Chapter 11 in
Engineering Design Reliability Handbook
Edited by E. Nikolaides, D. Ghiocel and S. Singhal, CRC Press, 2005

Info-gap Decision Theory For Engineering Design. Or: Why 'Good' is Preferable to 'Best'

Yakov Ben-Haim*
Yitzhak Moda'i Chair in Technology and Economics
Faculty of Mechanical Engineering
Technion — Israel Institute of Technology
Haifa 32000 Israel
yakov@technion.ac.il
http://tx.technion.ac.il/~yakov

Contents

papers crc gpb.tex 11.2.2003
* This paper was written while the author was a visiting professor at the Georgia Institute of Technology.

Abstract

Designers face an ineluctable trade-off of system-performance against robustness to information-gaps in the designer's knowledge base. A design which optimizes performance, based on the best available models and data, will have no immunity to deficiencies in those models and data. Immunity is obtained only by relinquishing aspiration for high performance. Info-gap uncertainty is ignorance or incomplete understanding of the systems and processes being optimized. This is broader than usually treated with probability theory. The strategy advocated here is one of robust-satisficing. The robustness function is the greatest horizon of info-gap uncertainty within which the performance is guaranteed to meet the aspirations. For fixed design: as the aspirations become more demanding, the immunity to uncertainty becomes lower and finally vanishes when maximal performance is demanded. Robustness can be recovered only by retreating from maximal aspiration. For fixed aspirations: one design is preferred over another if the first entails greater robustness than the latter. This search for robustifying designs is feasible (contains a non-empty search set) only when the aspiration is sub-optimal. Sub-optimal designs can have greater robustness than performance-optimal designs, when evaluated at the same performance requirement. We consider four examples: designing the shape of a cantilever, maneuvering a dynamic system, identifying a system model, and supervising a go/no-go decision. One theorem is presented which establishes the theoretical foundations of this analysis.

1 Introduction and Overview

We call ourselves *Homo sapiens,* in part because we value our ability to optimize, but our sapience is not limited to the persistent pursuit of unattainable goals. Rather, what eons have taught the species is the lesson of balancing goals against the constraints of resources, knowledge and ability. Indeed, the conjunction of reliability-analysis and system-design is motivated precisely by the need to balance idealized goals and realistic constraints.

The reliability-analyst/system-designer seeks to optimize the design, so the question is: What constitutes feasible optimization? First we must recognize that even our best models are wrong in ways we perhaps cannot even imagine. In addition, our most extensive data is incomplete and especially lacks evidence about surprises — catastrophes as well as windfalls — which impact the success and survival of the system. These model- and data-deficiencies are information-gaps or epistemic uncertainties. We will show that optimization of performance is always accompanied by minimization of robustness to epistemic uncertainty. That is, performance and robustness are antagonistic attributes and one must be traded-off against the other. A performance-maximizing option will have less robustness against unmodelled information-gaps than some sub-optimal option, when both are evaluated against the same aspiration for performance. The conclusion is that the performance-sub-optimal design is preferable over the performance-optimal design.

The principles just mentioned are explicated with four examples in sections 2–5. These sections can be read independently, though the examples supplement one another by emphasizing different applications, aspects of the problem, and methods of analysis. A theoretical framework is provided in section 6.

Section 2 considers the design of the profile of a cantilever which is subjected to uncertain static loads. We first formulate a traditional design analysis, in the absence of uncertainty, which leads to a family of performance-optimal designs. These designs are Pareto-efficient trade-offs between minimizing the stress and minimizing the weight of the beam. We then show that these Pareto-efficient designs in fact have no immunity to info-gaps in the load. Robustness to load-uncertainty is obtained only by moving off the Pareto-optimal design surface. We also consider the windfall gains which can be garnered from load uncertainty, and examine the relation between robust and opportune designs.

Section 3 examines the maneuvering of a vibrating system whose impulse response function is incompletely known. The emphasis is not on control technology, but rather on modelling and managing unstructured info-gaps in the design-base model of the system. We formulate a traditional performance-optimization of the control input based on the best-available model. We demonstrate that this performance-optimization has no immunity to the info-gaps which plague the design-base dynamic model. This leads to the analysis of performance-sub-optimal designs which magnify the immunity to uncertainty in the dynamic behavior of the system. A simple numerical example shows that quite large robustness can be achieved with sub-optimal designs, while satisficing the performance (making the performance good enough) at levels not too much less than the performance-optimum.

Section 4 differs from the previous engineering-design examples, and considers the process of up-dating the parameters of a system model, based on data, when the basic structure of the model is inaccurate. The case examined is the impact of unmodelled quadratic non-linearities. We begin by formulating a standard model up-dating procedure based on maximizing the fidelity between the model and the data. We then show that the result of this procedure has no robustness to the structural deficiencies of the model. We demonstrate, through a simple numerical example, that fidelity-sub-optimal models can achieve substantial robustness to model-structure errors, while satisficing the fidelity at levels not too far below the fidelity-optimum.

In section 5 we consider hybrid uncertainty: a combination of epistemic info-gaps and explicit (though imprecise) probability densities. A go/no-go decision is to be made based on the evaluation of the probability of failure. This evaluation is based on the best available probability density function. However, this probability density is recognized to be imperfect, which constitutes the info-gap which beleaguers the go/no-go decision. An info-gap analysis is used to address the question: how reliable is the probabilistic go/no-go decision, with respect to the unknown error in the probability density. In short, the info-gap robustness-analysis supervises the probabilistic decision.

The examples in sections 2–4 illustrate the assertion that performance-optimization will lead inevitably to minimization of immunity to information-gaps. This suggests that performance should be satisficed — made adequate but not optimal — and that robustness should be optimized. Section 6 provides a rigorous theoretical basis for these conclusions.

2 Design of a Cantilever with Uncertain Load

In this section we will formulate a simple design problem, and solve it by finding the design which optimizes a performance criterion. We will then show that this solution has no robustness to uncertainty: infinitesimal deviations (of the load, in this example) can cause violation of the design criterion. This will illustrate the general conclusion (to be proven later) that optimizing the performance in fact minimizes the robustness to uncertainty. Stated differently, we will observe that a designer's aspiration for high performance must be accompanied by the designer's acceptance of low robustness to failure. Conversely, feasible solutions will entail sub-optimal performance. Again stated differently, the designer faces an irrevocable trade-off between performance and robustness-to-failure: demand for high performance is vulnerable to uncertainty; modest performance requirements are more immune to uncertainty.

2.1 Performance Optimization

Consider a uniform cantilever of length L [m] subject to a continuous uniform load density $\widetilde{\phi}$ [N/m] applied in a single plane perpendicular to the beam axis, as in fig. 1. The beam is rectangular in cross section. The beam width, w [m], is uniform along the length and determined by prior constraints, but the thickness in the load plane, $T(x)$ [m], may be chosen by the designer to vary along the beam. The beam is homogeneous and its density is known. The designer wishes to choose the thickness profile to minimize the mass of the beam and also to minimize the maximum absolute bending stress

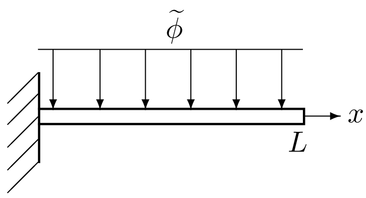

Figure 1: Cantilever with uniform load.

in the beam. These two optimization criteria for selecting the thickness profile $T(x)$ are:

$$\min_{T(x)>0} \int_0^L T(x)\,\mathrm{d}x \tag{1}$$

$$\min_{T(x)>0} \max_{0\le x\le L} |\sigma_T(x)| \tag{2}$$

where $\sigma_T(x)$ is the maximum bending stress in the beam section at x. The positivity constraint on $T(x)$ arises because the thickness must be positive at every point, otherwise the "beam" is not a beam.

These two design criteria are in conflict, so a trade-off between mass- and stress-minimization will be needed. To handle this we will solve the stress-minimization with the mass constrained to a fixed value. We will then vary the beam mass. Consider the following set of thickness profiles corresponding to fixed mass:

$$\Theta(\theta) = \left\{ T(x) : \int_0^L T(x)\,\mathrm{d}x = \theta \right\} \tag{3}$$

For any given value of θ, which determines the beam mass, we choose the thickness profile from $\Theta(\theta)$ to minimize the maximum stress. The performance criterion by which a design proposal $T(x)$ is evaluated is:

$$R(T) = \max_{0\le x\le L} |\sigma_T(x)| \tag{4}$$

The design which optimizes the performance from among the beams in $\Theta(\theta)$, which we denote $\widehat{T}_\theta(x)$, is implicitly defined by:

$$R(\widehat{T}_\theta) = \min_{T(x)\in\Theta(\theta)} \max_{0\le x\le L} |\sigma_T(x)| \tag{5}$$

From the small-deflection static analysis of a beam with thickness profile $T(x)$, one finds the magnitude of the maximum absolute bending stress at section x to be:

$$|\sigma_T(x)| = \frac{3\widetilde{\phi}(L-x)^2}{wT^2(x)} \quad [\mathrm{Pa}] \tag{6}$$

where $x = 0$ at the clamped end of the beam.

In light of the integral constraint on the thickness profile, $T(x) \in \Theta(\theta)$ in eq.(3), and of the demand for optimal performance, eq.(5), we find that the optimal design makes the stress uniform along the beam and as small as possible. The optimal profile is a linear taper:

$$\widehat{T}_\theta(x) = \frac{2\theta(L-x)}{L^2} \quad [\mathrm{m}] \tag{7}$$

The performance obtained by this design is:

$$R(\widehat{T}_\theta) = \frac{3\widetilde{\phi}L^4}{4w\theta^2} \quad [\mathrm{Pa}] \tag{8}$$

To understand eq.(7) we note from eq.(6) that the linear taper is the only thickness profile which achieves the same maximum stress at all sections along the beam. From eq.(6) we know that we could reduce the stress in some regions of the beam by increasing the thickness profile in those regions. However, the mass-constraint, eq.(3), would force a lower thickness elsewhere, and in those other regions the stress would be augmented. Since the performance requirement is to minimize the maximal stress along the beam, the uniform stress profile is the stress-minimizing solution at this beam mass. Eq.(8), the performance obtained by this optimal design, is the value of the stress in eq.(6) with the optimal taper of eq.(7).

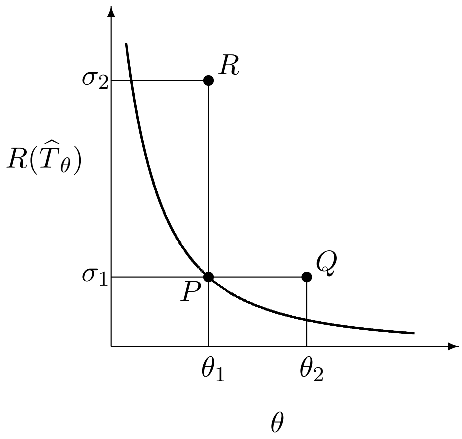

Figure 2: Optimal (min-max) stress $R(\widehat{T}_\theta)$ vs. mass-parameter θ, eq.(8).

We may think of eq.(8) as a curve, $R(\widehat{T}_\theta)$-vs.-θ, representing the trade-off between minimal stress and minimal mass, as shown in fig. 2. As the beam mass, θ, is reduced, the least possible maximum stress, $R(\widehat{T}_\theta)$, increases. Every point along this curve is optimal in the Pareto sense that either of the design criteria—mass or stress minimization—can be improved only by detracting from the other criterion.

