984; 1996; 1999; Orloff, 2006). The reverse strategy is at the heart of many inventions and Altshuller's TRIZ system (Altshuller, 1984; 1996; 1999) identified a number of useful applications of this technique.

Reverse strategy is often used in planning projects, operations or processes where the success is critically dependent on whether people, equipment and resources are available when and where needed. This is particularly relevant to projects where critical deadlines need to be met. In order to reduce the risk of failure, the planning of the project/process is often started from the desired outcome and progressed backwards. Starting from the desired end result and planning backwards helps to determine what resources are necessary to achieve the goal and to ensure how to make the necessary resources available at the right place and at the right time.

Despite the progress made in using the reverse strategy for problem solving and planning, almost none of the known sources focuses on the application of the reverse strategy for reliability improvement and risk reduction. In addition, the structure needed for applying this technique for reliability improvement and risk reduction is missing.

The reverse strategy is a powerful method for improving the reliability of engineering designs and will be referred to as *the method of inversion*.

Inverse states based on compressive residual stresses introduced by shot peening and their role in improving the reliability of automotive suspension springs have been discussed in (Todinov, 1999) and Todinov (2000). Recently, Fu et al. (2015) presented a comprehensive review of methods for introducing compressive residual stresses by a cold expansion, to improve the fatigue life of connection holes in aircraft structures.

Reverse strategy featuring introduction of inverse states has already been presented in (Todinov 2015). In this paper, the method of inversion is developed further, by introducing the inverting technique, a technique based on changing the focus to the opposite and a technique based on moving backwards to contributing factors. The method of inversion has been applied to reduce the risk of collision between ships and a general classification of inversion techniques has been proposed.

Another risk-reduction method referred to as 'risk reduction by self-strengthening' has also been introduced.

A typical self-strengthening (self-reinforcing) behaviour is present in racing cars where wings are used for creating an aerodynamic downforce which enhances the road grip and allows the car to travel faster around corners. Altshuller (1999) describes an invention where the hydrodynamic force produced by the sprinkler nozzles has been used to reduce the weight and increase the span of the wings of an irrigation machine. This is a self-strengthening effect, where an internal force is used to generate a self-strengthening effect.

Another example of improving reliability by self-strengthening (self-reinforcement) is the deformation self-strengthening of Hadfield steels discussed in (Lindroos, 2015). Introducing an impact loading improves the resistance of the steel to abrasive wear. A similar self-strengthening effect caused by macrostructural changes is the case of increasing the resistance against crack propagation in cubic zirconia ceramic with dispersed tetragonal zirconia particles (Kelly and Francis-Rose, 2002). For an existing crack in the ceramic material, increasing the crack tip opening load intensifies the stress field around the tip of the crack. The intensified stress field triggers the transformation of the embedded tetragonal zirconia particles to the monoclinic zirconia which has a larger specific volume.

The volume expansion associated with the phase transformation exerts a crack-closing force on the tip of the crack, which effectively resists the crack propagation. These are examples of self-strengthening resulting from a change of the properties of materials caused by internal stresses.

Self-strengthening is often a result of using a negative feedback control. Negative feedback control samples the output of the system and information about the deviation of the output is used to control the input in such a way that extreme deviations of the output, from a specified level, are avoided. In this way, negative feedback loops work towards stabilising the system's output and equilibrium and reduce deviations in the output. The larger the deviation of the output from the specified level the larger is the input correction counteracting the deviation of the output. Negative feedback loops are used widely in controlled systems to guarantee stable temperature, pressure, speed, etc. The negative feedback loop can be found, for example, in the coupling of induction motor with a machine. An increase in the angular velocity of the rotor from the point of stable operation causes the driving torque to drop below the resisting torque which slows down the rotor and decreases its angular velocity. A decrease in the angular velocity of the rotor causes the driving torque from the induction motor to increase above the resisting torque. The result is an increase of the angular velocity of the rotor. Commonly, in negative feedback control systems, the output is measured and fed back to an error detector at the input. A controller is then correcting the parameters of the system /process, so that the deviations from the expected output become as close to zero as possible. The availability of low-cost electrical devices and sensors makes it possible to provide more flexibility and regulate mechanical systems to a finer degree compared to all-mechanical governors.

Self-locking devices, such as self-locking screws, self-locking grips, self-locking hooks, self-energizing breaks, etc., have been known for a long time and are continuously used in modern designs. Costache et al., (2016) for example, recently introduced self-locking grips for anchoring fibre-reinforced tendons. Compensating factors leading to self-balancing have also been exploited as a self-strengthening process. The 'Roly-Poly' toy which rights itself when pushed over is a well-known example. Recently, self-balancing inspired by the 'Roly-Poly' toy has been suggested by Zhang et al. (2013) for improving the side rolling stability of an e-bike.

Finally, there exist self-healing materials which show self-strengthening effect with time. For example, some ceramic materials have natural ability to autogenously repair cracks over a period of time. Some metals also have the ability to repair a high-temperature creep damage.

Often, mechanisms which are in fact only 'strengthening' are referred to as 'self-strengthening.' Such is for example the strengthening mechanism of a rope made of twisted strands described in (Mattheck, 1998). According to the author, self-strengthening occurs if the direction of twist creates tensile stresses which load the twisted strands in tension. A twist in the opposite direction fails the rope easily.

It should be pointed out that increasing reliability through self-strengthening (self-reinforcement) is present if increasing the levels of external or internal forces/states also increases the levels of third factors which strengthen the system's response. An important feature of self-strengthening which distinguishes it from mere 'strengthening' is that increasing the external/internal force magnitudes increases the self-strengthening effect. For a simple strengthening mechanism, such a magnification of the strengthening effect is absent. Self-locking hooks for example, will not open under load and *increasing the load makes them more difficult to open*. For strengthened hooks that can only be locked, increasing the loading force does not make them more difficult to open.

Despite that isolated applications of self-strengthening have been known for a long time, this paper introduces the self-strengthening (self-reinforcement) method for the first time as a systematic reliability improvement and risk-reduction method.

2. IMPROVING RELIABILITY AND REDUCING RISK BY INVERTING

2.1. Inverting the Relative Location of Objects and Features

There are numerous cases where inverting the relative location of objects and features improves reliability and reduces risk.

A simple example are the flanged wheels of the railroad trains. The flanged wheels are needed to prevent the train from sliding off the track. Initially, the railroad tracks were flanged instead of the wheels and thousands of miles of unsafe railroad tracks were manufactured, with thousands of miles of unnecessary flanges. This situation changed dramatically when inverting had been applied to the location of the flange. The location of the flange was changed from the railway track to the wheel. In this way, the basic function of keeping the train from sliding off the track was preserved and the railway safety improved significantly.

2.2. Inverting the Orientation of Objects and the Direction of Motion

A well-known example of inverting the orientation of objects comes from drilling vertical blind holes in components, by a robot, on a manufacturing line. This operation requires cleaning the blind holes from metal chips which stick to the bottom because of gravity. However, if the hole is drilled on the component positioned upside-down; the problem associated with cleaning the hole is eliminated because gravity now helps to clean the hole.

Often, making a moving object stationary and moving the stationary object in the opposite direction (inverting the direction of motion) results in a significantly improved performance and risk reduction. This idea underlies the Cosworth® sand casting process (Campbel, 2015) where the molten metal is never poured down into the sand mould as is the case in the classical sand casting process. The molten metal flows in the opposite direction (uphill) into the mould. This type of metal flow eliminates turbulence and reduces the risk of trapping oxides into the metal which would reduce significantly the fatigue strength of the cast component.

Inverting the direction of motion can also be done to reduce the number of moving parts and improve reliability

Moving parts exhibit more failures compared to stationary parts. This is usually due to the increased kinetic energy, wear, fatigue, vibration, heat generation and erosion associated with moving parts.

Reducing the number of moving parts is an efficient way of improving the reliability of a system. Inverting the motion can be used directly to reduce the number of moving parts, which results in improving the reliabilities of assemblies.

Suppose that the required function is guaranteed by the relative motion of two components with respect to each other. If an extra uniform motion with opposite direction to the motion of the first component is imparted to both components, the first component will become stationary and the motion of the second component will become a composition of its motion before the inversion and the motion of the first component taken with a negative sign. The relative motion of the two components will not change but the reliability will improve because the number of moving parts have been decreased by one.

2.3. Variables Inversion Technique for Minimising Objective Functions Whose First-Order Partial Derivatives Retain the Same Sign on a Rectangular Domain

An important new application of the method of inversion is presented here, in determining the global minimum of multivariable functions whose first-order partial derivatives maintain the same sign on a rectangular domain.

Suppose that for a differentiable multivariable function $f(x_1, x_2, \dots, x_n)$, the design variables x_1, x_2, \dots, x_n vary within a particular rectangular domain:

$$\begin{aligned}
 a_1 &\leq x_1 \leq b_1 \\
 a_2 &\leq x_2 \leq b_2 \\
 &\dots\dots \\
 a_n &\leq x_n \leq b_n
 \end{aligned} \tag{1}$$

and the first-order partial derivatives do not change sign in the specified domain (1).

In this case, the global minimum of the objective function $f(x_1, x_2, \dots, x_n)$ can be found at a location \mathbf{x}^* , obtained by an inversion of the design variables. The inversion of the design variables consists of the following. If the first-order partial derivative with respect to design variable x_i is positive in the rectangular domain, the design variable x_i takes the smallest value from its permitted variation interval (the left end a_i of the interval $(x_i^* = a_i)$). If the partial derivative with respect to design variable x_i is negative, the design variable x_i takes the largest value from its permitted variation interval (the right end b_i of the interval $(x_i^* = b_i)$). At the end of this process, a corner point \mathbf{x}^* with coordinates $x_1^* = a_1, x_2^* = b_2, x_3^* = a_3, x_4^* = a_4, \dots, x_n^* = b_n$ will be defined, which corresponds to the global minimum $y^* = f(x_1^*, x_2^*, \dots, x_n^*)$ of the function $f(x_1, x_2, \dots, x_n)$ in the entire rectangular domain (1).

Proof: Because the function $f(x_1, x_2, \dots, x_n)$ is defined on a rectangular domain, the total function increment associated with any possible location \mathbf{x}^P in the rectangular domain can always be obtained from the defined corner point \mathbf{x}^* by adding a finite number k of local increments Δf_i from k sequential changes Δx_i of each design variable (coordinate) while keeping the rest of the design variables (coordinates) constant. Hence, the value y^P of the function associated with location \mathbf{x}^P is given by:

$$y^P = y^* + \Delta f_1 + \Delta f_2 + \dots + \Delta f_k \tag{2}$$

where Δf_i is the increment resulting from the change Δx_i of design variable x_i while keeping the rest of the design variables constant ($i=1,2,\dots,k$) and y^* is the value of the function which corresponds to the corner location \mathbf{x}^* .

