Sensitivity analysis methods for reliability problems

K.W. Breitung
Technical University of Munich, Germany

ABSTRACT: In reliability problems an important aspect is the study of the influence of parameter changes on the target reliability which is often more an operational value. Since in complex problems such quantities have to be computed often, efficient methods for doing this are desirable. Here some methods for sensitivity analysis are outlined. For a FORM/SORM analysis it is possible to get sensitivity measures from the Lagrange multiplier at the beta point. This allows simple estimates for the sensitivities with respect to parameters without additional computations. In the general case the partial derivatives and sensitivities of the failure probability with respect to parameters are given by surface integrals over the limit state surface. Such integrals can be transformed into domain integrals over the safe domain using the divergence theorem (Gauss-Ostrogradsky theorem). By modifying the integrands in a suitable way, it is possible to modify the integration domains such that the integrals can be estimated in a more efficient way.

1 INTRODUCTION

In many reliability problems one has parameters which can be varied or are not exactly known. Therefore the influence of changes in these parameters is an important information in studying such problems.

The basic reliability problem is in the form

$$P = \int_{g(\boldsymbol{x})\le 0} f(\boldsymbol{x})\ \mathrm{d}\boldsymbol{x} \tag{1}$$

where P is the failure probability, $f(\boldsymbol{x})$ is the probability density function and $g(\boldsymbol{x})$ the limit state function. If we consider the existence of parameters, the more general form of this problem can be written as:

$$P(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \int_{g(\boldsymbol{x},\boldsymbol{\theta}_1)\le 0} f(\boldsymbol{x}, \boldsymbol{\theta}_2)\ \mathrm{d}\boldsymbol{x} \tag{2}$$

Here the first parameter vector $\boldsymbol{\theta}_1$ includes the parameters of the limit state function and the second vector $\boldsymbol{\theta}_2$ includes the parameters of the probability density function.

For sake of simplicity we will study here only the case that the limit state function depends on a single parameter τ, i.e. integrals of the form:

$$P(\tau) = \int_{g(\boldsymbol{x},\tau)\le 0} f(\boldsymbol{x})\ \mathrm{d}\boldsymbol{x} \tag{3}$$

These results can be generalized for the more complex cases.

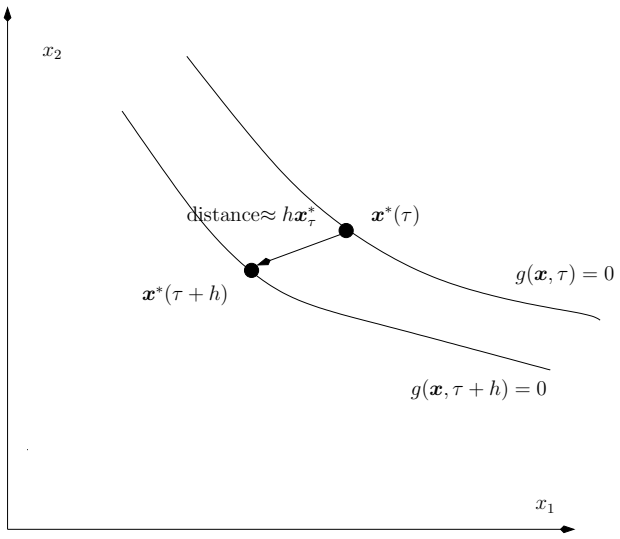

Figure 1: The movement of the PML-point as function of τ

2 PARAMETER DEPENDENCE OF THE PML-POINT

For the functions f and $g: \mathbb{R}^n \to \mathbb{R}$ in eq. (3) let the gradients of these functions be denoted by $\nabla_{\boldsymbol{x}} f$ (resp. $\nabla_{\boldsymbol{x}} g$) and the partial derivative with respect to τ by g_τ. The Hessian of f is written as $\boldsymbol{H}_f$ and the Hessian of the function $g(\boldsymbol{x},\tau)$ with respect to the first n variables is written as $\boldsymbol{H}_g$. We assume that for a fixed value of τ there is a unique PML (Point of maximum likelihood) $\boldsymbol{x}^*(\tau)$ and we write in a shorthand notation $\boldsymbol{x}^*$. The vector of the first derivatives of $\boldsymbol{x}^*$ with respect to τ is written as $\boldsymbol{x}^*_\tau$.