Consider a point P on this optimal design curve, corresponding to the min-max stress σ_1 of a beam of mass θ_1. That is: $\sigma_1 = R(\widehat{T}_{\theta_1})$. Let Q be a point to the right of P. Q represents beams whose mass is $\theta_2 > \theta_1$ and whose min-max stress is still only σ_1. These beams are sub-optimal: at this min-max stress they have excessive mass. Alternatively, consider the point R lying above P, which represents beams of mass θ_1 whose min-max stress exceeds the optimum for this mass: $\sigma_2 > R(\widehat{T}_{\theta_1})$. This again is sub-optimal. We can interpret any sub-optimal beam as either mass-excessive for its min-max stress, or stress-excessive for its mass. Finally, because the curve is a Pareto frontier, there are no beams corresponding to points below the curve.

2.2 Robustness to Uncertain Load

Now we depart from the performance-optimization analysis described above. Any designer wants better performance rather than worse, but aspirations are tempered by the need for feasibility, the need for reliable design. We will now consider the very common situation in which the load profile is uncertain and, in response, we will develop a **robust satisficing** design strategy. We will be particularly interested in the relation between the optimal design under this strategy, and the performance-maximizing design described in section 2.1.

The designer will choose a thickness profile, $T(x)$, which **satisfices** the aspiration for good performance, that is, which attempts to guarantee that the maximum stress is no greater than a specified level, for a given beam mass. Satisficing is not an optimization, so an element of design freedom still remains. The designer will use this degree of freedom to maximize the immunity to error in the design-base load. The design specification is *satisficed,* while the robustness to failure is *maximized.*

In order to implement this we will first define an info-gap model of uncertainty, and then define the robustness function.

Let $\phi(x)$ [N/m] represent the unknown actual load-density profile, and let $\widetilde{\phi}(x)$ denote the designer's best estimate of $\phi(x)$. For instance, the nominal estimate may be the constant load density used in section 2.1. Let $\mathcal{U}(\alpha, \widetilde{\phi})$ be a set of load profiles $\phi(x)$, containing the nominal estimate $\widetilde{\phi}(x)$. An **info-gap model** for the designer's uncertainty about $\phi(x)$ is a *family of nested sets* $\mathcal{U}(\alpha, \widetilde{\phi})$, $\alpha \geq 0$. As α grows, the sets become more inclusive:

$$\alpha \leq \alpha' \quad \text{implies} \quad \mathcal{U}(\alpha, \widetilde{\phi}) \ \subseteq \ \mathcal{U}(\alpha', \widetilde{\phi}) \tag{9}$$

Also, the nominal load belongs to all the sets in the family:

$$\widetilde{\phi}(x) \ \in \ \mathcal{U}(\alpha, \widetilde{\phi}) \quad \text{for all} \quad \alpha \geq 0 \tag{10}$$

The nesting of the uncertainty-sets imbues α with its meaning as a **horizon of uncertainty**. A large α entails great variability of the potential load profiles $\phi(x)$ around the nominal estimate $\widetilde{\phi}(x)$. Since α is unbounded ($\alpha \geq 0$), the family of uncertainty sets is likewise unbounded. This means that we cannot identify a worse case, and the subsequent analysis is *not* a worst-case analysis in the ordinary sense, and does *not* entail a min-max as in eq.(5). We will see an example of an info-gap model of uncertainty shortly. (Info-gap models may obey additional axioms as well [4].)

Now we define the **info-gap robustness function**. The designer's aspiration (or requirement) for performance is that the bending stress not exceed the critical value $\sigma_{\rm c}$ anywhere along the beam of specified mass. That is, the condition $\sigma(x) \leq \sigma_{\rm c}$ is needed for 'survival'; better performance ($\sigma(x) \ll \sigma_{\rm c}$) is desirable but is not a design requirement. The designer will choose the critical stress $\sigma_{\rm c}$ as small as necessary, but no smaller than needed. The designer attempts to satisfy the design specification with the choice of the thickness profile $T(x)$ from the set $\Theta(\theta)$ in eq.(3) but, since the actual load profile $\phi(x)$ is unknown when $T(x)$ is chosen, the maximum bending stress is also unknown. The **robustness** of thickness profile $T(x)$ is the greatest horizon of uncertainty, α, at which the maximum stress is guaranteed to be no greater than the design requirement:

$$\widehat{\alpha}(T, \sigma_{\rm c}) = \max \left\{ \alpha : \max_{\phi \in \mathcal{U}(\alpha, \widetilde{\phi})} \rho(T, \phi) \leq \sigma_{\rm c} \right\} \tag{11}$$

where:

$$\rho(T, \phi) = \max_{0 \leq x \leq L} |\sigma_{\phi, T}(x)| \tag{12}$$

which is the analog of eq.(4) for the current case of unknown load profile. $\sigma_{\phi,T}(x)$ denotes the maximum stress in the beam section at x, given load profile $\phi(x)$ and thickness profile $T(x)$.

We can 'read' eq.(11) from left to right: The robustness $\widehat{\alpha}(T, \sigma_{\rm c})$ of thickness profile $T(x)$, given design specification (or aspiration) $\sigma_{\rm c}$, is the maximum horizon of uncertainty α such that the worst performance $\rho(T(x), \phi(x))$, for any realization $\phi(x)$ of the actual load profile up to α, is no greater than $\sigma_{\rm c}$. This is a worst-case-up-to-α analysis, but since α is unknown, what we are doing is determining the greatest α which does not allow failure.

More robustness to failure is better than less, provided the design requirements are satisfied. An info-gap robust-optimal design is an allowed thickness profile, $\widehat{T}_\theta(x) \in \Theta(\theta)$, which maximizes the robustness while also satisficing the performance:

$$\widehat{\alpha}(\widehat{T}_\theta, \sigma_{\rm c}) = \max_{T \in \Theta(\theta)} \widehat{\alpha}(T, \sigma_{\rm c}) \tag{13}$$

This is the info-gap analog of the optimal design criterion in eq.(5). Note that, unlike eq.(5), we are not minimizing the maximum stress. Rather, we are maximizing the robustness to uncertainty; the stress-requirement is satisficed to σ_c by the robustness function $\widehat{\alpha}(T, \sigma_c)$.

Let's consider a concrete example. Suppose that the known nominal load density is the constant non-negative value $\widetilde{\phi}$, and that we are also aware that the actual load profile $\phi(x)$ may deviate from $\widetilde{\phi}$, but we have no information about this deviation. One representation of this load-uncertainty is the envelope-bound info-gap model, which is the following family of nested sets of load profiles:

$$\mathcal{U}(\alpha, \widetilde{\phi}) = \left\{ \phi(x) : \ \left| \phi(x) - \widetilde{\phi} \right| \le \alpha \right\}, \quad \alpha \ge 0 \tag{14}$$

$\mathcal{U}(\alpha, \widetilde{\phi})$ is the set of load profiles whose deviation from the nominal profile $\widetilde{\phi}$ is bounded by α. Since the horizon of uncertainty, α, is unbounded, we have an unbounded family of nested sets of load profiles. Note that $\mathcal{U}(\alpha, \widetilde{\phi})$ satisfies the nesting and inclusion properties of eqs.(9) and (10).

The maximum absolute stress in the beam section at x, $|\sigma_{\phi,T}(x)|$, given load profile $\phi(x)$ and thickness profile $T(x)$, is found to be:

$$|\sigma_{\phi,T}(x)| = \frac{6}{wT^2(x)} \int_x^L (v - x)\phi(v) \, \mathrm{d}v \tag{15}$$

Employing this relation with eq.(12) in eq.(11) yields, after some manipulation, the following expression for the robustness of thickness profile $T(x)$ with design specification σ_c:

$$\widehat{\alpha}(T, \sigma_c) = \frac{w\sigma_c/3}{\displaystyle\max_{0 \le x \le L} \left(\frac{L - x}{T(x)} \right)^2} - \widetilde{\phi} \tag{16}$$

provided that this expression is positive. A negative value arises if the maximum stress in response to the nominal load exceeds the design requirement, σ_c. A negative value means that, even without uncertainty, the design requirement cannot be achieved. In this case the robustness to uncertainty vanishes and we define $\widehat{\alpha}(T, \sigma_c) = 0$.

2.3 Info-gap Robust-optimal Design: Clash with Performance-optimal Design

We now discuss the robustness-maximizing design, $\widehat{T}_\theta(x)$ in eq.(13). We will find the beam-shape which maximizes the robustness, for any choice of the stress-requirement σ_c. Significantly, this beam-shape will be the linear taper which maximizes the performance. However, we will find that when σ_c is chosen on the Pareto-optimal curve, the robustness of this linear taper is precisely zero. That is: performance-optimization entails robustness-minimization. This motivates the choice of performance-sub-optimal designs as the only way to obtain positive robustness to uncertainty.

Consider beams of mass θ whose maximum bending stress is no greater than σ_c. We are considering *any* σ_c, so (θ, σ_c) is not necessarily Pareto-optimal and does not necessarily fall on the optimal-design curve of fig. 2. From examination of eq.(16), we find that the thickness profile in the allowed-mass set $\Theta(\theta)$ which maximizes the robustness and satisfices the stress (at stress requirement σ_c) is precisely the profile which maximizes the performance: the linear taper $\widehat{T}_\theta(x)$ in eq.(7). With this robustness-maximizing thickness profile, the robustness in eq.(16) becomes:

$$\widehat{\alpha}(\widehat{T}_\theta, \sigma_c) = \frac{4w\theta^2\sigma_c}{3L^4} - \widetilde{\phi} \tag{17}$$

Fig. 3 illustrates this optimal robustness versus the maximum-stress design requirement, σ_c.

Fig. 3 demonstrates one of the most important universal properties of the robustness function: robustness decreases monotonically as the performance-requirement becomes more stringent. A small

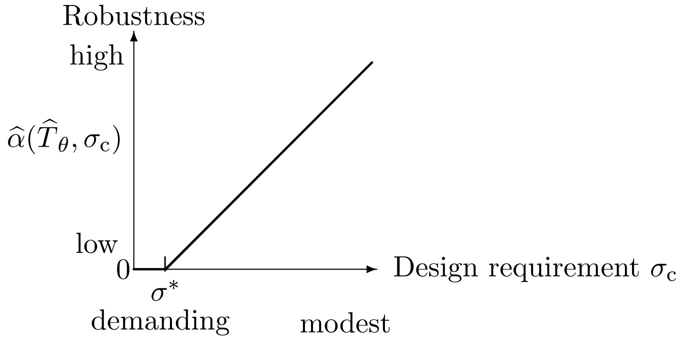

Figure 3: Optimal robustness curve, $\widehat{\alpha}(\widehat{T}_\theta, \sigma_c)$, versus the maximum-stress design requirement σ_c, eq.(17).

value of σ_c is a demanding specification, while a large value of σ_c is more lenient. A modest stress requirement will be quite robust, while a demanding design will be prone to failure. The value of σ_c at which the robustness becomes zero is denoted in fig. 3 by σ^*. This is such an exacting requirement that even infinitesimal deviations of the actual load profile from the nominal profile may entail violation of the design requirement. Clearly, choosing the requirement $\sigma_c = \sigma^*$ is infeasible and unrealistic since σ_c is defined as a stress level which must not be exceeded.

The value of σ^* is obtained by equating $\widehat{\alpha}(\widehat{T}_\theta, \sigma_c)$, in eq.(17), to zero and solving for σ_c. The result is:

$$\sigma^* = \frac{3\widetilde{\phi}L^4}{4w\theta^2} \ \text{[Pa]} \tag{18}$$

which is precisely the minimal stress obtained by the performance-optimizing design, eq.(8). We see here an instance of another general phenomenon of great importance: a design which **optimizes the performance** (as in section 2.1), also **minimizes the robustness**. That is, the locus of (θ, σ) values on the optimal-design curve of fig. 2 coincides with the zero-robustness points (σ^*, 0) in fig. 3. A point with positive robustness on the curve in fig. 3 ($\widehat{\alpha}(\widehat{T}_\theta, \sigma_2) > 0$) corresponds to a point *above* the Pareto-optimal curve like R on fig. 2 for which $\sigma_2 > \sigma(\theta)$. Performance optimization leads to the least feasible of all realizable designs. The designer is therefore strongly motivated to **satisfice the performance and maximize the robustness,** as we have done in this section. We explore this further in the next subsection.