Because of the selection of the corner location \mathbf{x}^* in reaching the arbitrary location \mathbf{x}^P , there will be no design variable change whose corresponding increment Δf_i is negative. In other words, the sequential design variable changes (while keeping the rest of the variables constant) from location \mathbf{x}^* to any chosen location \mathbf{x}^P from the domain (1), result only in $\Delta f_i > 0$. As a result, the corner location \mathbf{x}^* is associated with the smallest value y^* of the function $f(x_1, x_2, \dots, x_n)$.

Indeed, suppose that a location exists where the function accepts a value y^P which is smaller than the value y^* of the function associated with the corner location \mathbf{x}^* ($y^P < y^*$). Because $y^P < y^*$, from Equation (2) it follows that:

$$\Delta f_1 + \Delta f_2 + \dots + \Delta f_k < 0 \tag{3}$$

must necessarily hold. Consequently, there must be at least one design variable change for which $\Delta f_i < 0$. This means that at least one of the additive increments must be negative ($\Delta f_i < 0$). This however, contradicts the fact that all permitted design variable changes from the initial corner point x^* result only in positive increments $\Delta f_i > 0$. This completes the proof.

Here is a simple example, illustrating this technique. Consider the objective function:

$$f(x, y) = -x^2 + 4x - y^3 - \sin y + \cos y - 4 \quad (4)$$

The first-order partial derivative $\frac{\partial f}{\partial x} = -2x + 4$ of this function is with a positive sign in the domain $0.1 \leq x \leq \pi/2$; $0.1 \leq y \leq \pi/2$. While the partial derivative $\frac{\partial f}{\partial y} = -3y^2 - \cos y - \sin y$ is with a negative sign in the same domain.

Consider the initial state $x^* = 0.1$ and $y^* = \pi/2$ obtained by an inversion of the design variables. Because $\frac{\partial f}{\partial x} > 0$, design variable x accepts the smallest value ($x^* = 0.1$) of its variation interval while design variable y accepts its largest value ($y^* = \pi/2$) because $\frac{\partial f}{\partial y} < 0$. From the initial location $x^* = 0.1$, $y^* = \pi/2$ any other location can be reached by increasing x by Δx and decreasing y by Δy . The signs of both function increments Δf_x and Δf_y , corresponding to the changes in x and y are positive, therefore at $x^* = 0.1$ and $y^* = \pi/2$ the function $f(x, y)$ has a global minimum.

The variables inversion technique provides the valuable opportunity to find quickly a global extremum of a function for which the first-order partial derivatives retain their sign on a rectangular domain. As a result, the stated technique provides an efficient strategy for guaranteeing a global extremum in many engineering problems related to optimising a performance characteristic which improves reliability and reduces risk. The proposed simple technique, similar to the pigeonhole principle in combinatorics (Brualdi, 2010), has deep and non-trivial applications.

In many engineering optimisation problems, the permitted intervals of variation of the design variables are relatively small, and the first-order partial derivatives retain their sign within their relatively small intervals of variation. In this case, the first-order partial derivatives do not change their signs and the variables inversion technique can be applied.

The variables inversion technique can also be applied to improve the performance of heuristic methods searching for a global extremum of a multivariable function $f(x_1, x_2, \dots, x_n)$ in a large domain. Many heuristic optimisation methods such as the particle swarm optimisation (Lazinica 2009, Kiranyaz et al., 2014) for example, incorporate a local search in the vicinity of a point, to find the local best value. Because within a relatively small vicinity of a point, the first-order partial derivatives do not normally change sign, the proposed technique can be used for determining quickly the extreme local value in the vicinity of a specified point. If the local search space is defined by n -dimensional rectangular domain:

$$a_1 \leq x_1 \leq b_1; a_2 \leq x_2 \leq b_2; \dots; a_n \leq x_n \leq b_n$$

determining the function values in each of the corner points of this domain involves 2^n evaluations of the function. This is a very expensive and time consuming step if n is relatively large. With the proposed method, the signs of the partial derivatives can be checked by obtaining the sign of the following finite differences:

$$f(b_1, a_2, a_3, \dots, a_n) - f(a_1, a_2, a_3, \dots, a_n), f(a_1, b_2, a_3, \dots, a_n) - f(a_1, a_2, \dots, a_n), \\ f(a_1, a_2, b_3, \dots, a_n) - f(a_1, a_2, a_3, \dots, a_n), \dots, f(a_1, a_2, a_3, \dots, b_n) - f(a_1, a_2, a_3, \dots, a_n)$$

which involves only $n+1$ evaluations of the function $f(x_1, x_2, x_3, \dots, x_n)$. Depending on the signs of these finite differences, the corner point corresponding to the smallest value of the function is determined very quickly. One more final evaluation of the function is finally needed to determine the value of the extremum. This results in only $n+2$ evaluations of the function $f(x_1, x_2, x_3, \dots, x_n)$, overall. As a result, the time complexity of this critical step is reduced from exponential $O(2^n)$ to linear $O(n)$, which results in a significantly improved performance of the optimisation algorithm.

2.4. Inverting by Maintaining an Invariant

This new technique is introduced through several basic applications: (1) reducing the risk of catastrophic failure of pressure vessels (2) improving the reliability of screw joints and (3) evaluating the risk of collision by a repeated inversion of the direction of velocities.

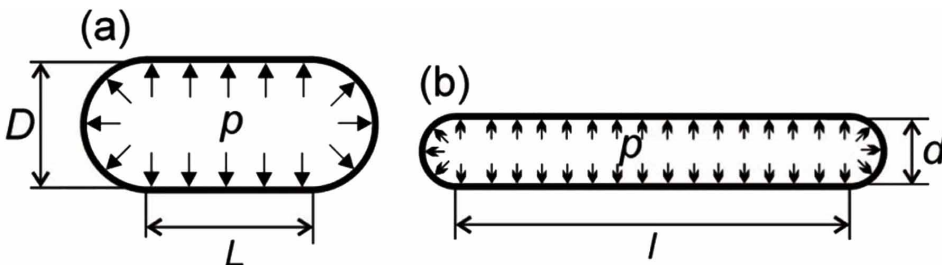
An invariant is a quantity that does not change despite that some parameters of the system do change. Well-known examples are: (1) the total momentum of an isolated system; (2) the total energy of an isolated system; (3) the relative distances between several bodies with respect to a uniform motion imparted to each of the bodies, etc.

2.4.1. Inverting Geometrical Parameters

Consider the pressure vessel in Figure 1 with a specified volume V , diameter D , length L and thickness t of the shell. The vessel contains fluid exerting pressure p on the inside of the shell (Figure 1a). The shape of the pressure vessel can be modified by inverting the dimensions D and L in Figure 1a while maintaining the same volume. The relatively large diameter D is reduced to a diameter d (Figure 1b) and the relatively small length L is increased to a length l (Figure 1b). During this inversion, the volume is kept constant, equal to the required volume of the pressure vessel:

$$V = V_1 = \pi D^2 (D/6 + L/4) = V_2 = \pi d^2 (d/6 + l/4)$$

Figure 1. Inverting the dimensions by preserving an invariant volume significantly reduces the hoop stress



The largest tensile stress acting in the pressure vessel from Figure 1a is the hoop stress σ_{H1} , determined from:

$$\sigma_{H1} = \frac{pD}{2t} \quad (5)$$

The largest tensile stress acting in the wall of the modified pressure vessel from Figure 1b is the hoop stress σ_{H2} , determined from:

$$\sigma_{H2} = \frac{pd}{2t} \quad (6)$$

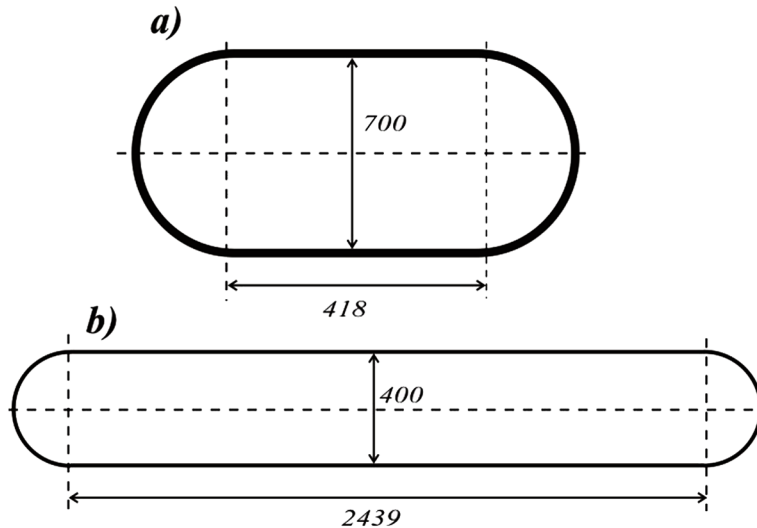
Despite that the volume of the pressure vessel remained unchanged, the inversion of the dimensions reduced significantly the hoop stress, which is the largest principal tensile stress acting in the shell of the pressure vessel. The axial principal tensile stress is also reduced.

The reduced hoop stress in the second design, makes it even possible to reduce the wall thickness t which permits obtaining not only a safer design but also a design with a significantly reduced weight.

Indeed, suppose that an alloy with fracture toughness of $K_{Ic} = 112 \text{ MPa}\sqrt{\text{m}}$ and yield stress $\sigma_s = 1100 \text{ MPa}$ has been used for the design of a cylindrical pressure vessel with semi-spherical ends, containing gas with pressure 20 MPa . The dimensions of the pressure vessel (in mm) are according to Figure 2a. The radius of the spherical cap is 350mm and the length L is 418mm . The volume is $V = 0.34 \text{ m}^3$. The thickness of the wall for the initial design A is $t_1 = 12 \text{ mm}$ but can be reduced down to $t_2 = 5 \text{ mm}$.

By applying inversion of the geometric parameters preserving the volume, the thickness of the shell is reduced to 5 mm , the diameter of the pressure vessel is reduced to 400mm and the length of

Figure 2. Modifying the shape of the pressure vessel by inverting the geometric parameters, while keeping the same volume, improves reliability and reduces the weight of the design



the vessel is increased to 2439mm. As a result, a modified design is obtained with the same volume $V = 0.34 m^3$ (Figure 2b).