In the case of a standard normal distribution the PML is the beta point. This point $\boldsymbol{x}^*$ is a stationary point of the Lagrangian function $L(\boldsymbol{x},\lambda,\tau)$ defined by

$$L = f - \lambda g \tag{4}$$

Therefore for the point $\boldsymbol{x}^*$ the following equation system must be fulfilled for arbitrary τ:

$$\begin{aligned} \nabla_{\boldsymbol{x}} f - \lambda \nabla_{\boldsymbol{x}} g &= \boldsymbol{o}_n, \\ g &= 0 \end{aligned} \tag{5}$$

with $\boldsymbol{o}_n$ the n-dimensional zero row vector.

To find now the derivatives of the coordinates of the PML with respect to changes in the parameter τ, we differentiate this system with respect to τ and set all derivatives equal to zero. Differentiating the term in the i-th row ($1 \le i \le n$) gives

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}\tau} \left[\frac{\partial f(x_1^*(\tau),\ldots,x_1^*(\tau))}{\partial x_i} - \lambda \frac{\partial g(x_1^*(\tau),\ldots,x_1^*(\tau),\tau)}{\partial x_i}\right] \\ &= \sum_{j=1}^{n} \left[\frac{\partial^2 f(x_1^*(\tau),\ldots,x_n^*(\tau))}{\partial x_i \partial x_j} - \lambda \frac{\partial^2 g(x_1^*(\tau),\ldots,x_n^*(\tau),\tau)}{\partial x_i \partial x_j}\right] \frac{\mathrm{d}x_j^*(\tau)}{\mathrm{d}\tau} \\ &- \lambda_\tau \frac{\partial g(x_1^*(\tau),\ldots,x_1^*(\tau),\tau)}{\partial x_i} - \lambda \frac{\partial g(x_1^*(\tau),\ldots,x_1^*(\tau),\tau)}{\partial x_i \partial \tau} = 0 \end{aligned} \tag{6}$$

and for the last line we get:

$$\sum_{i=1}^{n} \frac{\partial g(x_1^*(\tau),\ldots,x_1^*(\tau),\tau)}{\partial x_i} \frac{\mathrm{d}x_i^*(\tau)}{\mathrm{d}\tau} + \frac{\partial g(x_1^*(\tau),\ldots,x_n^*(\tau),\tau)}{\partial \tau} = 0 \tag{7}$$

This yields written in vector notation:

$$\begin{aligned} \boldsymbol{H}_f \boldsymbol{x}_\tau^* - \lambda \boldsymbol{H}_g \boldsymbol{x}_\tau^* - \lambda_\tau \nabla_{\boldsymbol{x}} g - \lambda \nabla_{\boldsymbol{x}} g_\tau &= \boldsymbol{o}_n, \\ (\nabla_{\boldsymbol{x}} g)^T \boldsymbol{x}_\tau^* + g_\tau &= 0 \end{aligned} \tag{8}$$

Rearranging the terms we get:

$$\begin{aligned} (\boldsymbol{H}_f - \lambda \boldsymbol{H}_g) \boldsymbol{x}_\tau^* - \lambda \nabla_{\boldsymbol{x}} g_\tau - \lambda_\tau \nabla_{\boldsymbol{x}} g &= \boldsymbol{o}_n \\ (\nabla_{\boldsymbol{x}} g)^T \boldsymbol{x}_\tau^* + g_\tau &= 0 \end{aligned} \tag{9}$$