2.4 Resolving the Clash

We will seek a thickness profile, $T^*(x)$, which has two properties:

1. The design has positive robustness to load-uncertainty, so that $\widehat{\alpha}(T^*, \sigma_c)$ in eq.(16) is positive.
2. The design has sub-optimal performance, so it is not on the optimal trade-off curve between mass and stress, eq.(8). This is necessary in order to enable positive robustness.

If the beam mass is constrained to $\Theta(\theta)$, then the min-max stress is given by eq.(8). Let us adopt this value, σ^* in (18), as the design requirement. We know from our analysis in section 2.3 that we must accept a beam-mass in excess of θ in order to satisfice this stress requirement with positive robustness. That is, following our discussion of eq.(8) and fig. 2, we must choose a design point to the right of the optimum-performance design curve, such as point Q in fig. 2.

From eq.(16), the condition for positive robustness is:

$$\frac{w\sigma^*/3}{\max_{0\le x\le L}\left(\frac{L-x}{T(x)}\right)^2} \;>\; \widetilde{\phi} \tag{19}$$

When $T(x)$ is the performance-optimal linear taper, $\widehat{T}_\theta(x)$ of eq.(7), and σ^* is given by eq.(18), we obtain equality in (19) and hence zero robustness, so we must choose the thickness profile so that:

$$T(x) \ge \frac{2\theta(L-x)}{L^2} = \widehat{T}_\theta(x) \tag{20}$$

with strict inequality over at least part of the beam. There are many available solutions; we consider one simple class of solutions:

$$T^*(x) = \widehat{T}_\theta(x) + \gamma \tag{21}$$

where $\gamma \ge 0$. From eq.(16), the robustness of this profile is:

$$\widehat{\alpha}(T^*, \sigma^*) = \frac{w\sigma^*}{3L^2}\left(\frac{2\theta}{L}+\gamma\right)^2 - \widetilde{\phi} \tag{22}$$

With σ^* from eq.(18), this robustness becomes:

$$\widehat{\alpha}(T^*, \sigma^*) = \frac{\widetilde{\phi}L^2}{4\theta^2}\left(\frac{2\theta}{L}+\gamma\right)^2 - \widetilde{\phi} \tag{23}$$

whose positivity is controlled by γ which also controls the extent of deviation of T^* from the performance-optimum solution $\widehat{T}_\theta$. When γ is zero, the beam shape lies on the performance-optimal curve, eq.(8): it has minimal mass for the stress-requirement σ_c. However, at $\gamma = 0$ the robustness to load-uncertainty is zero. As γ becomes larger, the robustness increases but the beam also becomes more mass-excessive.

In eqs.(20)–(23) we have derived a performance-sub-optimal design which has positive robustness. This design is formulated as a point Q to the right of the performance-optimal curve in fig. 2: we have increased the mass while holding the stress-requirement fixed. Another approach to defining designs which are performance-sub-optimal and yet have positive robustness is to seek points R which are above the curve. Positive robustness is obtained if inequality eq.(19) is satisfied. This can be obtained with the linear beam shape of eq.(7), and with a stress-requirement σ_c in excess of the minimal (optimal) stress for this beam mass given by σ^* in eq.(18) (or equivalently by $R(\widehat{T}_\theta)$ in eq.(8)). One way to understand this alternative approach is in terms of the relation between points P and R in fig. 2. The mass-optimal linear taper is used, eq.(7) with θ_1, but the aspiration for stress-performance is weakened: rather than adopting the optimal-stress requirement (σ_1 in fig. 2), the designer adopts a less demanding stress requirement (σ_2 in fig. 2).

The added beam thickness, γ in eq.(21), can be thought of as a safety factor.[1] The designer who proceeded according to the performance-optimization procedure of section 2.1 may add the thickness γ as an ad hoc protection. However, eq.(23) enables one to evaluate this deviation from the optimum-performance design in terms of the robustness-to-uncertainty which it entails. The robustness, $\widehat{\alpha}(T^*, \sigma^*)$ in eq.(23), is the greatest value of the uncertainty parameter, α in the info-gap model of eq.(14), which does not allow the maximum bending stress to exceed $\sigma^*(\theta)$ of eq.(18). $\widehat{\alpha}$ is the interval of load-amplitude within which the actual load profile, $\phi(x)$, may deviate from the design-base nominal load profile, $\widetilde{\phi}$, without exceeding the stress requirement. Suppose the designer

[1]The author is indebted to Prof. Eli Altus, of the Technion, for suggesting this interpretation.

desires that the robustness, $\widehat{\alpha}$, be equal to a fraction f of the nominal load: $\widehat{\alpha} = f\widetilde{\phi}$. The thickness safety factor which is needed is obtained by inverting eq.(23) to obtain:

$$\gamma = \frac{2\theta}{L}\left(\sqrt{1+f} - 1\right) \tag{24}$$

$f = 0$ means that no robustness is needed, and this causes $\gamma = 0$, meaning that the optimal-performance design, $\widehat{T}_\theta$, is obtained. f is increased to represent greater demanded robustness to uncertainty, causing the safety factor, γ, to increase from zero.

2.5 Opportunity from Uncertain Load

In section 2.2 we defined the robustness function: $\widehat{\alpha}(T, \sigma_{\rm c})$ is the greatest horizon of uncertainty which design $T(x)$ can tolerate without failure, when $\sigma_{\rm c}$ is the maximum allowed stress. The robustness function addresses the adverse aspect of load uncertainty. It evaluates the immunity to failure, which is why a large value of robustness is preferred to a small value, when the stress-limit is fixed.

In this section we explore the idea that uncertainty may be propitious: unknown contingencies may be favorable. The 'opportunity function' which we will formulate is also an immunity function: it assesses the immunity against highly desirable windfall outcomes. Since the opportunity function is the immunity against sweeping success, a small value is preferred over a large value. Each immunity function—robustness and opportunity—generates its own preference ranking on the set of available designs. We will see that, in general, these rankings may or may not agree.

As before, $\sigma_{\rm c}$ is the greatest acceptable bending stress. Let $\sigma_{\rm w}$ be a smaller stress which, if not exceeded at any point along the beam, would be a desirable 'windfall' outcome. It is not necessary that the stress be as small of $\sigma_{\rm w}$, but this would be viewed very favorably. The nominal load produces a maximum bending stress which is greater than $\sigma_{\rm w}$. However, favorable fluctuations of the load could produce a maximum bending stress as low as $\sigma_{\rm w}$. The **opportunity function** is the lowest horizon of uncertainty at which the maximum bending stress at any section of the beam can be as low as $\sigma_{\rm w}$:

$$\widehat{\beta}(T, \sigma_{\rm w}) = \min\left\{\alpha : \min_{\phi \in \mathcal{U}(\alpha, \widetilde{\phi})} \rho(T, \phi) \le \sigma_{\rm w}\right\} \tag{25}$$

where $\rho(T, \phi)$ is the maximum absolute bending stress occurring in the beam, specified in eq.(12). $\widehat{\beta}(T, \sigma_{\rm w})$ is the lowest horizon of uncertainty which must be accepted in order to enable maximum stress as low as $\sigma_{\rm w}$. $\widehat{\beta}(T, \sigma_{\rm w})$ is the immunity to windfall: a small value implies that windfall performance is possible (though not guaranteed) even at low level of uncertainty. The opportunity function is the dual of the robustness function in eq.(11).

Employing eq.(12), we can write the opportunity function more explicitly as:

$$\widehat{\beta}(T, \sigma_{\rm w}) = \min\left\{\alpha : \min_{\phi \in \mathcal{U}(\alpha, \widetilde{\phi})} \max_{0 \le x \le L} |\sigma_{\phi, T}(x)| \le \sigma_{\rm w}\right\} \tag{26}$$

The evaluation of the opportunity function requires a bit of caution because, in general, the order of the inner 'min' and 'max' operators can not be reversed. In the current example, however, a simplification occurs.

Let us define a constant load profile, $\phi^* = \widetilde{\phi} - \alpha$, which belongs to $\mathcal{U}(\alpha, \widetilde{\phi})$ for all $\alpha \ge 0$. One can readily show that this load profile minimizes the maximum stress at all sections, x. That is:

$$|\sigma_{\phi^*, T}(x)| = \min_{\phi \in \mathcal{U}(\alpha, \widetilde{\phi})} |\sigma_{\phi, T}(x)| \tag{27}$$

Since this minimizing load profile, ϕ^*, is the same for all positions x, we can reverse the order of the operators in eq.(26) as:

$$\min_{\phi\in\mathcal{U}(\alpha,\widetilde{\phi})} \max_{0\le x\le L} |\sigma_{\phi,T}(x)| = \max_{0\le x\le L} \min_{\phi\in\mathcal{U}(\alpha,\widetilde{\phi})} |\sigma_{\phi,T}(x)| \tag{28}$$

$$= \max_{0\le x\le L} |\sigma_{\phi^*,T}(x)| \tag{29}$$

$$= \max_{0\le x\le L} \frac{3(\widetilde{\phi}-\alpha)(L-x)^2}{wT^2(x)} \tag{30}$$

The opportunity function is the smallest horizon of uncertainty, α, for which the righthand side of eq.(30) is no greater than $\sigma_{\rm w}$. Equating this expression to $\sigma_{\rm w}$ and solving for α yields the opportunity for design $T(x)$ with windfall aspiration $\sigma_{\rm w}$:

$$\widehat{\beta}(T,\sigma_{\rm w}) = \widetilde{\phi} - \frac{w\sigma_{\rm w}/3}{\displaystyle\max_{0\le x\le L}\left(\frac{L-x}{T(x)}\right)^2} \tag{31}$$

This expression is non-negative unless the righthand side of eq.(30) is less than $\sigma_{\rm w}$ at $\alpha=0$, which occurs if and only if the nominal load entails maximal stress less than $\sigma_{\rm w}$; in this case we define $\widehat{\beta}(T,\sigma_{\rm w})=0$.

We have already mentioned that each immunity function—robustness and opportunity—generates its own preference ranking of available designs. Since "bigger is better" for the robustness function, we will prefer T over T' if the former design is more robust than the latter. Concisely:

$$T \succ_{\rm r} T' \quad \text{if} \quad \widehat{\alpha}(T,\sigma_{\rm c}) > \widehat{\alpha}(T',\sigma_{\rm c}) \tag{32}$$

The opportunity function is the immunity against windfall performance, so "big is bad". This means that we will prefer T over T' if the former design is more opportune than the latter:

$$T \succ_{\rm o} T' \quad \text{if} \quad \widehat{\beta}(T,\sigma_{\rm w}) < \widehat{\beta}(T',\sigma_{\rm w}) \tag{33}$$

The generic definitions of the immunity functions do not imply that the preference-rankings in eqs.(32) and (33) agree. The immunity functions are said to be **sympathetic** when their preference rankings agree; they are **antagonistic** otherwise. Both situations are possible. In the present example the immunities are sympathetic, as we see by combining eqs.(16) and (31) as:

$$\widehat{\beta}(T,\sigma_{\rm w}) = -\underbrace{\frac{\sigma_{\rm w}}{\sigma_{\rm c}}\widehat{\alpha}(T,\sigma_{\rm c})}_{A} + \underbrace{\left(1-\frac{\sigma_{\rm w}}{\sigma_{\rm c}}\right)\widetilde{\phi}}_{B} \tag{34}$$

Expression 'B' does not depend upon the design, $T(x)$, and expression 'A' is non-negative. Consequently, any change in the design which causes $\widehat{\alpha}$ to increase (that is, robustness improves), causes $\widehat{\beta}$ to decrease (which improves opportunity). Likewise, robustness and opportunity deteriorate together. These immunity functions are sympathetic for any possible design change, though they do not necessarily improve at the same rate; marginal changes may be greater for one than for the other.