The hoop stress in the shell of the pressure vessel is given by $\sigma = \frac{pd}{2t}$, where p is the internal pressure and d is the internal diameter of the vessel. For a through crack oriented in the worst possible direction (perpendicular to the hoop stress), the stress intensity factor is: $K_I = \sigma \sqrt{\pi a} = \frac{pd}{2t} \sqrt{\pi a}$. In order for the *leak before break condition to exist*, the material must be able to support a through crack with length $2a=2t$. As a result, the leak before break condition becomes:

$$\frac{pd}{2t} \sqrt{\pi t} < K_{Ic}$$

where K_{Ic} is the fracture toughness of the material.

The hoop stress characterising the design in Figure 2a is:

$$\frac{pd_1}{2t_1} = \frac{20 \times 10^6 \times 0.7}{2 \times 0.012} = 583.3 MPa$$

The hoop stress characterising the design from Figure 2b is:

$$\frac{pd_2}{2t_2} = \frac{20 \times 10^6 \times 0.4}{2 \times 0.005} = 800 MPa$$

and both hoop stresses are smaller than the yield stress of the material. However, the condition ‘leak before break’ for the first design:

$$\frac{p_1 d_1}{2t_1} \sqrt{\pi t_1} = 583.3 \times 10^6 \sqrt{\pi \times 0.012} = 113.26 MPa\sqrt{m} > K_{Ic} = 112 MPa\sqrt{m}$$

is not fulfilled. The design from Figure 2a is not safe.

The condition ‘leak before break’ for the second design:

$$\frac{p_2 d_2}{2t_2} \sqrt{\pi t_2} = 800 \times 10^6 \sqrt{\pi \times 0.005} = 100.27 MPa\sqrt{m} < K_{Ic} = 112 MPa\sqrt{m}$$

is fulfilled and the design is safe. The volume of material required by the design in Figure 2a is:

$$V_1 = 2\pi (d_1/2) \times L_1 \times t_1 + 4\pi (d_1/2)^2 t_1 = (2\pi \times 0.35 \times 0.418 + 4\pi \times 0.35^2) \times 0.012 \approx 0.0295 m^3$$

while the volume of material required by the design in Figure 2b is:

$$V_2 = 2\pi (d_2/2) \times L_2 \times t_2 + 4\pi (d_2/2)^2 t_2 = (2\pi \times 0.2 \times 2.439 + 4\pi \times 0.2^2) \times 0.005 \approx 0.0178 m^3$$

As a result of the inversion of the geometric parameters preserving an invariant volume, the safety of the design in Figure 2a has been significantly improved and simultaneously, the weight of the design has been reduced by 40%.

2.4.2. Inverting Mechanical Properties

For a bolt and flange joint where the load on the bolt fluctuates, the force-deformation diagram for a stiff bolt and elastic flange is shown in Figure 3a. The elastic constants of the bolt and the flange are $k_b = \tan(\alpha)$ and $k_f = \tan(\beta)$, correspondingly. The bolted joint is subjected to a required preload (clamping) force of magnitude P . Because of a variation of the load on the bolt, the bolted joint is subjected to a pulsating external force with magnitude F , which causes a pulsating force of magnitude F_b in the bolt (Figure 3a).

The range of the pulsating force in the bolt can be reduced by making an inversion of the elastic constant of the bolt and the flange whilst keeping the required clamping force P constant. The relatively large elastic constant of the bolt can be reduced and the relatively small elastic constant of the flange can be increased (Figure 3b). The net effect is a smaller range of the pulsating force in the bolt (Figure 3b) which results in a larger fatigue life (reliability).

This effect can be quantified through the amplitude of the pulsating force F_b , by expressing the length L_{AB} of the segment AB (Figure 3a), in two different ways (Figure 3a):

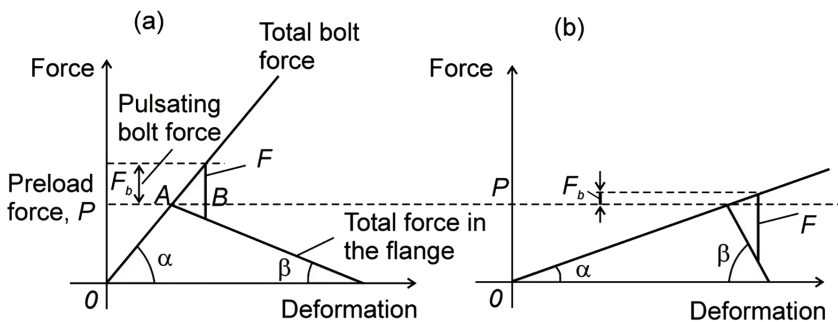
$$L_{AB} = \frac{F_b}{\tan(\alpha)} = \frac{F - F_b}{\tan(\beta)} \quad (7)$$

Considering that $k_b = \tan(\alpha)$ and $k_f = \tan(\beta)$, from (7) we have:

$$F_b = \frac{k_b}{k_b + k_f} F \quad (8)$$

Equation (8) shows that inverting the elastic properties of the bolt and the flange (reducing the elastic constant k_b of the bolt and increasing the elastic constant of the flange) reduces the magnitude

Figure 3. Range F_b of the pulsating force acting in the bolt of a preloaded bolted joint a) a stiff bolt and elastic flange; b) elastic bolt and a stiff flange



of the pulsating force F_b in the bolt and increases its fatigue life. The elastic constant of the bolt is given by $k_b = EA / L$, where L is the length of the bolt, A is the cross sectional area and E is the Young modulus of the material. Reducing the elastic constant of the bolt can be done by selecting a bolt with a larger length L or by reducing the cross-sectional area A of the bolt by thinning down the shank diameter of the bolt.

2.4.3. Repeated Inversion of the Direction of Velocities

The technique of inversion by maintaining an invariant property can be used with success for reducing the risk of collision of ships, and fishing boats. The as-described risk of collision is an important problem already discussed widely in the relevant literature (Oh et al, 2015; Lebkowski et al, 2005). This problem is particularly acute for a large volume of sea traffic, high number of shipping lanes, and increased speeds and dimensions of newly-built ships (Lebkowski et al, 2005). Consequently, the ability to make fast decisions about manoeuvres of ships is a safety-critical activity.

The risk of a dangerous approach to a travelling ship can be assessed through the value r_{sd} where r_{sd} is the radius the safety domain surrounding the ship. This radius makes a safety zone for the moving ship which cannot be entered by any other navigational object without a collision threat. A comprehensive discussion related to the safety domain and its definition has been provided in (Lebkowski et al, 2005).

The inversion method applied for assessing the risk of collision uses the fact that the relative distances between a number of navigating vessels (Figure 4a) does not change if the same uniform velocity is imparted to each vessel, in a particular common direction. In other words, the relative distances between the ships are invariant with respect to a common uniform velocity imparted to all ships.

If the common velocity is selected to be equal in magnitude and opposite to the velocity of one of the ships (e.g. the first ship), this ship will become stationary ($\mathbf{v}'_1 = \mathbf{v}_1 - \mathbf{v}_1 = 0$) and the rest of the ships will move with altered velocities equal to the vector sum of their old velocity and the inverted velocity of the first ship. In Figure 4b, the altered velocities of the ships become: $\mathbf{v}'_2 = \mathbf{v}_2 - \mathbf{v}_1$, $\mathbf{v}'_3 = \mathbf{v}_3 - \mathbf{v}_1$ and $\mathbf{v}'_4 = \mathbf{v}_4 - \mathbf{v}_1$. The relative distances between the ships however, will not change.

Essentially, the problem is more complex because the ships actually start and end their journeys at different times. This time dependence is presented in Figure 5 for two ships which start their travel at times ts_1 and ts_2 , and end it at times te_1 , te_2 , correspondingly. It is assumed that if a ship is

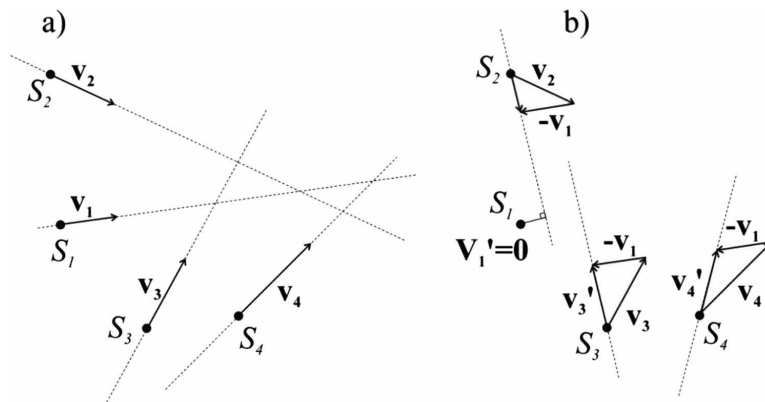


Figure 4. Determining the closest distance among ships travelling with constant speed, by using a repeated inversion of velocities

located safely in port, no collision with other ships could possibly occur. Consequently, a collision between two navigating vessels can only occur during the time interval of parallel travel (in Figure 5, during the time interval ts_2, te_1). This observation significantly simplifies the problem, because the time intervals during which one of the ships is in port are excluded from consideration.

Another complication is that during their travel the ships traverse not infinite straight lines but finite segments. Determining the smallest distances between the first ship S_1 and the second ship S_2 (Figure 6a) during their parallel time interval of navigation can be done by applying the method of inversion.

A uniform velocity $-v_1$ is imparted to the first ship (Figure 6b). As a result, the first ship becomes stationary and the second ship S_2 starts moving with a new velocity $v'_2 = v_2 - v_1$ along a modified segment S_2E_2 . The modified segment S_2E_2 travelled by the second ship, with the altered velocity, is oriented along the new trajectory defined by the vector $v'_2 = v_2 - v_1$ and the length of the travelled modified segment is equal to $|v_2 - v_1| t_{p12}$, where $|v_2 - v_1|$ is the magnitude of the modified velocity and t_{p12} is the time of parallel travel for ships S_1 and S_2 (Figure 6b). Next, the smallest distance between the first and second ship is determined. It is the shortest distance between the point S_1 and the modified segment S_2E_2 (Figure 6b).

If both angles $\alpha \equiv S_1S_2E_2$ and $\beta \equiv S_1E_2S_2$ are acute (Figure 6b), the shortest distance is the length of the perpendicular to S_2E_2 (Figure 7).

Denote by (x_1, y_1) the coordinates of the ship S_1 , by (x_2, y_2) the coordinates of the ship S_2 , by (vx_1, vy_1) the components of the velocity v_1 and by (vx_2, vy_2) the components of the velocity v_2 .