This is a linear equation system with the $n+1$ unknowns $\boldsymbol{x}_\tau^*$ and λ_τ

$$\begin{aligned} (\boldsymbol{H}_f - \lambda \boldsymbol{H}_g) \boldsymbol{x}_\tau^* - \lambda_\tau \nabla_{\boldsymbol{x}} g &= \lambda \nabla_{\boldsymbol{x}} g_\tau \\ -(\nabla_{\boldsymbol{x}} g)^T \boldsymbol{x}_\tau^* + \lambda_\tau \cdot 0 &= g_\tau \end{aligned} \tag{10}$$

The solution is then:

$$\begin{pmatrix} \boldsymbol{x}_\tau^* \\ \lambda_\tau \end{pmatrix} = \begin{pmatrix} \boldsymbol{H}_f - \lambda \boldsymbol{H}_g & -\nabla_{\boldsymbol{x}} g \\ -\nabla_{\boldsymbol{x}} g^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} \lambda \nabla_{\boldsymbol{x}} g_\tau \\ g_\tau \end{pmatrix}$$

The result for standard normal densities was derived by Enevoldsen (1994). This can be generalized for the case of several equality and inequality constraints.

3 SURFACE INTEGRALS OVER BOUNDARIES OF STAR-SHAPED DOMAINS

For some special cases it is easy to calculate surface integrals, i.e. convex and star-shaped domains. For a given star-shaped domain D (i.e, for every point in the domain D all points on the straight line between the origin and this point are in D) we consider a surface integral over the boundary of the domain given by limit state function $g(\boldsymbol{x})$

$$I = \int_{g(\boldsymbol{x})=0} h(\boldsymbol{x}) \, \mathrm{d}s(\boldsymbol{x}) \tag{11}$$

with $\mathrm{d}s(\boldsymbol{x})$ denoting surface integration. This integral can be computed using directional sampling. Let $\boldsymbol{u}_i$, $i = 1, \ldots, k$ be independent random unit vectors each with a uniform distribution over the n-dimensional sphere. Then an estimate of the integral is given by

$$\hat{I} = k^{-1} \sum_{i=1}^{k} F(\boldsymbol{u}_i) \tag{12}$$

with

$$F(\boldsymbol{u}_i) = \left(\frac{2\pi^{n/2}}{\Gamma(n/2)} \right) h(\boldsymbol{x}_i) \frac{|\boldsymbol{x}_i|^{n-1}}{|\boldsymbol{u}_i^T \cdot \boldsymbol{n}(\boldsymbol{x}_i)|} \tag{13}$$

where $\boldsymbol{x}_i$ is the point where the ray from the origin in the direction of the normal $\boldsymbol{u}_i$ hits the surface and $\boldsymbol{n}(\boldsymbol{x}_i) = |\nabla g(\boldsymbol{x})|^{-1} \nabla g(\boldsymbol{x})$ is the surface normal of the surface $\{g(\boldsymbol{x}) = 0\}$ at this point. The term in the round brackets is the surface of the unit sphere in the n-dimensional Euclidean space. The term in the second denominator is the cosine between the directions of $\boldsymbol{u}_i$ and $\boldsymbol{n}(\boldsymbol{x}_i)$. The denominator adjusts the weight of the sample point to take into account the projection onto the unit sphere (see for example Thomas and Finney (1988), p. 1040).

For the general case Straub (2011) proposed a method based on replacing the surface integral by a domain integral using an auxiliary variable. Here we describe an alternative which allows also to calculate such surface integrals with holes in the safe or unsafe domain. Given is a limit state function $g(\boldsymbol{x}, \tau)$ depending on a parameter τ. To find the derivative of the integral with respect to the parameter τ, we write the difference

$$P(\tau + h) - P(\tau) = \int_{g(\boldsymbol{x},\tau+h)\leq 0} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} - \int_{g(\boldsymbol{x},\tau)\leq 0} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{14}$$

This can be written making a coordinate transformation (see figure 2) as

$$\int_{g(\boldsymbol{x},\tau)=0} \int_0^{\delta(\boldsymbol{x})} f(\boldsymbol{x} + \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}))J(\boldsymbol{x} + \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}))\mathrm{d}s_\tau(\boldsymbol{x}) \tag{15}$$