In general, robustness and opportunity functions are not necessarily sympathetic. Their sympathy in the current example is guaranteed because B in eq.(34) is independent of the design. This need not be the case. If B increases due to a design-change which causes $\widehat{\alpha}$ to increase, the net effect may be an increase in $\widehat{\beta}$, which constitutes a decrease in opportunity. For an example see [4, p.52].

3 Maneuvering a Vibrating System with Uncertain Dynamics

In section 2 we considered the design and reliability analysis of a static system subject to uncertain loads. We now consider the analysis and control of a simple vibrating system whose dynamic equations are uncertain. That is, the best model is known to be wrong or incomplete in some poorly understood way, and an info-gap model represents the uncertainty in this system-model. Despite the uncertainty in the system-model, the designer must choose a driving function which efficiently "propels" the system as far as possible.

3.1 Model Uncertainty

Consider a one-dimensional linear system whose displacement $x(t)$ resulting from forcing function $q(t)$ is described by Duhamel's relation:

$$x(t;q,h) = \int_0^t q(\tau)h(t-\tau)\,\mathrm{d}\tau \tag{35}$$

where $h(t)$ is the impulse response function (IRF).

The best available model for the IRF is denoted $\widetilde{h}(t)$, which may differ substantially from $h(t)$ due to incomplete or inaccurate representation of pertinent mechanisms. For instance, for the undamped linear harmonic oscillator:

$$\widetilde{h}(t) = \frac{1}{m\omega}\sin\omega t \tag{36}$$

where m is the mass and ω is the natural frequency. This IRF is seriously deficient in the presence of damping, which is a complicated and incompletely understood phenomenon.

Let $\mathcal{U}(\alpha,\widetilde{h})$ be an info-gap model for uncertainty in the IRF. That is, $\mathcal{U}(\alpha,\widetilde{h})$, $\alpha \ge 0$, is a family of nested sets of IRFs, all containing the nominal best-model, $\widetilde{h}(t)$. That is:

$$\alpha < \alpha' \quad \text{implies} \quad \mathcal{U}(\alpha,\widetilde{h}) \;\subset\; \mathcal{U}(\alpha',\widetilde{h}) \tag{37}$$

and

$$\widetilde{h}(t) \in \mathcal{U}(\alpha,\widetilde{h}) \quad \text{for all} \quad \alpha \ge 0 \tag{38}$$

As an example we now construct a Fourier ellipsoid-bound info-gap model of uncertainty in the system dynamics. Actual IRFs are related to the nominal function by:

$$h(t) = \widetilde{h}(t) + \sum_i c_i\sigma_i(t) \tag{39}$$

where the $\sigma_i(t)$ are known expansion functions (e.g. cosines, sines, polynomials, etc.) and the c_i are unknown expansion coefficients. Let c and $\sigma(t)$ denote the vectors of expansion coefficients and expansion functions, respectively, so that eq.(39) becomes:

$$h(t) = \widetilde{h}(t) + c^T\sigma(t) \tag{40}$$

A Fourier ellipsoid-bound info-gap model for uncertainty in the IRF is a family of nested ellipsoids of coefficient vectors:

$$\mathcal{U}(\alpha,\widetilde{h}) = \left\{ h(t) = \widetilde{h}(t) + c^T\sigma(t) :\; c^TVc \le \alpha^2 \right\}, \quad \alpha \ge 0 \tag{41}$$

where V is a known, real, symmetric, positive definite matrix which determines the shape of the ellipsoids of c-vectors. V is based on fragmentary information about the dispersion of the expansion coefficients. The size of each ellipsoid is determined by the (unknown) horizon-of-uncertainty parameter α.

3.2 Performance Optimization with the Best Model

We now consider the performance-optimal design of the driving function $q(t)$ based on the model $\widetilde{h}(t)$ of eq.(36) which is, for the purpose of this example, the best-known IRF of the system. The goal of the design is to choose the forcing function $q(t)$ to achieve large displacement $x(T; q, \widetilde{h})$ at specified time T with low control effort $\int_0^T q^2(t)\,\mathrm{d}t$. Specifically, we would like to select $q(t)$ so as to achieve an optimal balance between the following two conflicting objectives:

$$\max_{q(t)} x(T; q, \widetilde{h}) \tag{42}$$

$$\min_{q(t)} \int_0^T q^2(t)\,\mathrm{d}t \tag{43}$$

Let $\mathcal{Q}(E)$ denote the set of all control functions $q(t)$ whose control effort equals E:

$$\mathcal{Q}(E) = \left\{ q(t):\ E = \int_0^T q^2(t)\,\mathrm{d}t \right\} \tag{44}$$

Using the Schwarz inequality, one can readily show that the q-function in $\mathcal{Q}(E)$ which maximizes the displacement $x(T; q, \widetilde{h})$ at time T is:

$$q_E^*(t) = \frac{\sqrt{f_0 E}}{m\omega} \sin \omega(T - t) \tag{45}$$

where:

$$f_0 = \frac{4m^2\omega^3}{2\omega T - \sin 2\omega T} \tag{46}$$

From this one finds that the greatest displacement at time T, obtainable with any control function in $\mathcal{Q}(E)$, is:

$$x(T; q_E^*, \widetilde{h}) = \sqrt{\frac{E}{f_0}} \tag{47}$$

Eq.(47) expresses the trade-off between control effort E and maximal displacement x: large displacement is obtained only at the expense of large effort, as shown in fig. 4. Like fig. 2, this relationship expresses the Pareto-optimal design options: any improvement in control effort (making E smaller) is obtained only by relinquishing displacement (making x smaller).

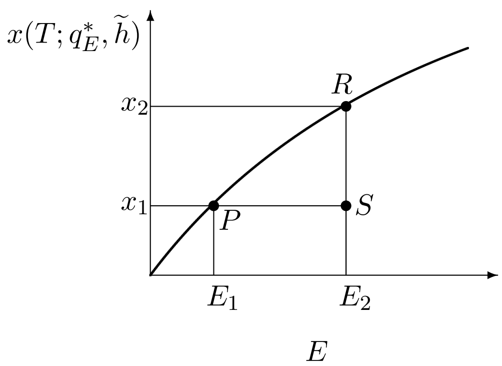

Figure 4: Maximal displacement $x(T; q_E^*, \widetilde{h})$ vs. control effort E, eq.(47).

Points above the curve in fig. 4 are inaccessible: no design can realize those (E, x) combinations. Points on the curve are Pareto-optimal and points below the curve are sub-optimal designs. For instance, point P on the curve is Pareto-optimal: E_1 is the lowest control effort which can achieve displacement as large as x_1. Point S is sub-optimal and represents excessive control effort ($E_2 > E_1$) to achieve displacement x_1. Likewise, point R is Pareto-optimal: x_2 is the greatest displacement which can be attained with control effort E_2. So again S is sub-optimal: greater displacement ($x_2 > x_1$) could be achieved with effort E_2.

3.3 Robustness Function

We now develop an expression for the robustness, of the displacement $x(T;q,h)$, to uncertainty in the system dynamics $h(t)$.

The first design goal, eq.(42), implies that a large value of displacement is needed. The second design goal, eq.(43), conflicts with the first and calls for small control effort. In section 3.2 we found that performance optimization leads to a Pareto trade-off between these two criteria, expressed in eq.(47) and fig. 4. In this section, in light of the uncertainty in the IRF, we take a different approach. For any given control function $q(t)$, we "satisfice" the displacement by requiring that the displacement be at least as large as some specified and satisfactory value $x_{\rm c}$. Since $x(T;q,h)$ depends upon the unknown IRF we cannot guarantee that the displacement will be satisfactory. However, we can answer the following question: for given forcing function $q(t)$, by how much can the best model $\widetilde{h}(t)$ err without jeopardizing the achievement of adequate displacement? More specifically, given $q(t)$, what is the greatest horizon of uncertainty, α, up to which every model $h(t)$ causes the displacement to be at least as large as $x_{\rm c}$? The answer to this question is the robustness function:

$$\widehat{\alpha}(q, x_{\rm c}) = \max \left\{ \alpha : \min_{h \in \mathcal{U}(\alpha, \widetilde{h})} x(T;q,h) \ge x_{\rm c} \right\} \tag{48}$$

We can "read" this relation from left to right: the robustness $\widehat{\alpha}(q, x_{\rm c})$ of control function $q(t)$ with displacement-aspiration $x_{\rm c}$ is the greatest horizon of uncertainty α such that every system-model $h(t)$ in $\mathcal{U}(\alpha, \widetilde{h})$ causes the displacement $x(T;q,h)$ to be no less than $x_{\rm c}$. If $\widehat{\alpha}(q, x_{\rm c})$ is large then the system is robust to model-uncertainty and $q(t)$ can be relied upon to bring the system to at least $x_{\rm c}$ at time T. If $\widehat{\alpha}(q, x_{\rm c})$ is small then this driving function cannot be relied upon and the system is vulnerable to uncertainty in the dynamics.

Using Lagrange optimization one can readily show that the smallest displacement, for any system model $h(t)$ up to uncertainty α, is:

$$\min_{h \in \mathcal{U}(\alpha, \widetilde{h})} x(T;q,h) = x(T;q,\widetilde{h}) - \alpha \sqrt{b^T V b} \tag{49}$$

where we have defined the following vector:

$$b = \int_0^T q(t) \sigma(T-t) \, {\rm d}t \tag{50}$$

Eq.(49) asserts that the least displacement, up to uncertainty α, is the nominal, best-model, displacment $x(T;q,\widetilde{h})$, decremented by the uncertainty term $\alpha\sqrt{b^T V b}$. If the nominal displacement falls short of the demanded displacement $x_{\rm c}$, then uncertainty only makes things worse and the robustness to uncertainty is zero. If $x(T;q,\widetilde{h})$ exceeds $x_{\rm c}$ then the robustness is found by equating the righthand side of eq.(49) to $x_{\rm c}$ and solving for α. That is, the robustness of driving function $q(t)$ is:

$$\widehat{\alpha}(q, x_{\rm c}) = \left\{ \begin{array}{cl} 0 & \text{if } \ x(T;q,\widetilde{h}) \le x_{\rm c} \\ \dfrac{x(T;q,\widetilde{h}) - x_{\rm c}}{\sqrt{b^T V b}} & \text{else} \end{array} \right. \tag{51}$$

Eq.(51) documents the trade-off between robustness, $\widehat{\alpha}(q, x_{\rm c})$, and aspiration for performance, $x_{\rm c}$, as shown in fig. 5. A large and demanding value of $x_{\rm c}$ is accompanied by a low value of immunity to model-uncertainty, meaning that aspirations for large displacements are unreliable and infeasible. Modest requirements (small values of x_c) are more feasible since they have greater immunity to uncertainty. The value of x_c at which the robustness vanishes, x^* in the figure, is precisely the displacement predicted by the best model, $x(T;q,\widetilde{x})$. That is:

$$\widehat{\alpha}\left(q,\, x(T;q,\widetilde{h})\right) = 0 \tag{52}$$

This means that, for *any* driving function $q(t)$, the displacement predicted by the best available model $\widetilde{h}(t)$ cannot be relied upon to occur. Short-fall of the displacement may occur due to infinitesimally small error of the model. Since this is true for any $q(t)$ it is also true for the performance-optimum control function $q^*_E(t)$ in eq.(45):

$$\widehat{\alpha}\left(q^*_E,\, x(T;q^*_E,\widetilde{h})\right) = 0 \tag{53}$$

While $q^*_E(t)$ is, according to $\widetilde{h}(t)$, the most effective driving function of energy E, and while $x(T;q^*_E,\widetilde{h})$ is, again according to $\widetilde{h}(t)$, the resulting displacement, eq.(53) shows that this prediction has no immunity to modelling errors.