Figure 5. Collision between two navigating vessels can only occur during the time interval of a parallel travel

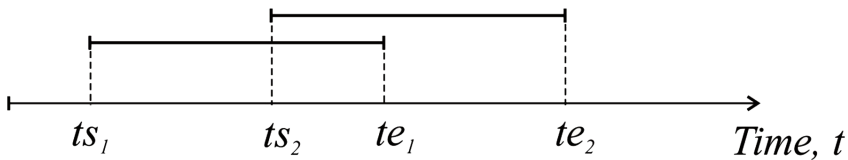


Figure 6. Determining the closest distance among ships travelling with constant speed, by using a repeated inversion of velocities

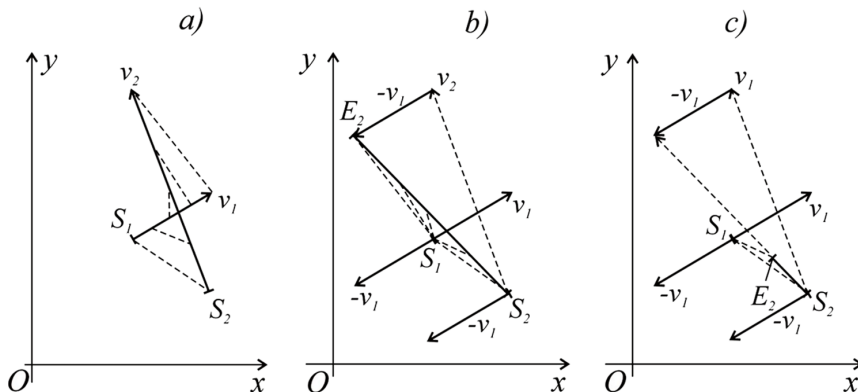
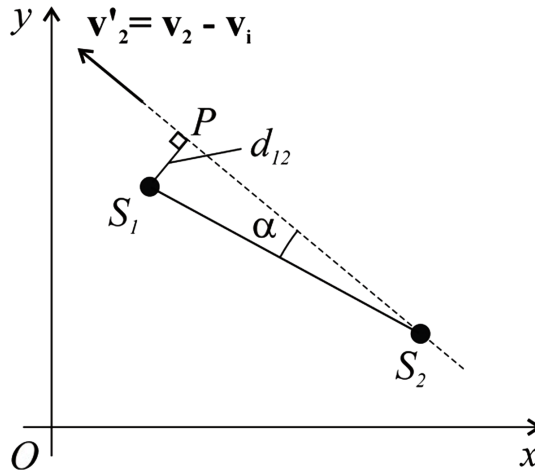


Figure 7. Determining the shortest distance to the modified travelled segment when both angles α and β are acute



Because the components of the modified velocity vector $\mathbf{v}' = \mathbf{v}_2 - \mathbf{v}_1$ of the second ship and the components of the vector $\mathbf{S}_2\mathbf{S}_1$ are known, the angle α (Figure 7) is determined from:

$$\cos \alpha = \frac{\mathbf{S}_2\mathbf{S}_1 \cdot \mathbf{v}'_2}{\|\mathbf{S}_2\mathbf{S}_1\| \cdot \|\mathbf{v}'_2\|} \quad (9)$$

where $\mathbf{S}_2\mathbf{S}_1 \cdot \mathbf{v}'_2 = (x_1 - x_2)(vx_2 - vx_1) + (y_1 - y_2)(vy_2 - vy_1)$ is the scalar product of the vectors $\mathbf{S}_1\mathbf{S}_2$ and \mathbf{v}'_2 , $\|\mathbf{S}_1\mathbf{S}_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ is the length of the distance $\mathbf{S}_1\mathbf{S}_2$ and $\|\mathbf{v}'_2\| = \sqrt{(vx_2 - vx_1)^2 + (vy_2 - vy_1)^2}$ is the magnitude of the altered velocity \mathbf{v}'_2 .

In a similar fashion, the cosine of angle $\beta \equiv \mathbf{S}_1\mathbf{E}_2\mathbf{S}_2$ ($\cos \beta$) is determined. If both $0 \leq \cos \alpha \leq 1$ and $0 \leq \cos \beta \leq 1$, the shortest distance is the perpendicular S_1P to the modified segment S_2E_2 .

From expression (9), $\sin \alpha = \sqrt{1 - \cos^2 \alpha}$ can be determined and the length of the perpendicular S_1P becomes:

$$S_1P \equiv d_{12} = \|\mathbf{S}_2\mathbf{S}_1\| \sin \alpha = \frac{1}{\|\mathbf{v}'_2\|} \sqrt{\|\mathbf{S}_2\mathbf{S}_1\|^2 \cdot \|\mathbf{v}'_2\|^2 - (\mathbf{S}_2\mathbf{S}_1 \cdot \mathbf{v}'_2)^2} \quad (10)$$

If $0 \leq \cos \alpha \leq 1$ is not fulfilled, the angle α is obtuse and the shortest distance is the length S_2S_1 ; If $0 \leq \cos \beta \leq 1$ is not fulfilled, the angle β is obtuse and the shortest distance is the length S_1E_2 (Figure 6b).

Once the smallest distances between the first ship and the remaining ships have been determined, the same procedure can be applied to the second ship and the remaining ships. To each of the remaining

ships (except the first ship whose distances have been determined), a uniform velocity of $-\mathbf{v}_2$ is imparted. As a result, the velocity of the second ship becomes zero ($\mathbf{v}'_2 = \mathbf{v}_2 - \mathbf{v}_2 = 0$) and the velocities of the rest of the ships become $\mathbf{v}'_3 = \mathbf{v}_3 - \mathbf{v}_2$, $\mathbf{v}'_4 = \mathbf{v}_4 - \mathbf{v}_2$ and so on. Now, the smallest distances between the second ship and the rest of the ships can be determined. This procedure is extended to the rest of the ships.

Finally, comparing the as-determined individual shortest distances d_{ij} , $i < j$, between the ships, reveals the absolute shortest distance between a pair of ships. The smallest absolute distance $d_{\min} = \min \{d_{ij}\}$ is determined and the possibility of collision is checked by comparing it with the radius r_{sd} of the safety zone. On the basis of the described method, a simple algorithm can be developed for determining the smallest distance among moving ships (see Algorithm 1).

Initially, the smallest distance d_{\min} between the travelling ships is initialised with a very big number. As each smallest distance d between a pair of ships is calculated sequentially, it is compared with the current minimum distance d_{\min} . As a result, at the end of the calculations, the absolute

Algorithm 1

```
dmin= Big_number; ind1=0; ind2=0;

for i=1 to n-1 do
  for j=i+1 to n do
    {

      Determine the time interval of parallel travel
      between ships i and j;
      if (there is an interval of parallel travel)
        {Stop ship i by imparting a velocity with the same
          magnitude and opposite in direction to its current
          velocity.

          Determine the modified velocity of the jth ship by subtracting
          from its velocity  $v_j$ , the velocity  $v_i$  of the first ship;

          Determine the length of the modified segment  $S_jE_j$  travelled by the jth ship;

          Evaluate the angles alfa and beta;

          if (alfa is an obtuse angle) d =  $S_iS_j$ ;
          else if (beta is an obtuse angle) d= $S_iE_i$ ;
          else {
            Determine the length p= $S_iP$  of the perpendicular  $S_iP$ ;
            d=p;
          }

          if (d<dmin) {dmin=d; ind1=i; ind2=j;}
        }
    }
  }
```

minimum distance remains in the variable d_{min} and the indices of the ships closest to each other remain in the variables 'ind1' and 'ind2.' The smallest distance d_{min} is determined after $n(n-1)/2$ comparisons. The technique *inverting by maintaining an invariant* yielded a simple algorithm for assessing the risk of collision among ships.

The problem can be generalized in the case where the ships change their velocity magnitudes and directions with time $\mathbf{v}_i = \mathbf{v}_i(t)$ (Figure 8a). In this case, when considering the shortest distance between two ships, for example between ships S_1 and S_2 , the updating of the velocity of the second ship must be done regularly at the times t_{11}, t_{12}, \dots , at which the first ship changes the magnitude or the direction of its velocity and also at the times t_{21}, t_{22}, \dots at which the second ship changes the magnitude or the direction of its velocity. During the updates, the updating velocity for the second ship is always equal and directly opposite to the velocity of the first ship. In this case, the modified travelled path of the second ship will no longer be a single straight-line segment but a series of connected line segments (see Figure 8b).

The first ship in Figure 8a changed its velocity magnitude and direction at time t_{11} while the second ship changed its velocity magnitude and direction at time t_{21} . The modified path of the second ship is composed of the connected line segments S_2M_1 , M_1M_2 and M_2E_2 . The shortest distance between the ships is determined by determining the shortest distance between point S_1 , which corresponds to the position of the first ship (which never changes during the repeated inversion of velocity) and the series of connected line segments $S_2M_1M_2E_2$. The shortest distance is obtained by determining sequentially the shortest distances to the individual segments and selecting the smallest distance.

3. IMPROVING RELIABILITY AND REDUCING RISK BY MOVING BACKWARDS AND CHANGING THE FOCUS TO THE OPPOSITE

The classification of the techniques for improving reliability and reducing risk based on the method of inversion, proposed in this paper, is given in Figure 9. This classification introduces a structure in implementing the method of inversion and simplifies the process of searching for an appropriate solution. In addition, the structure provided by the classification makes the application of the method of inversion less dependent on inspiration and lateral thinking leaps compared to a design not supported by such classification. The classification is easy to use and does not require a significant investment

Figure 8. Determining the smallest distance in the case of ships travelling with variable velocities

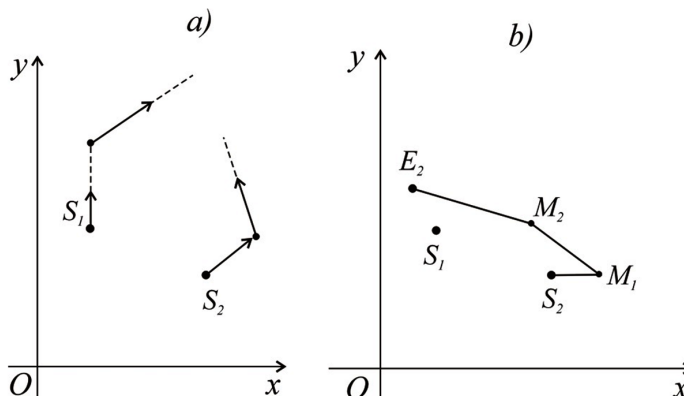
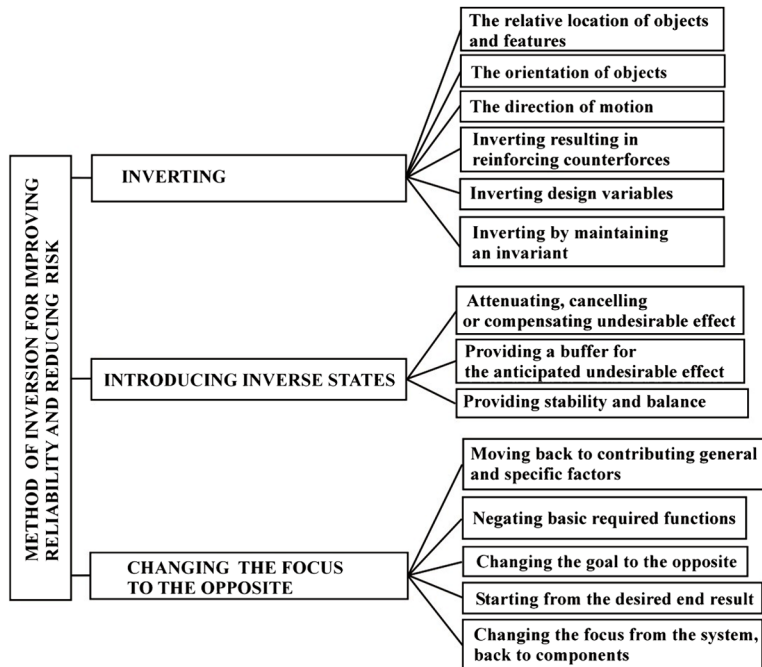


Figure 9. The method of inversion for improving reliability and reducing risk



of time. In addition, the classification acts as a useful checklist which helps avoid the omission of a relevant technique based on the method of inversion.