$J(\boldsymbol{x}+\delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}))$ is the Jacobian of the coordinate transformation and the outward point-

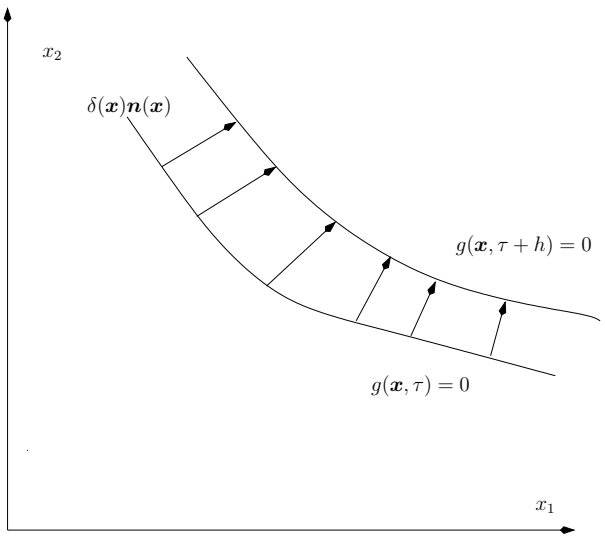

Figure 2: The difference $P(\tau + h) - P(\tau)$

ing surface normal is $\boldsymbol{n}(\boldsymbol{x}) = \nabla_{\boldsymbol{x}} g(\boldsymbol{x}, \tau)|\nabla_{\boldsymbol{x}} g(\boldsymbol{x}, \tau)|^{-1}$. Further $\mathrm{d}s_\tau(\boldsymbol{y})$ denotes surface integration over the surface $\{g(\boldsymbol{x}, \tau) = 0\}$. Due to the choice of the coordinate transformation we have that $J(\boldsymbol{x}) = 1$. The value of δ is implicitly defined by

$$g(\boldsymbol{x} + \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}), \tau + h) = 0 \tag{16}$$

To find an explicit formula for $\delta(\boldsymbol{x})$, we make a first order Taylor expansion of $g(\boldsymbol{x}, \tau)$.

$$g(\boldsymbol{x} + \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x}), \tau + h) \approx \underbrace{g(\boldsymbol{x}, \tau)}_{=0} + \delta(\boldsymbol{x})\boldsymbol{n}(\boldsymbol{x})^T \nabla_{\boldsymbol{x}} g(\boldsymbol{x}, \tau) + h g_\tau(\boldsymbol{x}, \tau) \tag{17}$$

Setting the lefthand side equal to zero yields then

$$\delta(\boldsymbol{x}) \approx h \cdot \frac{-g_\tau(\boldsymbol{x}, \tau)}{|\nabla_{\boldsymbol{x}} g(\boldsymbol{x}, \tau)|} \ . \tag{18}$$

This derivation is outlined in more detail in Breitung (1994), p. 23-5.

For an integral in the form as in eq. (3) the partial derivative with respect to τ is given by:

$$P'(\tau) = \int_{g(\boldsymbol{y},\tau)=0} f(\boldsymbol{y}) \frac{-g_\tau(\boldsymbol{y}, \tau)}{|\nabla_{\boldsymbol{y}} g(\boldsymbol{y}, \tau)|} \, \mathrm{d}s_\tau(\boldsymbol{y}) \tag{19}$$

Now we will try using the divergence theorem (also called Gauss-theorem or Gauss-Ostrogradsky-theorem) for transforming this surface integral into a domain integral. For a domain $\{g(\boldsymbol{x},\tau) > 0\}$ and a vector function $\boldsymbol{\Phi}(\boldsymbol{x})$ defined on this domain, this theorem says that:

$$\int_{g(\boldsymbol{y},\tau)=0} \boldsymbol{\Phi}(\boldsymbol{y})^T \boldsymbol{n}(\boldsymbol{y}) \mathrm{d}s_\tau(\boldsymbol{y}) = \int_{g(\boldsymbol{x},\tau)>0} \mathrm{div}(\boldsymbol{\Phi}(\boldsymbol{x}))\ \mathrm{d}\boldsymbol{x} \tag{20}$$

where $\boldsymbol{n}$ is the outward pointing normal as defined before.