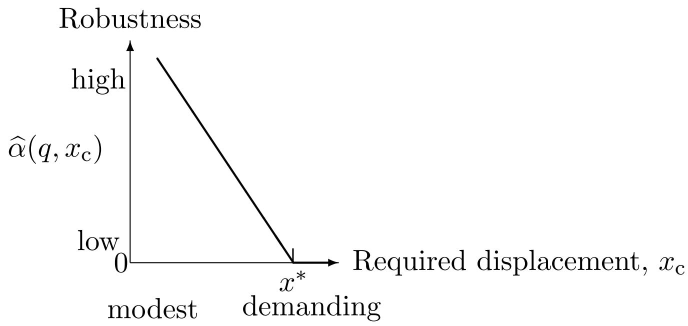

Figure 5: Robustness $\widehat{\alpha}(q,x_c)$ versus the demanded displacement x_c, eq.(51).

It is important to recognize that ordered pairs such as $\left(E,\ x(T;q^*_E,\widetilde{h})\right)$ correspond to points such as P and R on the Pareto-optimal design surface in fig. 4. That is, E and $x(T;q^*_E,\widetilde{h})$ are related by eq.(47). Hence eq.(53) shows that all of the performance-optimal designs on the Pareto surface have no immunity to errors in the design-base model of the system. These Pareto-efficient designs are not feasible or reliable predictions of the system performance.

The conclusion from eqs.(52) and (53) is likely to be that, since $x(T;q^*_E,\widetilde{h})$ cannot be relied upon to occur, one must moderate one's aspirations and accept a lower value of displacement. The designer might "travel" up and to the left on the robustness curve in fig. 5 until finding a value of $x_c < x^*$ at which the robustness is satisfactorily large. But then the question arises: what driving function $q(t)$ *maximizes* the robustness at this selected aspiration for displacement? The optimization studied in section 3.2 was optimization of *performance* (displacement and control effort). We now consider satisficing these quantities and optimizing the *robustness*. Specifically, for any displacement-aspiration x_c, the **robust-optimal** control function $\widehat{q}_E(t)$ of energy E maximizes the robustness function:

$$\widehat{\alpha}(\widehat{q}_E,x_c) = \max_{q(t)\in\mathcal{Q}(E)} \widehat{\alpha}(q,x_c) \tag{54}$$

This robust optimum may not always exist, or it may be inaccessible for practical reasons. In any case one will tend to prefer more robust over less robust solutions. More specifically, if $q_1(t)$ is more robust than $q_2(t)$, while satisficing the performance at the same level x_c, then $q_1(t)$ is preferred over $q_2(t)$:

$$q_1(t) \succ q_2(t) \quad \text{if} \quad \widehat{\alpha}(q_1,x_c) > \widehat{\alpha}(q_2,x_c) \tag{55}$$

Relating this to our earlier discussion, suppose that E is an accessible control effort and that $x_{\rm c}$ is a satisfactory level of performance. In order for $x_{\rm c}$ to be feasible it must be less than the best performance obtainable with effort E, namely $x_{\rm c} < x(t; q_E^*, \widetilde{h})$. This assures that the robustness of the performance-optimal control function, $q_E^*(t)$, will be positive: $\widehat{\alpha}(q_E^*, x_{\rm c}) > 0$. However, we might well ask if there is some other control function in $\mathcal{Q}(E)$ whose robustness is even greater. This will often be the case, as we illustrate in the next subsection.

3.4 Example

To keep things simple suppose that $\sigma(t)$ in the unknown part of the IRF in eq.(40) is a single linearly decreasing function:

$$\sigma(t) = \eta(T - t) \tag{56}$$

where η is a positive constant. Thus c is a scalar and the shape-matrix in the info-gap model of eq.(41) is simply $V = 1$.

The performance-optimal control function of effort E is $q_E^*(t)$ in eq.(45), which is a sine function at the natural frequency of the nominal IRF, $\widetilde{h}(t)$. From eq.(51), the robustness of this control function is:

$$\widehat{\alpha}(q_E^*, x_{\rm c}) = \frac{\dfrac{1}{m\omega}\displaystyle\int_0^T q_E^*(t)\sin\omega(T-t)\,{\rm d}t \;-\; x_{\rm c}}{\left|\eta\displaystyle\int_0^T q_E^*(t)(T-t)\,{\rm d}t\right|} \tag{57}$$

Similarly, the robustness of any arbitrary control function $q(t)$ is:

$$\widehat{\alpha}(q, x_{\rm c}) = \frac{\dfrac{1}{m\omega}\displaystyle\int_0^T q(t)\sin\omega(T-t)\,{\rm d}t \;-\; x_{\rm c}}{\left|\eta\displaystyle\int_0^T q(t)(T-t)\,{\rm d}t\right|} \tag{58}$$

(Presuming the numerator is positive.) In light of our discussion of eq.(55), we would like to find a control function $q(t)$ in $\mathcal{Q}(E)$ whose robustness is substantially greater than the robustness of $q_E^*(t)$.

We will illustrate that very substantial robustness benefits can be achieved by abandoning the performance-optimal function $q_E^*(t)$. We will not consider the general maximization of $\widehat{\alpha}(q, x_{\rm c})$, but only a parametric case. Consider functions of the form:

$$q_\mu(t) = A\sin\mu(T-t) \tag{59}$$

where $0 < \mu < \omega$ and A is chosen to guarantee that $q(t)$ belongs to $\mathcal{Q}(E)$ (which was defined in eq.(44)):

$$A = \sqrt{\frac{4\mu E}{2\mu T - \sin 2\mu T}} \tag{60}$$

For the special case that $\omega T = \pi$ the robustness functions of eqs.(57) and (58) become:

$$\widehat{\alpha}(q_E^*, x_{\rm c}) \;=\; \frac{\pi\sqrt{\omega E} - \sqrt{2\pi}\, m\omega^2 x_{\rm c}}{2m\eta\sqrt{\omega E}} \tag{61}$$

$$\widehat{\alpha}(q_\mu, x_{\rm c}) \;=\; \frac{A\mu^2\left(\dfrac{\sin[\pi(\mu-\omega)/\omega]}{2(\mu-\omega)} - \dfrac{\sin[\pi(\mu+\omega)/\omega]}{2(\mu+\omega)}\right) - m\omega\mu^2 x_{\rm c}}{\eta m A\left[\omega\sin(\pi\mu/\omega) - \pi\mu\cos(\pi\mu/\omega)\right]} \tag{62}$$

Fig. 6 shows the ratio of the robustnesses of the sub-optimal to the performance-optimal control functions, vs. the frequency of the control function. The robustness at control frequencies μ

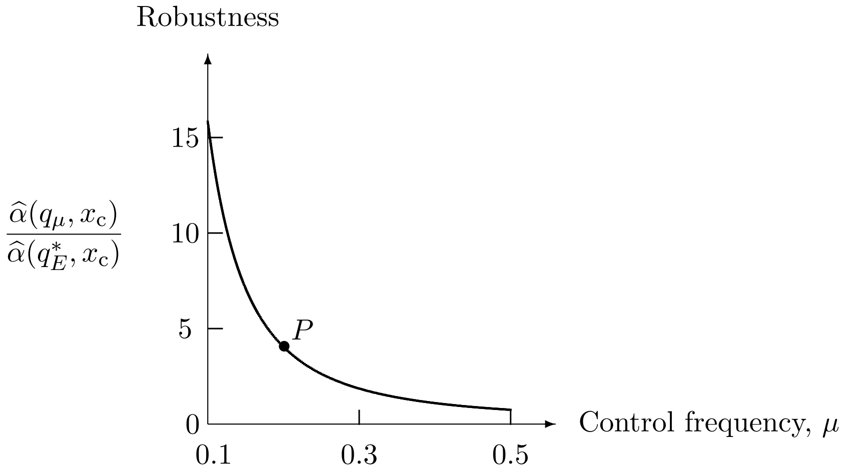

Figure 6: Ratio of the robustnesses of the sub-optimal to performance-optimal control functions, vs. frequency of control function. $x_c = 0.5$, $\omega = E = \eta = m = 1$. $\widehat{\alpha}(q_E^*, x_c) = 0.94$

much less than the nominal natural frequency ω, is substantially greater than the robustness of the performance-maximizing function. For instance, at point P, $\mu = 0.2$ and the robustness ratio is $\widehat{\alpha}(q_\mu, x_c)/\widehat{\alpha}(q_E^*, x_c) = 4.0$, meaning that $q_\mu(t)$ can tolerate a horizon of model-uncertainty 4 times greater than the uncertainty which is tolerable for $q_E^*(t)$, when satisficing the displacement at x_c. $q_E^*(t)$ maximizes the displacement according to the best IRF, $\widetilde{h}(t)$. However, this displacement-optimization of the control function leaves little residual immunity to uncertainty in the IRF at this value of x_c. A control function such as $q_\mu(t)$, for $\mu < \omega$, belongs to $\mathcal{Q}(E)$, as does $q_E^*(t)$, but $q_\mu(t)$ is sub-optimal with respect to displacement. That is, $x(T; q_\mu, \widetilde{h}) < x(T; q_E^*, \widetilde{h})$. However, because $q_\mu(t)$ is sub-optimal, there are many functions with control effort E which cause displacement as large as $x(T; q_\mu, \widetilde{h})$. In other words, there is additional design freedom with which to amplify the immunity to uncertainty. What fig. 6 shows is that large robustness-amplification can be achieved. This answers, by way of illustration, the question raised at the end of section 3.3.

4 System Identification

A common task encountered by engineering analysts is the up-dating of a system-model, based on measurements. The question we consider in this section is, given that the structure of the model is imperfect, what constitutes optimal estimation of the parameters? More precisely, is it sound procedure to maximize the fidelity between the model and the measurements, if the model structure is wrong (in unknown ways, of course)?

4.1 Optimal Identification

We begin by formulating a fairly typical framework for optimal identification of a model for predicting the behavior of a system. We then consider an example.

Let y_i be a vector of measurements of the system at time or state i, for $i = 1, \ldots, N$. Let $f_i(q)$ denote the model-prediction of the system in state i, which should match the measurements if the model is good. The vector q, containing real and linguistic variables, denotes the parameters and properties of the model which can be modified to bring the model into agreement with the measurements. We will denote the set of measurements by $Y = \{y_1, \ldots, y_N\}$ and the set of corresponding model-predictions by $F(q) = \{f_1(q), \ldots, f_N(q)\}$.

The overall performance of the predictor is assessed by a function $R[Y, F(q)]$. For example, this might be a mean-squared prediction error:

$$R[Y, F(q)] = \frac{1}{N} \sum_{i=1}^{N} \| f_i(q) - y_i \|^2 \tag{63}$$

A performance-optimal model, q^*, minimizes the performance-measure:

$$R[Y, F(q^*)] = \min_q R[Y, F(q)] \tag{64}$$

4.2 Uncertainty and Robustness

The model $f_i(q)$ is undoubtedly wrong, perhaps fundamentally flawed in its structure. There may be basic mechanisms which act on the system but which are not represented by $f_i(q)$. Let us denote more general models, some of which may be more correct, by:

$$\phi_i = f_i(q) + u_i \tag{65}$$

where u_i represents the unknown corrections to the original model, $f_i(q)$. We have very little knowledge about u_i; if we had knowledge of u_i we would most likely include it in $f_i(q)$. So, let us use an info-gap model of uncertainty to represent the unknown variation of possible models:

$$\phi_i \in \mathcal{U}(\alpha, f_i(q)), \quad \alpha \geq 0 \tag{66}$$

The centerpoint of the info-gap model, $f_i(q)$, is the known model, parameterized by q. The horizon of uncertainty, α, is unknown. This info-gap model is a family of nested sets of models. These sets of models become ever more inclusive as the horizon of uncertainty increases. That is:

$$\alpha \leq \alpha' \quad \text{implies} \quad \mathcal{U}(\alpha, f_i(q)) \subset \mathcal{U}(\alpha', f_i(q)) \tag{67}$$

In addition, the up-date model is included in all of the uncertainty sets:

$$f_i(q) \in \mathcal{U}(\alpha, f_i(q)), \quad \text{for all} \quad \alpha \geq 0 \tag{68}$$

As before, the model-prediction of the system output in state i is $f_i(q)$, and the set of model-predictions is denoted $F(q) = \{f_1(q), \ldots, f_N(q)\}$. More generally, the set of model-predictions with unknown terms $u_1, \ldots, u_N$ is denoted $F_u(q) = \{f_1(q) + u_1, \ldots, f_N(q) + u_N\}$.