One of the simplest yet very efficient group of techniques based on the method of inversion is ‘moving backwards to the contributing factors’ and ‘changing the focus to the opposite.’

3.1. Moving Backwards to the Contributing Factors

Starting from an end state and moving backwards towards the general and specific contributing factors is a powerful technique for reliability improvement and risk reduction (Figure 10). Moving backwards to the contributing factors creates conceptual anchors for the thinking and generates a full range of alternative reliability improvement measures contributing to the isolated general factor. In this way, a full set of reliability improvement measures are generated and the likelihood of missing a relevant reliability improvement measure is reduced significantly. This process has been illustrated in Figure 10.

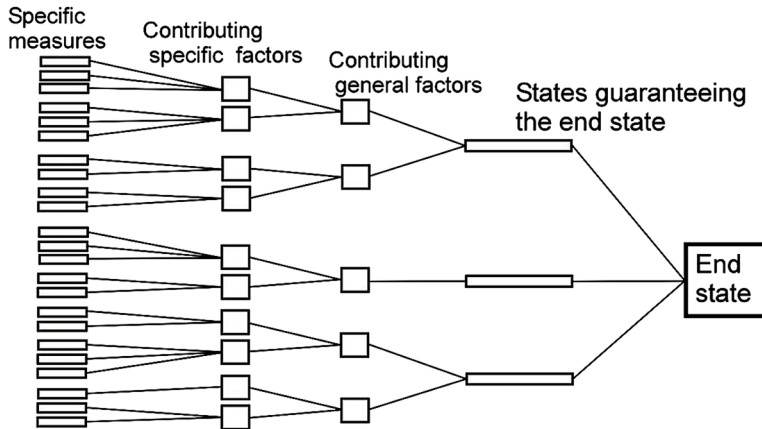
In the common case of a loaded component, high reliability is brought by the state ‘low probability that the load will exceed strength.’ The general factors contributing to this state are (1) ‘decreased load’ (2) ‘increased strength.’

Contributing specific reliability improvement measures to the general factor ‘decreased load’ are: decreasing the mean value of the load by derating, decreasing the variability of the load by smooth loading and eliminating the high values of the load by stress limiters.

Contributing specific reliability improvement measures to the general factor ‘increased strength’ are increasing the mean value of the strength by selecting appropriate materials, increasing the mean value of the strength by increasing the load carrying sections, decreasing the strength variability by better technological processes, manufacturing control and inspection control, eliminating the lower tail of the strength distribution by a burn-in and environmental testing.

Often it is beneficial to change the goal from reliability improvement to unreliability and failure. In this case, the failure modes at a system level are provided by negating basic required system

Figure 10. Shifting the focus by starting from the end state and moving backwards to the contributing factors and specific measures



functions. Moving back to the general contributing factors and the specific failure modes leading to these factors is a powerful technique for discovering and eliminating failure modes.

Here is an example featuring a lubrication system. The purpose of a lubrication system is: (1) reducing wear and resisting friction forces, (2) removing heat from friction zones and cooling the contact surfaces, (3) cleaning the contact surfaces from abrasion particles and dirt and (4) protecting the lubricated parts from corrosion.

The required functions from a generic lubrication system is to supply constantly clean oil at a specified pressure, temperature, debit, composition and viscosity to contacting and moving parts. Not fulfilling any of the required system functions constitutes a system failure. Negating 'clean oil' results in 'unclean oil.' Moving back to the contributing general factors to unclean oil reveals the general contributing factors 'oil degradation' and 'oil contamination.'

Specific factors leading to oil degradation for example are 'depletion of the oil additives' and 'oxidation of the oil,' 'inappropriate oil selected,' 'inappropriate viscosity of the selected oil.' Specific factors leading to the general contributing factor 'oil contamination' are for example 'oil with suspended particles due to a damaged filter,' 'oil with moisture' etc.

Negating 'specified pressure' results in 'too high pressure of the supplied oil' and 'too low pressure of the supplied oil.' Moving back to the state 'too low pressure of the supplied oil' reveals the contributing general factors 'due to the oil used,' 'due to component failures,' 'due to excessive wear,' 'due to faulty design.'

Specific contributing factors leading to the general factor 'too low pressure due to the oil used' are 'oil with too large viscosity' and 'too small quantity of circulating oil in the lubricating system.' Specific contributing factors leading to the general factor 'too low oil pressure' are 'failure of the oil pump,' 'blockage of the oil filter,' 'blockage of the cooler,' 'failure of the 'pressure relief valve.'

A specific factor leading to the general contributing factor 'too low oil pressure due to excessive wear' is 'excessive clearances in the journal bearings due to wear' while a specific factor leading to the general factor too low oil pressure due to faulty design' is 'inappropriately designed oil galleries.'

Negating 'specified temperature' reveals the state 'too high temperature' of the supplied oil. Moving backwards reveals the contributing general factors 'too high temperature due to inappropriate cooling' and 'too high temperature due to the oil used.'

Specific factors leading to the contributing general factor 'too high temperature due to inappropriate cooling' are 'failure of the cooling circuit' and 'cooler clogged with debris or lined with plaques.' Specific factors leading to the contributing general factor too high temperature due to the oil used are 'insufficient quantity of the circulating oil,' and 'inappropriate oil selection.'

If factors are contributing to a more general factor, it is necessary to move back to a lower level in order to reveal the more specific contributing factors. Thus, moving back from the factor ‘insufficient quantity of the circulating oil’, reveals the more specific contributing factors ‘incorrectly sized sump’ and ‘too low oil level.’

3.2. Changing the Focus to the Opposite

Instead of looking at how to improve performance, the focus of this technique is shifted on inducing bad performance. Instead of focussing on how to produce a reliable design, the focus is on common mistakes made in design which promote early-life failures. The benefit from changing the focus to the opposite is the elimination of factors which promote bad performance. Changing the focus to the opposite essentially reframes the problem, sets another perspective, breaks the usual pattern of thought and brings new exploration avenues.

In addition, focusing on how to make the component or system fail (instead of how to make it more reliable), yields important ideas for reliability enhancement and risk reduction. Playing devil’s advocate and asking not only questions about what could possibly go wrong but also questions about how to make the system/component malfunction is a useful technique for identifying rare and unexpected failure modes.

A simple example of risk reduction achieved by changing the focus to the opposite is reducing the risk of unauthorised access to a valuable service through a computer programme controlling the access (e.g. the access to a bank account). An important route to reducing the risk of unauthorised access is focusing on how to compromise the computer programme controlling the access and make it fail. Invariably, this exercise reveals a number of software vulnerabilities which could be exploited to gain unauthorised access. Common vulnerabilities discovered by using this technique are: ‘intercepting the control over the programme execution after induced division by zero,’ ‘intercepting the programme execution after induced stack overflow,’ ‘stealing the access password because of a failure to clear the keyboard buffer,’ inadequate protection against numerous simultaneous attacks using different password combinations, etc. Focusing on how to induce failure of the software programme controlling the access is a very important route to eliminating software vulnerabilities and reducing risk.

3.3. Shifting the Focus from the Components to the System

The inductive approach which is at the heart of the FMEA (Failure modes and effects analysis, MIL-STD-1629A 1977), has been an important basic technique for identifying possible failure modes at a system level during the FMEA exercise. It consists of considering sequentially the failure modes of all parts and components building the system and tracking their effect on the system’s performance.

This inductive method for revealing failure modes provided by FMEA is not sufficient to reveal the failure modes. For example, for a lubrication system, the failure modes ‘too small size of the sump’ and ‘too large clearances between lubricated contact surfaces’ cannot be discovered by the inductive approach.

Shifting the focus to the system, instead of its individual components is very useful in identifying new failure modes and is a valuable addition to FMEA. It consists of considering an already identified failure mode at a system level and investigating what kind of component-level failures could cause the system-level failure mode or contribute to it. For example, answering the question what could possibly contribute to the system failure mode ‘too low oil pressure’ identifies the important failure mode ‘too large clearances between lubricated contact surfaces due to wear out.’ Shifting the focus from the components to the system also helps to discover the additional low-level failure modes ‘leaks from seals and gaskets’ and ‘inappropriate oil with high viscosity being used.’

Changing the focus and asking the question what could possibly contribute to the system failure mode ‘too high oil pressure,’ leads to the low-level failure mode ‘incorrect design of the oil galleries.’ Asking the question what could possibly contribute to a too high oil temperature points to the cause ‘a small amount of circulating oil in the system’ which helps to identify the low-level failure modes

‘too low oil level’ and ‘too small size of the sump.’ Undersized sumps lead to a high oil temperature which constitutes a failure mode at the system level.

3.4. Starting from the Desired Ideal End Result

Formulating the desired ideal end result and looking for suitable available resources to achieve it is often an important step in finding a reliability enhancing solution.

A sprinkler system for putting out fires must be activated if the temperature reaches a particular critical threshold. Adding temperature sensors and a sprinkler activation system is an expensive option. In addition, this is a high-risk solution because the sprinkler activation system may fail and become non-operational when needed. An ideal sprinkler system operates automatically if the temperature reaches a particular critical threshold and is not operational if the temperature is below this threshold. Starting from the desired ideal sprinkler system and looking for suitable resources to achieve it, suggests using an alloy with melting temperature equal to the critical threshold temperature activating the sprinklers. This solution yields a sprinkler system which operates whenever it is needed, avoids the use of unreliable activation systems and has a very few components. The result is a significantly improved reliability and reduced risk of malfunctioning.