This result can be understood easily if the vector function $\boldsymbol{\Phi}(\boldsymbol{x})$ is interpreted as the flow velocity of a fluid. The surface integral gives the total flow out of the domain $\{g(\boldsymbol{x},\tau) \geq 0\}$, the domain integral the local changes in the flow. This result is valid even if there are holes in the domain as shown in figure 3. The integral of the divergence over the domain by the horizontal stripes gives the flows over the two dashed curves. If the integral of the divergence of the function is known over the domain with the horizontal stripes and the flow over one of the curves also we can calculate the flow over the other curve. By

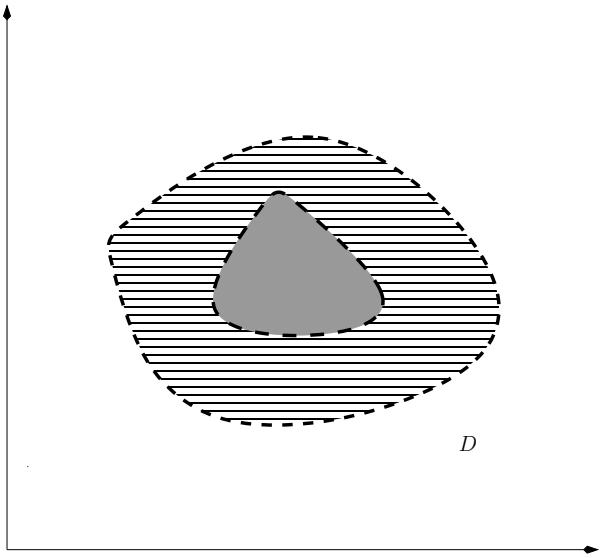

Figure 3: The divergence theorem for domains with holes

choosing an appropriate form for the vector function $\boldsymbol{\Phi}$ we can ensure that the scalar product of it with the surface normal on the limit state surface is equal to the function $f(\boldsymbol{y}) g_\tau(\boldsymbol{y},\tau) |\nabla g_{\boldsymbol{y}}(\boldsymbol{y},\tau)|^{-1}$ there. This is the case if we set

$$\boldsymbol{\Phi}(\boldsymbol{x}) = -f(\boldsymbol{x}) \frac{g_\tau(\boldsymbol{x},\tau) \nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)}{|\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)|^2} = -f(\boldsymbol{x}) \frac{g_\tau(\boldsymbol{x},\tau)}{|\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)|} \boldsymbol{n}(\boldsymbol{x}) \tag{21}$$

If we now try to apply the divergence theorem by calculating the divergence of the vector function $\boldsymbol{\Phi}(\boldsymbol{x})$, we see that the integrand is not defined in the whole integration domain. The domain is the set where $g(\boldsymbol{x},\tau) \geq 0$ and therefore the function will have at least one maximum there which means that the gradient $\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)$ vanishes there. The integrand will be undefined at this maximum point and also at all other points where $\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau) = \boldsymbol{o}_n$, see figure 4. To resolve this problem, first we can try to exclude small neighborhoods of these points from the integration domain. In Breitung (1994) some results, theorem 22 and corollary 23 are proven for solving this difficulty, but these results are valid only if the Hessians $\boldsymbol{H}_g$ at the points where the gradient vanishes are **definite**, not as erroneously stated there if they are regular. One way now would be to exclude all these points where the gradient is vanishing by introducing small spheres around them which are excluded from the integration domain.