We wish to choose a model, $f_i(q)$, for which the performance index, $R[Y, F_u(q)]$, is small. Let r_c represent an acceptably small value of this index. We would be willing, even delighted, if the prediction-error is smaller, but an error larger than r_c would be unacceptable.

The **robustness to model-uncertainty,** of model q with error-aspiration r_c, is the greatest horizon of uncertainty, α, within which all models provide prediction-error no greater than r_c:

$$\widehat{\alpha}(q, r_\text{c}) = \max \left\{ \alpha : \max_{\substack{\phi_i \in \mathcal{U}(\alpha, f_i(q)) \\ i=1, \ldots, N}} R[Y, F_u(q)] \leq r_\text{c} \right\} \tag{69}$$

When $\widehat{\alpha}(q, r_\text{c})$ is large, the model $f_i(q)$ may err fundamentally to a great degree, without jeopardizing the accuracy of its predictions; the model is robust to info-gaps in its formulation. When $\widehat{\alpha}(q, r_\text{c})$ is small, then even small errors in the model result in unacceptably large prediction errors.

Let q^* be an optimal model, which minimizes the prediction-error as defined in eq.(64), and let r_c^* be the corresponding optimal prediction error: $r_\text{c}^* = R(Y, F(q^*))$. By using model q^*, we can achieve prediction error as small as r_c^*, and no value of q can produce a model $f_i(q)$ which performs better. However, the robustness to model-uncertainty, of this optimal model, is zero:

$$\widehat{\alpha}(q^*, r_c^*) = 0 \tag{70}$$

This is a special case of the theorem to be discussed in section 6 that, by optimizing the performance, one minimizes the robustness to info-gaps. By optimizing the performance of the model-predictor, $f_i(q)$, we make this predictor maximally sensitive to errors in the basic formulation of the model.

In fact, eq.(70) is a special case of the following proposition. For any q, let $r_c = R[Y, F(q)]$ be the prediction-error of model $f_i(q)$. The preliminary lemma in section 6 shows that:

$$\widehat{\alpha}(q, r_c) = 0 \tag{71}$$

That is, the robustness of *any* model, $f_i(q)$, to uncertainty in the structure of that model, is precisely equal to zero, if the error-aspiration r_c equals the value of the performance function of that model. No model can be relied upon to perform at the level indicated by its performance function, if that model is subject to errors in its structure or formulation. $R[Y, F(q)]$ is an unrealistically optimistic assessment of model $f_i(q)$, unless we have reason to believe that no auxiliary uncertainties lurk in the mist of our ignorance.

4.3 Example

A simple example will illustrate the previous general discussion.

We begin by formulating a **mean-squared-error estimator** for a 1-dimensional linear model. The measurements y_i are scalars, and the model to be estimated is:

$$f_i(q) = iq \tag{72}$$

The performance function is the mean-squared error between model and measurements, eq.(63), which becomes:

$$R[Y, F(q)] = \frac{1}{N}\sum_{i=1}^{N}(iq - y_i)^2 \tag{73}$$

$$= \underbrace{\frac{1}{N}\sum_{i=1}^{N} y_i^2}_{\eta_2} - 2q \underbrace{\frac{1}{N}\sum_{i=1}^{N} iy_i}_{\eta_1} + q^2 \underbrace{\frac{1}{N}\sum_{i=1}^{N} i^2}_{\eta_0} \tag{74}$$

which defines the quantities η_0, η_1 and η_2. The performance-optimal model defined in eq.(64), which minimizes the mean-squared error, is:

$$q^* = \frac{\eta_1}{\eta_0} \tag{75}$$

Now we introduce **uncertainty in the model.** The model which is being estimated is linear in the 'time' or 'sequence' index i: $f_i = iq$. How robust is the performance of our estimator, to modification of the structure of this model? That is, how much can the model err in its basic structure, without jeopardizing its predictive power?

Suppose that the linear model of eq.(72) errs by lacking a quadratic term:

$$\phi_i = iq + i^2 u \tag{76}$$

where the value of u is unknown. The uncertainty in the quadratic model is represented by an interval-bound info-gap model, which is the following unbounded family of nested intervals:

$$\mathcal{U}(\alpha, iq) = \left\{ \phi_i = iq + i^2 u : \ |u| \le \alpha \right\}, \quad \alpha \ge 0 \tag{77}$$

The robustness of nominal model $f_i(q)$, with performance-aspiration r_c, is the greatest value of the horizon of uncertainty α at which the mean-squared error of the prediction is no greater than r_c for any model in $\mathcal{U}(\alpha, iq)$:

$$\widehat{\alpha}(q, r_c) = \max \left\{ \alpha : \max_{|u| \le \alpha} R[Y, F_u(q)] \le r_c \right\} \tag{78}$$

The mean-squared error of a model with non-linear term $i^2 u$ is:

$$R[Y, F_u(q)] = \frac{1}{N} \sum_{i=1}^{N} (iq + i^2 u - y_i)^2 \tag{79}$$

$$= \underbrace{\frac{1}{N} \sum_{i=1}^{N} (iq - y_i)^2}_{\xi_2} + 2u \underbrace{\frac{1}{N} \sum_{i=1}^{N} i^2 (iq - y_i)}_{\xi_1} + u^2 \underbrace{\frac{1}{N} \sum_{i=1}^{N} i^4}_{\xi_0} \tag{80}$$

which defines ξ_0, ξ_1 and ξ_2.

Some manipulations show that the maximum mean-squared error, for all quadratic models ϕ_i up to horizon of uncertainty α, is:

$$\max_{|u| \le \alpha} R[Y, F_u(q)] = \xi_2 + 2\alpha |\xi_1| + \alpha^2 \xi_0 \tag{81}$$

Referring to eq.(78), the robustness to an unknown quadratic non-linearity $i^2 u$, of the linear model $f_i(q)$, is the greatest value of α at which this maximum error is no greater than r_c.

First we note that the robustness is zero if r_c is small:

$$\widehat{\alpha}(q, r_c) = 0, \quad r_c \le \xi_2 \tag{82}$$

This is because, if $r_c \le \xi_2$, then $\max R$ in eq.(81) exceeds r_c for any positive value of α. One implication of eq.(82) is that some non-linear models have prediction errors in excess of ξ_2. If it is required that the fidelity between model and measurement be as good as or better than ξ_2, then no modelling errors of the quadratic type represented by the info-gap model of eq.(77) can be tolerated. Recall that ξ_2 is the mean-squared error of the nominal linear predictor, $f_i(q)$. Eq.(82) means that there is no robustness to model-uncertainty, if the performance-aspiration r_c is stricter or more exacting than the performance of the nominal, linear model.

For $r_c \ge \xi_2$, the robustness is obtained by equating the righthand side of eq.(81) to r_c and solving for α, resulting in:

$$\widehat{\alpha}(q, r_c) = \frac{|\xi_1|}{\xi_0} \left(-1 + \sqrt{1 + \frac{\xi_0 (r_c - \xi_2)}{\xi_1^2}} \right), \quad \xi_2 \le r_c \tag{83}$$

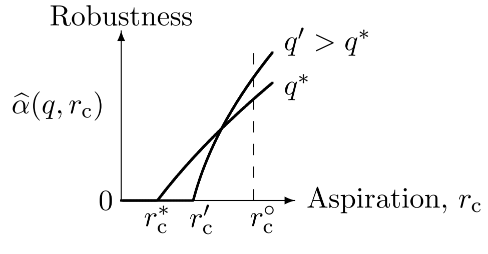

Figure 7: Robustness vs. prediction-aspiration, eq.(83). q fixed.

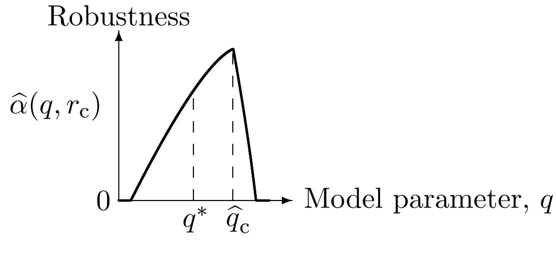

Figure 8: Robustness vs. model parameter, eq.(83). r_c fixed.

Relation (83) is plotted in fig. 7 for synthetic data[2] y_i and for two values of the model parameter q. The figure shows the robustness to model-uncertainty, $\widehat{\alpha}(q, r_{\rm c})$, against the aspiration for prediction error, $r_{\rm c}$. The robustness increases as greater error is tolerated. Two curves are shown, one for the optimal linear model, $q^* = 2.50$ in eq.(75), whose mean-squared error $r_{\rm c}^* = R[Y, F(q^*)] = 4.82$ is the lowest obtainable with any linear model. The other model, $q' = 2.60$, has a greater mean-squared error $r_{\rm c}' = R[Y, F(q')] = 4.94$, so $r_{\rm c}' > r_{\rm c}^*$. However, the best performance (the smallest $r_{\rm c}$-value) with each of these models, q' and q^*, has no robustness to model-uncertainty: $0 = \widehat{\alpha}(q^*, r_{\rm c}^*) = \widehat{\alpha}(q', r_{\rm c}')$.

More importantly, the robustness curves cross at a higher value of $r_{\rm c}$ (corresponding to lower aspiration for prediction-fidelity), as seen in fig. 7. If prediction-error $r_{\rm c}^\circ = 5.14$ is tolerable, then the sub-optimal model q' is more robust than, and hence preferable over, the mean-squared optimal model q^*, at the same performance-aspiration. In other words, since $0 = \widehat{\alpha}(q^*, r_{\rm c}^*)$, the analyst recognizes that performance as good as $r_{\rm c}^*$ is not reliable or feasible with the optimal linear model q^*, and some larger $r_{\rm c}$ value (representing poorer fidelity between model and measurement) must be accepted; the analyst is motivated to "move up" along the q^*-robustness curve. If $r_{\rm c}^\circ$ is an acceptable level of fidelity, then the sub-optimal model q' achieves this performance with greater robustness than the optimal model q^*. In particular, $\widehat{\alpha}(q', r_{\rm c}^\circ) = 0.021$ which is small but still twice as large as $\widehat{\alpha}(q^*, r_{\rm c}^\circ) = 0.011$. In this case, 'good' (that is, q') is preferable to 'best' (q^*).

We see the robustness-preference for a sub-optimal model explicitly in fig. 8, which shows the robustness versus the linear model parameter q, for fixed aspiration $r_{\rm c} = 5.5$ (which is larger than the $r_{\rm c}$-values in fig. 7, so q^* has positive robustness). The least-squares optimal parameter, $q^* = 2.50$, minimizes the mean-squared error $R[Y, F(q)]$, while the robust-optimal parameter, $\widehat{q}_{\rm c} = 2.65$, maximizes the robustness function $\widehat{\alpha}(q, r_{\rm c})$. q^* has lower robustness than $\widehat{q}_{\rm c}$, at the same level of model-data fidelity, $r_{\rm c}$. Specifically, $\widehat{\alpha}(q^*, r_{\rm c}) = 0.033$ is substantially less than $\widehat{\alpha}(\widehat{q}_{\rm c}, r_{\rm c}) = 0.045$. The mean-squared error of $\widehat{q}_{\rm c}$ is $R[Y, F(\widehat{q}_{\rm c})] = 5.09$ which is only modestly worse than the least-squares optimum of $R[Y, F(q^*)] = 4.82$. In short, the performance-sub-optimal model $\widehat{q}_{\rm c}$ has only moderately poorer fidelity to the data than the least-squares optimal model q^*, while the robustness to model-uncertainty of $\widehat{q}_{\rm c}$ is appreciably greater than the robustness of q^*.