4. REDUCING RISK THROUGH SELF-STRENGTHENING (SELF-REINFORCEMENT)

Reducing risk through self-strengthening (self-reinforcement) is present when external/internal forces or states are transformed/channelled to induce counterforces strengthening the system’s response. An important feature of self-strengthening is that increasing the external/internal forces intensifies the system’s response against these forces.

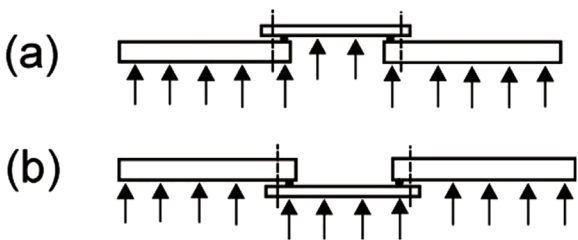
4.1. Channelling External/Internal Forces Into a Self-Strengthening Response

4.1.1. Channelling Pressure Into a Self-Strengthening Response

The self-strengthening is often closely linked with the inversion in the sense that the self-strengthening risk-reduction effect is actually derived from inverting the relative position of objects. Such a self-strengthening application is given in Figure 11, in the design of covers for containers under pressure.

For the design in Figure 11a, the loading stresses in the screws fixing the cover can be reduced and the reliability of the seal can be increased significantly by inverting the position of the cover. The design from Figure 11b channels the force due to the internal pressure into a self-strengthening force which helps to form a reliable seal and prevent leakage. Increasing the pressure increases the magnitude of the self-strengthening force, strengthens the seal and increases the resistance to the increased pressure.

Figure 11. Self-reinforcement in the design of a cover, for containers under pressure, achieved by inverting



4.1.2. Channelling Weight Into A Self-Strengthening Response

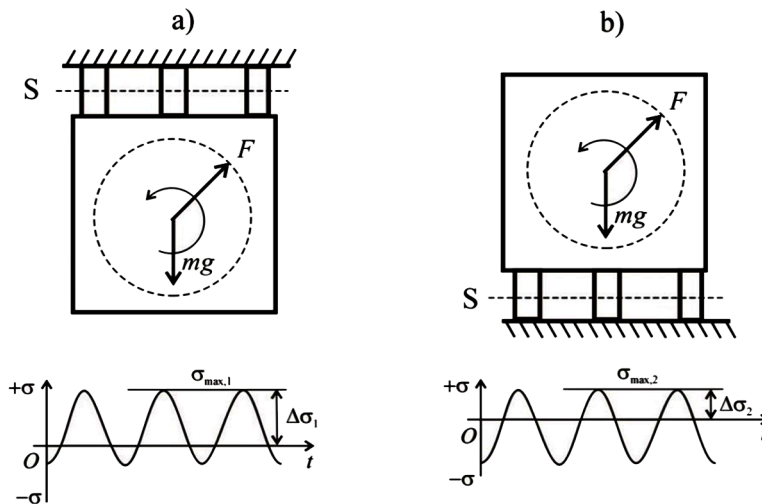
A powerful self-strengthening effect, achieved by channelling the weight as a compensating factor, by applying the inversion method, is given in Figure 12a. The brackets are supporting a part subjected to a varying force F (Figure 12a). The resultant effective stress range $\Delta\sigma_1$ present in the supporting bracket is shown in Figure 12a. The stress range $\Delta\sigma_1$ can be reduced significantly to the stress range $\Delta\sigma_2$ in Figure 12b, if the relative location of the supporting brackets and the component is inverted as it is shown in Figure 12b. In the inverted location, the weight of the component diminishes the stress range and increases the fatigue life.

Indeed, the fatigue life of the bracket is strongly dependent on the rate of fatigue crack growth. The fatigue crack growth rate da/dN is commonly estimated from the Paris power law (Paris et al., 1961; Paris and Erdogan, 1963):

$$da / dN = C \Delta K^m \quad (11)$$

where a is the crack length, N is the number of load cycles, C and m are constants depending on the material, environment, test temperature and load ratio $R = \sigma_{\min} / \sigma_{\max}$; $\Delta K = K_{\max} - K_{\min}$ is the range of the stress intensity factor. $K_{\max} = \sigma_{\max} Y \sqrt{\pi a}$ and $K_{\min} = \sigma_{\min} Y \sqrt{\pi a}$ are the stress intensity factors which correspond to the maximum (σ_{\max}) and minimum (σ_{\min}) value of the uniform tensile stress perpendicular to the crack plane. Consequently, $\Delta K = \Delta\sigma Y \sqrt{\pi a}$, where $\Delta\sigma = \sigma_{\max} - \sigma_{\min}$ is the stress range. Because fatigue cracks do not propagate if the minimum loading stress is compressive, in calculating the stress range, the minimum loading stress should be set to zero ($\sigma_{\min} = 0$) and the stress range becomes $\Delta\sigma = \sigma_{\max}$. Y is a dimensionless factor that depends on the geometry of the loaded crack.

Figure 12. Using the inversion method in the design of a bracket supporting mass experiencing variable inertia force



For relatively short cracks, it can be assumed that the geometry factor Y is independent of the crack length a and Equation (11) becomes:

$$da/dN = C \times \left(Y \Delta\sigma \sqrt{\pi a} \right)^m \quad (12)$$

where $\Delta\sigma = \sigma_{\max} - \sigma_{\min}$ is the range of the uniform tensile stress perpendicular to the crack plane. Equation (12) can be integrated by a separation of the variables. In the usual case where $m > 2$, integrating this equation results in N_f load cycles to failure, given by:

$$N_f = \frac{2}{(m-2)C \left(Y \Delta\sigma \sqrt{\pi} \right)^m} \left(\frac{1}{a_0^{(m-2)/2}} - \frac{1}{a_f^{(m-2)/2}} \right) \quad (13)$$

where a_0 is the initial crack size and a_f is the final crack size at which fast fracture is initiated. Assuming a fast fracture failure mode, the final crack size for which fracture will be initiated can be determined from Equation (14):

$$Y \sigma_{\max} \sqrt{\pi a_f} = K_{Ic} \quad (14)$$

which governs the onset of fast fracture. In Equation (14), K_{Ic} is the fracture toughness of the material and Y is the geometry factor. Equation (14) essentially states that fast fracture occurs when the stress intensity factor $Y \sigma_{\max} \sqrt{\pi a_f}$ becomes equal to the fracture toughness K_{Ic} of the material. From this equation, for the final crack size a_f , we have:

$$a_f = \frac{1}{\pi} \left(\frac{K_{Ic}}{Y \sigma_{\max}} \right)^2 \quad (15)$$

In other words, the larger the maximum stress σ_{\max} from the loading cycle, the smaller the final crack size which defines fast fracture. In other words, for the final crack sizes characterising the two designs 'a' (1) and 'b' (2) (Figure 12), we have $a_{f1} < a_{f2}$.

For the two competing designs, the effective stress ranges from the fatigue loading are different $\Delta\sigma_1 = \sigma_{\max,1} > \Delta\sigma_2 = \sigma_{\max,2}$ (Figure 12).

The different stress ranges characterizing the two designs result in different fatigue lives. According to Equation (13), the ratio of the fatigue lives characterising the two designs is:

$$\frac{N_{f1}}{N_{f2}} = \left(\frac{\Delta\sigma_2}{\Delta\sigma_1} \right)^m \left(\frac{A_1}{A_2} \right) \quad (16)$$

where $A_1 = \left(\frac{1}{a_0^{(m-2)/2}} - \frac{1}{a_{f1}^{(m-2)/2}} \right)$, $A_2 = \left(\frac{1}{a_0^{(m-2)/2}} - \frac{1}{a_{f2}^{(m-2)/2}} \right)$. Since $a_{f1} < a_{f2}$, $A_1 < A_2$. From this

result and also from $\Delta\sigma_2 < \Delta\sigma_1$, it follows that $\left(\frac{\Delta\sigma_2}{\Delta\sigma_1} \right)^m \left(\frac{A_1}{A_2} \right) < 1$. As a result, $\frac{N_{f1}}{N_{f2}} < 1$, which means that the second design is characterised by a larger fatigue life.

With increasing the weight of the component, the stress range decreases and the fatigue life increases. As a result, channeling the weight, by inverting the location of the component and the supporting brackets, results in a self-strengthened design characterised by an increased fatigue life and improved reliability. The larger the weight is, the larger is the reduction of the stress range, the larger is the increase of the fatigue life.

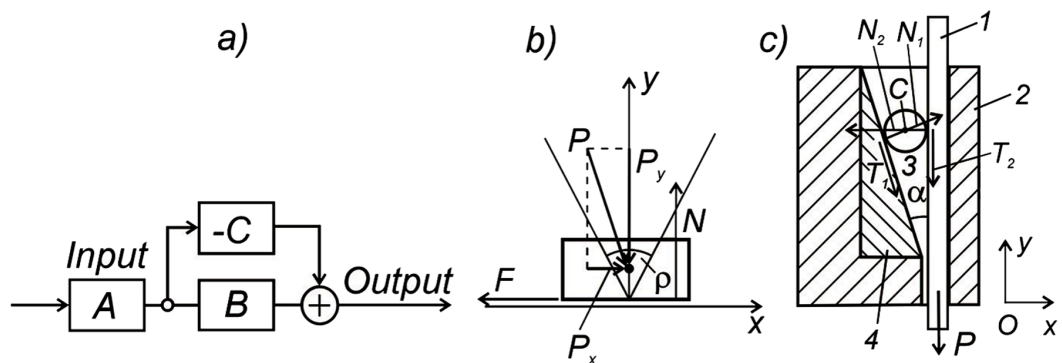
4.2. Reliability Improvement by Capturing Factors Providing a Compensating Proportional Response

The essence of this process has been captured with the abstract diagram in Figure 13a. Factor A results in the undesirable outcome B . A third factor C however, is generated whose magnitude is proportional to the magnitude of the undesirable outcome B . Now if the factor C is subtracted from the outcome B , the resultant undesirable outcome will be diminished.

This principle will be illustrated with friction forces used to produce a self-strengthening behaviour. Common applications are the self-locking screws and self-locking wedges. The larger the external loading force is, the larger are the excited friction forces in the assembly resisting the external force. Self-reinforcement is also found in other self-locking devices: (1) self-locking grips in tensile testing machines; (2) self-locking plate clamps; (3) self-locking hooks; self-locking climbing Equipment, (4) self-locking marine cleats, etc.