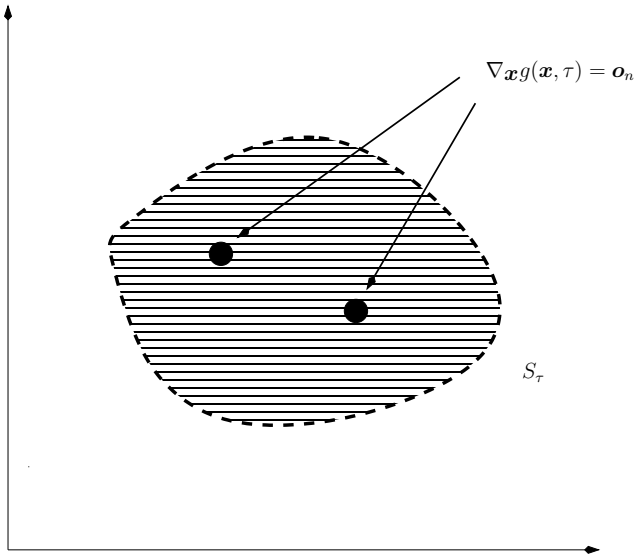

Figure 4: Points to be excluded

Another remedy for the problem of the zero gradients is to change the limit state function so that such zeros can be avoided. Let be given an upper bound for the norm of the gradient $\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)$, i.e.

$$M = \max_{\boldsymbol{x}\in S} |\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)| \tag{22}$$

We define a domain

$$S_{\epsilon,\beta_0} = \{|\boldsymbol{x}| > \beta_0,\ g(\boldsymbol{x},\tau) > \epsilon\} \tag{23}$$

Such a domain is shown in figure 5. This is a subset of the safe domain if β_0 is chosen not

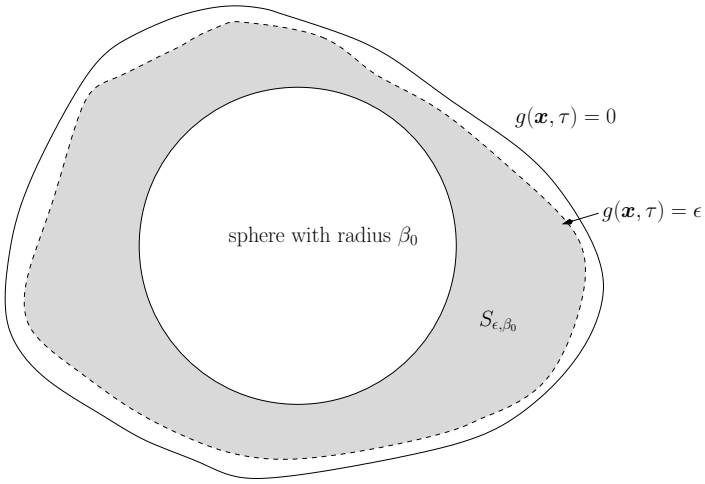

Figure 5: The domain S_{ϵ,β_0}

too large. Now a new limit state function g^* is defined by

$$g^*(\boldsymbol{x},\tau) = g(\boldsymbol{x},\tau)\exp\left(\mu|\boldsymbol{x}|^2/2\right) \tag{24}$$

where $\mu > 0$ is a parameter. This function defines the same limit state surface as the original limit state function g. The gradient of this new limit state function is:

$$\begin{aligned} \nabla_{\boldsymbol{x}} g^*(\boldsymbol{x},\tau) &= \exp(\mu|\boldsymbol{x}|^2/2)\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau) + \mu g(\boldsymbol{x},\tau)\exp(\mu|\boldsymbol{x}|^2/2)\boldsymbol{x} \\ &= \exp(\mu|\boldsymbol{x}|^2/2)\left(\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau) + \mu g(\boldsymbol{x},\tau)\boldsymbol{x}\right) \end{aligned} \tag{25}$$

Under which conditions does the gradient $\nabla_{\boldsymbol{x}} g^*$ vanish in the domain S_{ϵ,β_0}? It will be equal to the zero vector if

$$\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau) + \mu g(\boldsymbol{x},\tau)\boldsymbol{x} = \boldsymbol{o}_n. \tag{26}$$

A necessary condition for this is:

$$|\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)| = \mu |g(\boldsymbol{x},\tau)||\boldsymbol{x}| \tag{27}$$