In summary, we have established the following conclusions from this example.

First, the performance-optimal model, $f_i(q^*)$, has no immunity to error in the basic structure of the model. The model $f_i(q^*)$, which minimizes the mean-squared discrepancy between measurement and prediction, has zero robustness to modelling errors at its nominal prediction-fidelity, $r_{\rm c}^*$.

Second, this is actually true of *any* model, $f_i(q')$. The value of its mean-squared error is $r_{\rm c}'$ which, as in fig. 7, has zero robustness.

Third, the robustness curves of alternative linear models can cross, as in fig. 7. This shows that a sub-optimal model such as $f_i(q')$ can be more robust to model-uncertainty than the mean-squared optimal model $f_i(q^*)$, when these models are compared at the same aspiration for fidelity between model and measurement, $r_{\rm c}^\circ$ in the figure.

Fourth, the model $\widehat{q}_{\rm c}$ which maximizes the robustness can be substantially more robust than the optimal model q^* which minimizes the least-squared error function, as shown in fig. 8. This robustness curve is evaluated at a fixed value of the performance-satisficing parameter $r_{\rm c}$.

5 Hybrid Uncertainty: Info-gap Supervision of a Probabilistic Decision

5.1 Info-gap Robustness as a Decision-monitor

In sections 2 and 3 we considered the reliability of technological systems. We now consider the reliability of a decision algorithm itself. Many decisions are based on probabilistic considerations. A

[2] $N = 5$ and $y_1, \ldots, y_5 = 1.4, 2.6, 5.6, 8.6, 15.9$.

foremost class of examples entails acceptance tests based on the evaluation of a probability of failure. Paradigmatically, a go/no-go decision hinges on whether the probability of failure is below or above a critical threshold:

$$P_{\rm f}(p) \underset{\text{no-go}}{\overset{\text{go}}{\lessgtr}} P_{\rm c} \tag{84}$$

where p is a probability density function (PDF) from which the probability of failure, $P_{\rm f}(p)$, is evaluated.

This is a valid and meaningful decision procedure when the PDF is well known and when the probability of failure can be assessed with accurate system models. However, a decision algorithm such as eq.(84) will be unreliable if the PDF is uncertain (which will often be the case especially regarding the extreme tails of the distribution) and if the critical probability of failure $P_{\rm c}$ is small (which is typically the case with critical components). When the PDF is imprecisely known, the reliability of the probabilistic decision can be assessed by using the info-gap robustness function.

Let $\widetilde{p}$ be the best estimate of the PDF, which is recognized to be wrong to some unknown extent. For sake of argument let us suppose that, with $\widetilde{p}$, the probability of failure is acceptably small:

$$P_{\rm f}(\widetilde{p}) \le P_{\rm c} \tag{85}$$

That is, the nominal PDF implies 'all systems go'. However, since $\widetilde{p}$ is suspect, we would like to know how immune this decision is to imperfection of the PDF.

Let $\mathcal{U}(\alpha, \widetilde{p})$, $\alpha \ge 0$, be an info-gap model for the uncertain variation of the actual PDF with respect to the nominal, best estimate, $\widetilde{p}$. (We'll encounter an example shortly). The robustness, to uncertainty in the PDF, of decision algorithm eq.(84), is the greatest horizon of uncertainty up to which all PDFs lead to the same decision:

$$\widehat{\alpha}(P_{\rm c}) = \max \left\{ \alpha : \ \max_{p \in \mathcal{U}(\alpha, \widetilde{p})} P_{\rm f}(p) \le P_{\rm c} \right\} \tag{86}$$

$\widehat{\alpha}(P_{\rm c})$ is the greatest horizon of uncertainty in the PDF, up to which all densities p in $\mathcal{U}(\alpha, \widetilde{p})$ yield the same decision as $\widetilde{p}$. If $\widehat{\alpha}(P_{\rm c})$ is large, then the decision based on $\widetilde{p}$ is immune to uncertainty in the PDF and hence reliable. Alternatively, if $\widehat{\alpha}(P_{\rm c})$ is small, then decision based on $\widetilde{p}$ is of questionable validity. We see that the robustness function $\widehat{\alpha}(P_{\rm c})$ is a decision-evaluator: it supports the higher-level judgment (how reliable is the probabilistic algorithm?) which hovers over and supervises the ground-level go/no-go decision.

If the inequality in eq.(85) were reversed and $\widetilde{p}$ implied 'no-go', then we would modify eq.(86) to:

$$\widehat{\alpha}(P_{\rm c}) = \max \left\{ \alpha : \ \min_{p \in \mathcal{U}(\alpha, \widetilde{p})} P_{\rm f}(p) \ge P_{\rm c} \right\} \tag{87}$$

The meaning of the robustness function as a decision-monitor would remain unchanged. $\widehat{\alpha}(P_{\rm c})$ is still the greatest horizon of uncertainty up to which the decision remains constant.

We can formulate the robustness slightly differently as the greatest horizon of uncertainty at which the probability of failure does not dither more than $\pi_{\rm c}$:

$$\widehat{\alpha}(\pi_{\rm c}) = \max \left\{ \alpha : \ \max_{p \in \mathcal{U}(\alpha, \widetilde{p})} |P_{\rm f}(p) - P_{\rm f}(\widetilde{p})| \le \pi_{\rm c} \right\} \tag{88}$$

Other variations are also possible [5], but we now proceed to a simple example of the use of the info-gap robustness function in the supervision of a probabilistic decision with an uncertain PDF.

5.2 Non-linear Spring

Consider a spring with the following non-linear relationship between displacement x and force f:

$$f = k_1 x + k_2 x^2 \tag{89}$$

The spring fails if the magnitude of the displacement exceeds $x_{\rm c}$, and we require the probability of failure not to exceed $P_{\rm c}$.

The loading force f is non-negative but uncertain, and the best available PDF is a uniform density:

$$\widetilde{p}(f) = \left\{ \begin{array}{ccc} 1/F & \text{if} & 0 \le f \le F \\ 0 & \text{if} & F \le f \end{array} \right. \tag{90}$$

where the value of F is known. However, it is recognized that forces greater than F may occur. The probability of such excursions, though small, is unknown, as is the distribution of this high-tail probability. That is, the true PDF, shown schematically in fig. 9, is:

$$p(f) = \left\{ \begin{array}{ccc} \text{constant} & \text{if} & 0 \le f \le F \\ \text{variable} & \text{if} & F \le f \end{array} \right. \tag{91}$$

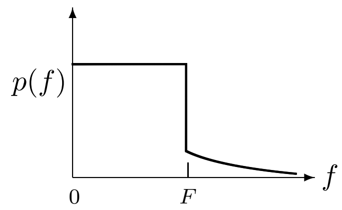

Figure 9: Uncertain probability density function of the load, eq.(91).

The first question we must consider is how to model the uncertainty in the pdf of the force f. What we **do know** is that f is non-negative, that $p(f)$ is constant for $0 \le f \le F$, and the value of F. What we **do not know** is the actual constant value of $p(f)$ for $0 \le f \le F$ and the behavior of $p(f)$ for $f > F$. We face an info-gap.

Let $\mathcal{P}$ denote the set of all non-negative and normalized PDFs on the interval $[0,\ \infty)$. Whatever form $p(f)$ takes, it must belong to $\mathcal{P}$. An info-gap uncertainty model which captures the information as well as the info-gaps about the PDF is:

$$\mathcal{U}(\alpha, \widetilde{p}) = \left\{ p(f): \quad p(f) \in \mathcal{P}; \quad \int_F^\infty p(f)\,{\rm d}f \le \alpha; \right.$$
$$\left. p(f) = \frac{1}{F}\left(1 - \int_F^\infty p(f)\,{\rm d}f\right),\ 0 \le f \le F \right\}, \quad \alpha \ge 0 \tag{92}$$

The first line of eq.(92) states that $p(f)$ is a normalized PDF whose tail above F has weight no greater than α. The second line asserts that $p(f)$ is constant over the interval $[0,\ F]$ and the weight in this interval is the complement of the weight on the tail.

The spring fails if x exceeds the critical displacement $x_{\rm c}$. This occurs if the force f exceeds the critical load $f_{\rm c}$ which is:

$$f_{\rm c} = k_1 x_{\rm c} + k_2 x_{\rm c}^2 \tag{93}$$

With PDF $p(f)$, the probability of failure is:

$$P_{\rm f}(p) = {\rm Prob}(f \ge f_{\rm c} | p) \tag{94}$$

We require that the failure probability not exceed the critical probability threshold:

$$P_{\rm f}(p) \le P_{\rm c} \tag{95}$$

The robustness of the determination of this threshold-exceedence, based on the nominal PDF $\widetilde{p}$, is $\widehat{\alpha}(P_{\rm c})$ given by eq.(86). The value of this robustness depends on the values of F and $f_{\rm c}$. After some algebra one finds:

$$\widehat{\alpha}(P_{\rm c}) = \begin{cases} 0 & \text{if} \quad f_{\rm c} \le (1-P_{\rm c})F \\ 1 - \dfrac{1-P_{\rm c}}{f_{\rm c}/F} & \text{if} \quad (1-P_{\rm c})F < f_{\rm c} \le F \\ P_{\rm c} & \text{if} \quad F < f_{\rm c} \end{cases} \tag{96}$$

The first line of eq.(96) arises when the critical force, $f_{\rm c}$, is small enough so that $P_{\rm f}(p)$ can exceed $P_{\rm c}$ even when the nominal PDF, $\widetilde{p}$, is correct. The third line arises when $f_{\rm c}$ is so large that only the tail could account for failure. The second line covers the intermediate case. $\widehat{\alpha}(P_{\rm c})$ is plotted schematically in fig. 10.

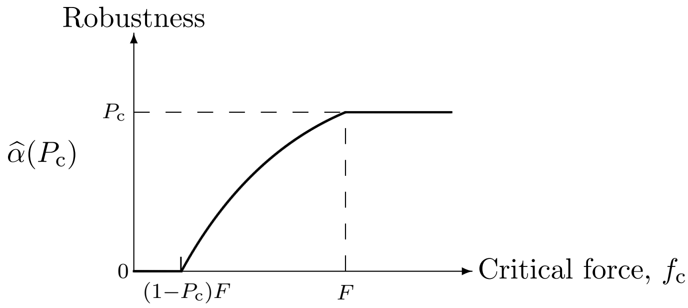

Figure 10: Robustness vs. critical force, eq.(96).

As we explained in section 5.1, the value of the robustness $\widehat{\alpha}(P_{\rm c})$ indicates whether the go/no-go threshold decision, based on the best-available PDF $\widetilde{p}$, is reliable or not. A large robustness implies that the decision is insensitive to uncertainty in the PDF, while a small value of $\widehat{\alpha}(P_{\rm c})$ means that the decision can err as a result of small error in $\widetilde{p}$. From eq.(96) and fig. 10 we see that the greatest value which $\widehat{\alpha}(P_{\rm c})$ can take is $P_{\rm c}$ itself, the critical threshold value of failure probability. In fact $\widehat{\alpha}(P_{\rm c})$ may be much less, depending on the critical force $f_{\rm c}$. From eq.(86) we learn that $\widehat{\alpha}(P_{\rm c})$ and α have the same units. The first line of eq.(92) indicates that α is a probability: the statistical weight of the upper tail. Consequently, $\widehat{\alpha}(P_{\rm c})$ is the greatest tolerable statistical weight of the unmodelled upper tail of $p(f)$. $\widehat{\alpha}(P_{\rm c})$ must be 'large' in order to warrant the go/no-go decision; the judgment whether $\widehat{\alpha}(P_{\rm c})$ is 'small' or 'large' depends on a judgment of how wrong $\widetilde{p}(f)$ could be. As the tolerable probability of failure, $P_{\rm c}$, becomes smaller, the tolerance against probability "leakage" into the upper tail becomes lower as well. $\widehat{\alpha}(P_{\rm c})$ establishes a quantitative connection between the critical force $f_{\rm c}$, the nominally maximum force F, the critical probability $P_{\rm c}$, and the reliability of the go/no-go decision.