Increasing the external force P increases proportionally the horizontal component P_x (undesirable outcome B) working towards disturbing the equilibrium of the body. It is a well-documented fact from statics that for the simple system in Figure 13b, no matter how large the loading force P is, if it lies within the friction cone defined by the friction angle ρ ($\mu = \tan \rho$, where μ is the static coefficient of friction), the body will remain in a state of equilibrium. Increasing the magnitude of the loading force P increases the component P_x pushing the component along the horizontal axis but also increases the normal force N which in turn increases the friction force F . F is proportional to the normal force N , $F = \mu N$ and can be interpreted as the third factor C , whose magnitude is

Figure 13. a) Self-strengthening caused by a compensating third factor whose magnitude is proportional to the negative outcome B; b) A cone of friction; c) A self-strengthening design based on using the external force to excite compensating friction forces



increasing with increasing the magnitude of the undesirable component P_x (Figure 13a) and is subtracted from the output. The two forces are combined and the friction force fully compensates the undesirable force P_x . The body maintains its equilibrium state. As a result, with increasing the external load, the resistance against moving along the horizontal axis is increased.

The principle of capturing a factor which compensates an undesirable outcome thereby generating a self-strengthening effect has been used in the design from Figure 13c of a friction grip. The friction grip includes the strip 1 which is being pulled downwards with a force P , the block 2, cylinder 3 and wedge 4, whose purpose is to prevent the strip 1 from being pulled through the block 2.

The static friction coefficient between the cylinder 3 and the wedge 4 is μ_1 ; the friction coefficient between the cylinder 3 and the strip 1 is μ_2 . The forces acting on the cylinder are: the normal force N_1 and the friction force $T_1 = \mu_1 N_1$ from the wedge 4, the normal force N_2 and the friction force $T_2 = \mu_2 N_2$ from the strip 1. The cylinder 3 has a radius r and its weight has been ignored as being too small compared to the rest of the forces.

Under certain conditions, the system will be a self-strengthening (self-reinforcing) system because increasing the external load P will increase the friction forces T_2 and T_1 (factor C from Figure 13a) thereby retaining the strip and the cylinder in equilibrium.

To determine the conditions under which the system will be self-strengthening, the equilibrium conditions for the cylinder 3 are considered. They include three equations: the sum of the projections of the forces along the x -axis must equal zero; the sum of the projections of the forces along the y -axis must equal zero and the sum of the moments of the forces with respect to the center C of the cylinder must equal zero:

$$N_1 \cos \alpha + \mu_1 N_1 \sin \alpha - N_2 = 0 \quad (17)$$

$$N_1 \sin \alpha - \mu_1 N_1 \cos \alpha - \mu_2 N_2 = 0 \quad (18)$$

$$\mu_1 N_1 r - \mu_2 N_2 r = 0 \quad (19)$$

From Equation (19), $N_2 = \frac{\mu_1}{\mu_2} N_1$. Substituting $N_2 = \frac{\mu_1}{\mu_2} N_1$ in Equations (17) and (18) followed by dividing the equations by N_1 (assuming that $N_1 > 0$) results in:

$$\cos \alpha + \mu_1 \sin \alpha - \frac{\mu_1}{\mu_2} = 0 \quad (20)$$

$$\sin \alpha - \mu_1 (1 + \cos \alpha) = 0 \quad (21)$$

From Equation (21), $\mu_1 = \frac{\sin \alpha}{1 + \cos \alpha} = \frac{2 \sin(\alpha / 2) \cos(\alpha / 2)}{2 \cos^2(\alpha / 2)} = \tan(\alpha / 2)$. From Equation (20), $\mu_2 = \frac{\mu_1}{\cos \alpha + \mu_1 \sin \alpha}$.

Self-strengthening (equilibrium of the cylinder) will only be present if for the coefficients of friction, the following inequalities are simultaneously fulfilled:

$$\mu_1 \geq \tan(\alpha / 2) \quad (22)$$

$$\mu_2 \geq \frac{\mu_1}{\cos \alpha + \mu_1 \sin \alpha} \quad (23)$$

If the two coefficients of friction are equal, equation (22) is obtained as a condition for self-strengthening. An example of self-strengthening derived from capturing the weight as a proportional compensating factor has been presented in Figure 14. Figure 14a features the overturning moment on a wall resisting the pressure from a large mass of liquid or granulated substance (water, concrete, sand, etc.). If an extension is attached to the lower end of the wall (Figure 14b), the distributed weight G of the mass of the liquid (granulated substance) will create an opposing moment (factor C in Figure 13a) which counteracts the overturning moment of the force P (factor B in Figure 13a). The result is a reduced total overturning moment and a reduced risk of overturning. Increasing the mass of the liquid (granulated substance) will increase the overturning moment but will also increase the opposing moment from the weight.

The solution related to reducing the risk of overturning in Figure 14b can be replicated to reduce the risk of overturning of crowd fences. A fence without self-reinforcement (Figure 14c) can be overturned relatively easily due to the forces P_i created by people pushing the fence (Figure 14c). If the lower end of the fence on the crowd side is made wider (Figure 14d), the people will have to stand on the fence when they push against it. As a result, the distributed weight of the people (the forces G_i in Figure 14d) standing on the fence will create an opposing moment which counteracts the overturning moment of the forces P_i from the people pushing the fence. The more people push on the fence, the more weight will counteract the overturning moment. The weight of the people has been channelled towards self-strengthening of the fence.

The next self-strengthening effect has been taken from an example discussed in (Pahl et al., 2007, Figure 15a,b).

The bending stresses at the base of a blade due to the tangential forces P (undesirable outcome B , Figure 13a) can be reduced if the internal centrifugal force resulting from the rotation of the rotor is channelled to create an opposing moment counteracting the moment created by the tangential forces P . Channelling the centrifugal force to create an opposing moment (proportional compensating factor C) can be done by inclining the blades at a certain angle from the vertical (Figure 15b). For the design

Figure 14. (a,b) Reducing the risk of overturning of a wall supporting liquid or granulated substance by self-strengthening. (c,d) Reducing the risk of overturning of a crowd fence by self-strengthening

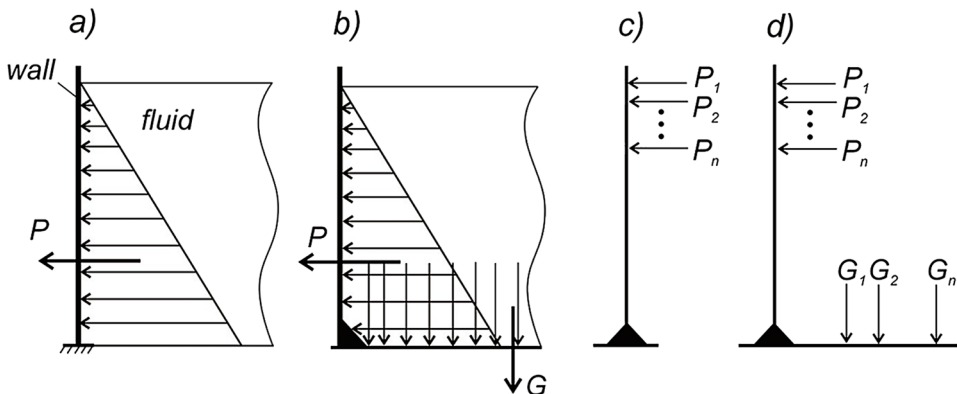
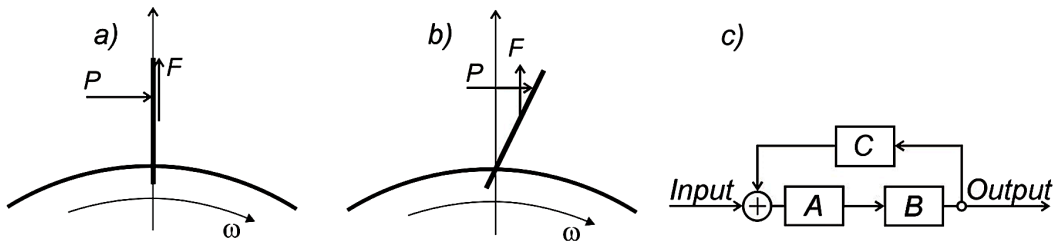


Figure 15. (a,b) Reducing the stresses at the base of a blade attached to a rotor by a proportional compensating factor ; c) a positive feedback loop



in Figure 15b, the centrifugal force F resulting from the rotation of the blade now creates a proportional moment counteracting the bending moment from the tangential forces P . This reduces the overall bending moment and the stresses at the base of the blade. The risk of failure is reduced and the life of the rotor is enhanced. The larger the tangential force P is, the larger is the rotational speed ω , the larger is the centrifugal force (proportional to ω^2), the larger is the counteracting moment created by the centrifugal force.

The example given in (Pahl et al., 2007), is essentially an example of self-strengthening by capturing a proportional compensating response.

4.3. Using Positive Feedback Loops as Self-Strengthening Mechanisms

While the application of negative feedback loops to counteract negative effects or restore the equilibrium of systems is well-understood and well-documented, there is a profound lack of discussion on the application of positive feedback loops to improve reliability. Positive feedback loops sample the output of the system and the deviation of the output is fed back to the input in such a way that a further deviation occurs in the same direction (Figure 15c). Increasing the input factor A leads to more of factor B and more of the factor B leads to a more of factor A which in turn leads to more of factor B (Figure 15c). A typical example is the exothermal chemical reaction. The heat released during the chemical reaction if not removed, causes the reaction to occur at a higher speed which in turn leads to a more released heat, etc. This positive feedback loop can easily lead to a thermal runaway and a chemical explosion. The positive feedback work towards destabilising the system's output and moving it away from equilibrium. The positive feedback loop is an important reason behind prolonged droughts, global warming, bank runs, etc. This is why it is widely believed that positive feedback loops are always associated with destruction and collapse.

Positive feedback loops however, can also be used with success for improving the reliability of systems. The self-energizing drum brakes can be given as an example of a positive feedback loop improving reliability. Once the self-energising drum brakes are engaged, the friction force between the brake pad and the drum increases the normal force on the drum, which in turn increases the friction force between the brake pad and the drum. As result of the positive feedback loop, the friction force quickly reaches a large value despite that the heavily loaded brake has initially been energized with a relatively small load. The result is an increased reliability of breaking under load.

Positive feedback can also be used in bistable systems to guarantee a fast transition to a more reliable state. If the input is slightly higher than the equilibrium state, the positive feedback causes the output to move in the same direction until the output reaches the nearest positive limit. If the input is slightly lower than the equilibrium state, the output moves fast in the same direction until it reaches the nearest negative limit. Once a positive or negative limit has been reached, the system remains in that state. The larger the deviation of the output from the specified level the larger is the input correction increasing the deviation of the output in the same direction.