Since we assumed that $|\boldsymbol{x}| > \beta_0$ this condition says:

$$|\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)| \geq \mu |g(\boldsymbol{x},\tau)|\beta_0 \tag{28}$$

Now, further in the domain of integration we have that $|g(\boldsymbol{x},\tau)| > \epsilon$; this yields then

$$|\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)| \geq \mu\epsilon\beta_0 \tag{29}$$

Now, since M is an upper bound for the norm of the gradient in S, we find that the inequality

$$M \geq \mu\epsilon\beta_0 \tag{30}$$

$$\mu \leq \frac{M}{\epsilon\beta_0} \tag{31}$$

has to be fulfilled if there is a point in $S_{\beta_0,\epsilon}$ where the gradient is zero. If μ is chosen larger than $M/(\beta_0\epsilon)$ there are no zeros of the gradient in the domain and the divergence theorem can be applied to calculate the integral of the divergence over S_{ϵ,β_0}.

For arbitrary ϵ we can integrate over the domain S_{ϵ,β_0} if the parameter μ is chosen large enough in dependence from ϵ. This allows to approximate the integral over the whole domain $\{|\boldsymbol{x}| > \beta_0, g(\boldsymbol{x},\tau) > 0\}$ by extrapolating from these results for $\epsilon \to 0$.

The divergence can now be written in a simpler form using the result of Gradshteyn and Ryzhik (1980), p. 1116, equation $4'$. If ϕ is a function and $\boldsymbol{\Psi}$ a vector function $\mathbb{R}^n \to \mathbb{R}^n$ then

$$\mathrm{div}(\phi \cdot \boldsymbol{\Psi}) = (\nabla\phi)^T \boldsymbol{\Psi} + \phi \cdot \mathrm{div}(\boldsymbol{\Psi}) \tag{32}$$

This gives with the definition $h(\boldsymbol{x}) = -f(\boldsymbol{x}) g_\tau(\boldsymbol{x},\tau) |\nabla_{\boldsymbol{x}} g(\boldsymbol{x},\tau)|^{-1}$ the following form for the integrand in eq.(21)

$$\mathrm{div}(h\boldsymbol{n}) = (\nabla h)^T \boldsymbol{n} + h \cdot \mathrm{div}(\boldsymbol{n}) \tag{33}$$

Now, the divergence of the normal vector is related to the mean curvature $H(\boldsymbol{x})$ of the surface at this point by (see for example, exercise 12.18, p. 94 in Thorpe (1979)):

$$\mathrm{div}(\boldsymbol{n}) = -(n-1)H(\boldsymbol{x}) \tag{34}$$

Therefore:

$$\mathrm{div}(h\boldsymbol{n}) = (\nabla h)^T \boldsymbol{n} - h \cdot (n-1)H(\boldsymbol{x}) \tag{35}$$

For the case that the safe domain is unbounded, we can get a solution by considering a second sphere so far outside that the integral over this sphere of the integrand is negligible. Then we can estimate the integral over the limit state surface between the larger sphere and a smaller sphere inside using the divergence theorem.

5 CONCLUSIONS

Several methods for computing surface integrals were outlined. The method based on the divergence theorem allows to compute such integrals for arbitrary shapes by transforming them to domain integrals.

REFERENCES

Breitung, K. (1994). *Asymptotic Approximations for Probability Integrals*. Berlin: Springer. Lecture Notes in Mathematics, Nr. 1592.

Enevoldsen, I. (1994). Sensitivity Analysis of Reliability-Based Optimal Solution. *Journal of the Engineering Mechanics Division ASCE 120*(1), 198–205.

Gradshteyn, I. and I. Ryzhik (1980). *Table of Integrals, Series and Products*. New York: Academic Press.

Straub, D. (2011). Reliability updating with equality information . *Probabilist Engineering Mechanics 26*(2), 254–258.

Thomas, G. and R. Finney (1988). *Calculus and Analytic Geometry* (7th ed.). Reading, Mass.: Addison-Wesley.

Thorpe, J. (1979). *Elementary Topics in Differential Geometry*. New York: Springer.