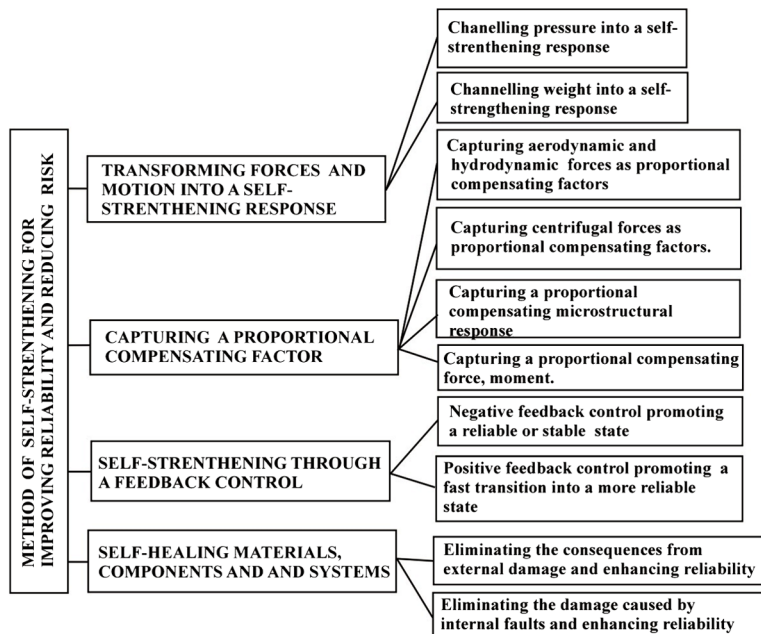
The Schmitt trigger circuit for example uses positive feedback to force the output in one of the two stable logic states. In electrical switches, the positive feedback minimises the length of arcing during the switching thereby minimising the wear out and prolonging the life of the switches.

A classification of the discussed methods for reliability improvement by self-strengthening has been presented in Figure 16. The structure provided by the classification makes the reliability improvement by the method of self-strengthening less dependent on inspiration and lateral thinking leaps.

5. CONCLUSION

1. New techniques for reliability improvement and risk reduction based on the inversion method, such as: 'inverting design variables,' 'inverting by maintaining an invariant,' 'inverting resulting in a reinforcing counter-force,' 'negating basic required functions' and 'moving backwards to general and specific contributing factors' have been formulated.
2. Inversion of the variables has been used to determine the global minimum of a multivariable function whose first partial derivatives maintain their sign in a rectangular domain. This result can be applied with success in all cases of design optimization where the first partial derivatives of the objective function retain their signs.
3. Inversion of the input variables can also be used to increase the performance of heuristic algorithms searching for a global minimum of a multivariable function. As a result of applying the variable inversion technique, the time complexity of the local search step of these algorithms can be reduced from exponential $O(2^n)$ to linear $O(n)$, which will bring a significantly improved performance.
4. By using detailed calculations, it has been demonstrated how the new technique 'repeated inversion maintaining an invariant' can be applied to reduce the risk of collision for multiple ships travelling at different times and with different speeds.

Figure 16. The method of self-strengthening for improving reliability and reducing risk



5. It has been demonstrated that for pressure vessels, inversion of the geometric parameters by maintaining an invariant volume could result not only in an increased safety but also in a significantly reduced weight.
6. The method of self-strengthening (self-reinforcement) has been introduced for the first time as a systematic method for improving reliability and reducing technical risk. An important feature of self-strengthening is that increasing the external/internal forces increases the self-strengthening effect.
7. The method of self-strengthening by capturing a proportional compensation factor has been proposed for the first time as an efficient reliability improvement tool.
8. By using detailed calculations, reliability improvement by channelling weight to reduce the stress range and improve fatigue life has been demonstrated.
9. Introducing a positive feedback loop has been proposed for the first time as an important design approach for achieving self-strengthening and reliability improvement.
10. For the first time, classifications have been proposed of methods and techniques for reliability improvement and risk reduction based on the method of inversion and the method of self-strengthening.

REFERENCES

- Altshuller, G. S. (1984). *Creativity as an exact science: The theory of the solution of inventive problems*. New York: Gordon and Breach Science Publishing.
- Altshuller, G. S. (1996). *And suddenly the inventor appeared, TRIZ, the theory of inventive problem solving, Translation from Russian by Lev Shulyak*. Worcester, MA: Technical Innovation Center.
- Altshuller, G. S. (1999). *The innovation algorithm, TRIZ, systematic innovation and technical creativity*. Worcester: Technical Innovation Center, Inc.
- Bellman, R. (1958). On a routing problem. *Quarterly of Applied Mathematics*, 16, 87–90.
- Bird, G. C., & Saynor, D. (1984). The effect of peening shot size on the performance of carbon-steel springs. *Journal of Mechanical Working Technology*, 10(2), 175–185. doi:10.1016/0378-3804(84)90065-2
- Brady, N., & Weil, R. (2007). *The Nature and Properties of Soils* (14th ed.). Prentice Hall.
- Brualdi, R. A. (2010). *Introductory Combinatorics* (5th ed.). Pentice Hall.
- Campbel, J. (2015). Complete casting handbook, Metal casting processes, Metallurgy, Techniques and design (2nd ed.). Amsterdam: Butterworth-Heinemann.
- Collins, J. A. (2003). *Mechanical design of machine elements and machines*. New York: John Wiley & Sons, Inc.
- Costache, A., Glejbøl, K., Sivebæk, I. M., & Berggreen, C. (2016). Improved Friction Joint With Self-Locking Grips. *Journal of Offshore Mechanics and Arctic Engineering*, 38(5).
- Ford, L. R. Jr. (1956). *Network Flow Theory. Paper P-923*. Santa Monica, California: RAND Corporation.
- Fu, Y., Ge, E., Su, H., Xu, J., & Li, R. (2015). Cold expansion technology of connection holes in aircraft structures: A review and prospect. *Chinese Journal of Aeronautics*, 28(4), 961–973. doi:10.1016/j.cja.2015.05.006
- Kelly P.M., & Francis-Rose L.R. (2002). Progress in Materials Science. *The Martensitic Transformation in Ceramics— its role in transformation toughening*, 47(5), 463-557.
- Kiranyaz, S., Ince, T., & Gabbouj, M. (2014). *Multidimensional Particle Swarm Optimization for Machine Learning and Pattern Recognition*. Springer. doi:10.1007/978-3-642-37846-1
- Lazinica, A. (Ed.), (2009). *Particle Swarm Optimization*. In-Tech. doi:10.5772/109
- Lebkowski, A., Tobiasz, M., Dziedzicki, K., Tomera, M., & Smierzchalski, R. (2005). Areas of dangerous approach to the ship defined with the aid of a domain. *Proceedings of ESREL 2005 conference*, Gdynia-Sopot-Gdansk.
- Lindroos, M., Apostol, M., Heino, V., Valtonen, K., Laukkanen, A., Holmberg, K., & Kuokkala, V. T. (2015). The deformation, strain hardening, and wear behaviour of chromium-alloyed hadfield steel in abrasive and impact conditions. *Tribology Letters*, 57(3), 24. doi:10.1007/s11249-015-0477-6
- Mattheck, C. (1998). *Design in nature*. Berlin: Springer Verlag.
- MIL-STD-1629A, *US Department of Defence procedure for performing a failure mode and effects analysis*. (1977). Washington, DC: US Department of Defence.
- Niku-Lari, A. (1981, September 14–17). Shot-peening, in the first international conference on shot peening, Paris. Oxford: Pergamon Press.
- Oh, J. H., Kim, K., & Jeong, J. S. (2015). A study on the risk analysis based on the trajectory of fishing vessels in the VTS area. *International journal of e-navigation and maritime economy*, 2, 38-46.
- Orloff, M. (2006). *Inventive thinking through TRIZ* (2nd ed.). Springer.
- Pahl, G., Beitz, W., Feldhusen, J., & Grote, K. H. (2007). *Engineering design*. Berlin: Springer. doi:10.1007/978-1-84628-319-2
- Paris, P. C., & Erdogan, F. (1963). A critical analysis of crack propagation laws. *Journal of Basic Engineering*, 85(4), 528–534. doi:10.1115/1.3656900

Paris, P. C., Gomez, M. P., & Anderson, W. P. (1961). A rational analytic theory of fatigue. *The Trend in Engineering*, 13, 9–14.

Potter, J. (Ed.). (1985, March 18-19). Fatigue in mechanically fastened composite and metallic joints. Proceedings of the symposium sponsored by ASTM ASTM special technical publication.

Todinov, M. T. (1999). Maximum principal tensile stress and fatigue crack origin for compression springs. *International Journal of Mechanical Sciences*, 41(3), 357–370. doi:10.1016/S0020-7403(98)00068-X

Todinov, M. T. (2000). Residual stresses at the surface of automotive suspension springs. *Journal of Materials Science*, 35(13), 3313–3320. doi:10.1023/A:1004887708822

Todinov, M. T. (2007). *Risk-based reliability analysis and generic principles for risk reduction*. Elsevier.

Todinov, M. T. (2015). Reducing risk through segmentation, permutations, time and space exposure, inverse states and separation. *International Journal of Risk and Contingency Management*, 4(3), 1–21. doi:10.4018/IJRCM.2015070101

Wirth, N. (1976). *Algorithms+data structures = programs*. Prentice-hall.

Zhang, L., Konyukhov, A., Mok, E., & Choo, H. (2013). Increasing the road safety of e-bike: design of protecting shell based on stability criteria during severe road accidents, *Proceedings of the 2013 International conference on electrical, control and automation engineering (ECAE '13)* (pp. 485-489).

Michael Todinov holds a PhD related to mathematical modelling of thermal and residual stresses and a higher doctorate Doctor of Engineering (DEng) which is the engineering equivalent of Doctor of Science (DSc) in the area of new probabilistic concepts and models in Engineering. M.Todinov's name is associated with creating the foundations of risk-based reliability analysis (driven by the cost of failure) and the theory of repairable flow networks and networks with disturbed flows. A sample of M.Todinov's results includes: the discovery of closed and dominated parasitic flow loops in real networks; the proof that the Weibull distribution is an incorrect model for the distribution of breaking strength of materials and the derivation of the correct alternative of the Weibull model; a theorem regarding the exact upper bound of properties from random sampling of multiple sources; a general equation for the probability of failure of brittle components with complex shape, the formulation and proof of the necessary and sufficient conditions of the Palmgren-Miner rule and Scheil's additivity rule and deriving the correct alternative of the Johnson-Mehl-Avrami-Kolmogorov equation. M.Todinov's research has been funded by research councils, the automotive industry, the nuclear industry and the oil and gas industry.