An additional use of the robustness function is in choosing technical modfications of the system itself which enhance the reliability of the go/no-go decision in the face of load uncertainty. Examination of eq.(96) reveals that $\widehat{\alpha}(P_{\rm c})$ is improved by increasing $f_{\rm c}$ (if $f_{\rm c} < F$). Let us consider the choice of the two stiffness coefficients, k_1 and k_2. From eq.(93) we note that:

$$\frac{\partial f_{\rm c}}{\partial k_1} = x_{\rm c} > 0 \tag{97}$$

$$\frac{\partial f_{\rm c}}{\partial k_2} = x_{\rm c}^2 > 0 \tag{98}$$

Thus f_c is increased, and thereby $\widehat{\alpha}(P_c)$ is improved, by increasing either k_1 or k_2 or both. Eq.(96) quantifies the robustness-enhancement from a design change in k_1 or k_2.

Eq.(94) implies that increasing f_c causes a reduction in the probability of failure, $P_f(p)$, regardless of how the load is distributed. Thus, a change in the system which reduces the probability of failure, also makes the prediction of the failure-probability more reliable. There is a **sympathy** between **system reliability** and **prediction reliability.** This is in fact not just a favorable quirk of this particular example. It is evident from the definition of robustness in eq.(86) that any system modification which decreases $P_f(p)$ will likewise increase (or at least not decrease) the robustness $\widehat{\alpha}(P_c)$.

6 Why 'Good' is Preferable to 'Best'

The examples in sections 2–4 illustrated the general proposition that optimization of performance is associated with minimization of immunity to uncertainty. This led to the conclusion that performance should be satisficed — made adequate but not optimal — and that robustness should be optimized. In the present section we put this conflict between performance and robustness on a rigorous footing.

6.1 The Basic Lemma

The designer must choose values for a range of variables. These variables may represent materials, geometrical dimensions, devices or components, design concepts, operational choices such as 'go' or 'no-go', etc. Some of these variables are expressible numerically, some linguistically. We will represent the collection of the designer's decisions by the **decision vector** q.

In contrast to q, which is under the designer's control, the designer faces uncontrollable uncertainties of many sorts. These may be uncertain material coefficients, unknown and unmodelled properties such as non-linearities in the design-base models, unknown external loads or ambient conditions, uncertain tails of a probability distribution, and so on. The uncertainties are all represented as vectors or functions (which may be vector-valued). We represent the uncertain quantities by the **uncertain vector** u. The uncertainties associated with u are represented by an info-gap model $\mathcal{U}(\alpha, \widetilde{u})$, $\alpha \geq 0$. The centerpoint of the info-gap model is the known vector $\widetilde{u}$, which is the nominal value of the uncertain quantity u.

Info-gap models are suitable for representing ignorance of u for both practical and fundamental reasons. Practically, probability models defined on multi-dimensional function spaces tend to be cumbersome and informationally intensive. More fundamentally, info-gap models entail no measure functions, while measure-theoretic representation of ignorance can lead to contradictions [7, chap. 4].

Many design specifications can be expressed as a collection of inequalities on scalar valued functions. For instance, the mechanical deflection must not exceed a given value, while each of the 3 lowest natural frequencies must be no greater than various thresholds. For a design choice q, and for a specific realization of the uncertainty u, the performance of the system is expressed by the real-valued **performance functions** $R_i(q, u)$, $i = 1, \ldots, N$, where the design specification is the following set of inequalities:

$$R_i(q, u) \leq r_{c,i} \quad \text{for all} \quad i = 1, \ldots, N \tag{99}$$

The $r_{c,i}$ are called **critical thresholds,** which are represented collectively by the vector r_c. These thresholds may be chosen either small or large, to express either demanding or moderate aspirations, respectively.

Definition 1 *A performance function $R_i(q, u)$ is* **upper unsatiated** *at design q if its maximum, up to horizon of uncertainty α, increases strictly as α increases:*

$$\alpha < \alpha' \quad \Longrightarrow \quad \max_{u \in \mathcal{U}(\alpha, \widetilde{u})} R_i(q, u) < \max_{u \in \mathcal{U}(\alpha', \widetilde{u})} R_i(q, u) \tag{100}$$

Upper unsatiation is a type of monotonicity: the maximum of the performance function strictly increases as the horizon of uncertainty increases. This monotonicity in α does not imply monotonicity of $R(q,u)$ in either q or u. Upper satiation results from the nesting of the sets in the info-gap model $\mathcal{U}(\alpha,\widetilde{u})$.

By definition, the robustness of design q, with performance requirements $r_{\rm c}$, is the greatest horizon of uncertainty at which all the performance functions satisfy their critical thresholds:

$$\widehat{\alpha}(q,r_{\rm c}) = \max\left\{\alpha : \max_{u\in\mathcal{U}(\alpha,\widetilde{u})} R_i(q,u) \le r_{{\rm c},i}, \ \text{for all } i=1,\ldots,N\right\} \tag{101}$$

We now assert the following basic lemma. (Proofs appear in section 6.4.)

Lemma 1 Given:
∘ *An info-gap model* $\mathcal{U}(\alpha,\widetilde{u})$, $\alpha \ge 0$.
∘ *Performance functions* $R_i(q,u)$, $i=1,\ldots,N$, *which are all upper unsatiated at* q .
∘ *Critical thresholds equalling the performance functions evaluated at the center-point of the info-gap model:*

$$r_{{\rm c},i} = R_i(q,\widetilde{u}), \quad i=1,\ldots,N \tag{102}$$

Then *the robustness-to-uncertainty of design* q *vanishes:*

$$\widehat{\alpha}(q,r_{\rm c}) = 0 \tag{103}$$

$\widetilde{u}$ is the centerpoint of the info-gap model: the known, nominal, 'best-estimate' of the uncertainties accompanying the problem. If $\widetilde{u}$ precisely represents the values of these auxiliary variables, then the performance aspirations in eq.(102) will be achieved by decision q. However, eq.(103) asserts that this level of performance has no immunity to unknown variations in the data and models upon which this decision is based. Any unmodelled factors, such as the higher-order terms in eq.(76), jeopardize the performance level vouched for in eq.(102). Probabilistically one would say that things can be worse than the expected outcome. However our assertion is stronger, since we are considering not only random uncertainties, but rather the info-gaps in the entire epistemic infrastructure of the decision, which may include info-gaps in model-structures and probability densities.

This result is particularly significant when we consider performance-optimization, to which we now turn.

6.2 Optimal-Performance vs. Optimal Robustness: The Theorem

We are particularly interested in the application of lemma 1 to optimal-performance design. The lemma will show that a design which optimizes the performance will have zero robustness to uncertainty. This means that high aspirations for performance are infeasible in the sense that these aspirations can fail to materialize due to infinitesimal deviations of the uncertain vector from its nominal value. It is true that failure to achieve an ultimate aspiration may entail only a slight reduction below optimal performance. Nonetheless, the gist of the theorem is that zenithal performance can not be relied upon to occur; a design specification corresponding to an extreme level of performance has no robustness to uncertainty. The designer can not "sign-off" on a performance-optimizing specification; at most one can hope that the short-fall will not be greatly below the maximum performance.

Let $\mathcal{Q}$ represent the set of available designs from which the designer must choose a design q. Let $\widetilde{u}$ denote the nominal, typical, or design-base value of the uncertain vector u. What is an optimal-performance design, from the allowed set $\mathcal{Q}$, and with respect to the design-base value $\widetilde{u}$?

If there is only one design specification, so $N=1$ in eq.(99), then an optimal-performance design q^* minimizes the performance function:

$$R(q^*,\widetilde{u}) = \min_{q\in\mathcal{Q}} R(q,\widetilde{u}) \tag{104}$$

If there are multiple design specifications ($N > 1$) then such a minimum may not hold simultaneously for all the performance functions. One natural extension of eq.(104) employs the idea of **Pareto efficiency.** Pareto efficiency is a 'short blanket' concept: if you pull up your bed covers to warm your nose, then your toes will get cold. A design q^* is Pareto efficient if any other design q' which improves (reduces) one of the performance functions, detracts from (increases) another:

$$\text{If:} \qquad R_i(q', \widetilde{u}) \; < \; R_i(q^*, \widetilde{u}) \quad \text{for some} \quad i \tag{105}$$

$$\text{Then:} \qquad R_j(q', \widetilde{u}) \; > \; R_j(q^*, \widetilde{u}) \quad \text{for some} \quad j \neq i \tag{106}$$

A Pareto efficient design does not have to be unique, so let us denote the set of all Pareto efficient designs by Q^*. For the case of a single design specification, let Q^* denote all the designs which minimize the performance function, as in eq.(104).

The following theorem, which is derived directly from lemma 1, asserts that any design which is performance-optimal (eq.(104)) or Pareto efficient (eqs.(105) and (106)) has no robustness to uncertain deviation from the design-base value $\widetilde{u}$.

Theorem 1 Given:

∘ *A set of Pareto efficient or performance-optimal designs Q^*, with respect to the design-base value $\widetilde{u}$.*

∘ *An info-gap model $\mathcal{U}(\alpha, \widetilde{u})$, $\alpha \geq 0$, whose center-point is the nominal or design-base value $\widetilde{u}$.*

∘ *Performance functions $R_i(q^*, u)$, $i = 1, \ldots, N$, which are all upper unsatiated at some $q^* \in Q^*$.*

∘ *Critical thresholds equalling the performance functions evaluated at the center-point of the info-gap model and at this q^*:*

$$r_{\mathrm{c},i} = R_i(q^*, \widetilde{u}), \quad i = 1, \ldots, N \tag{107}$$

Then *the robustness-to-uncertainty of this performance-optimal design q^* vanishes:*

$$\widehat{\alpha}(q^*, r_{\mathrm{c}}) = 0 \tag{108}$$

Theorem 1 is really just a special case of lemma 1. We know from lemma 1 that whenever the critical thresholds, $r_{\mathrm{c},i}$, are chosen at the nominal values of the performance-functions, $R_i(q, \widetilde{u})$, and when that nominal value, $\widetilde{u}$, is the center-point of the info-gap model, $\mathcal{U}(\alpha, \widetilde{u})$, then any design has zero robustness. Theorem 1 just specializes this to the case where q is Pareto-efficient or performance-optimal.

We can understand the special significance of this result in the following way. The functions $R_i(q^*, \widetilde{u})$ represent the designer's best available representation of how design q^* will perform. $R_i(q^*, \widetilde{u})$ is based on the best available models, and $\widetilde{u}$ is the best estimate of all residual (and possibly recalcitrant) uncertain factors or terms. We know from lemma 1 that any performance-optimal design which we choose will have zero robustness to the vagaries of those residual uncertainties. However, the lesson to learn is **not** to choose the design whose performance is optimal; theorem 1 makes explicit that this design also can not be depended upon to fulfil our expectations. To state it harshly, choosing the performance-optimum is just wishful thinking, unless we are convinced that no uncertainties lurk behind our models. The lesson to learn, if we seek a design whose performance can be reliably known in advance, is to move off the surface of Pareto-efficient or performance-optimal solutions. For example, referring again to fig. 2, we must move off the optimal-performance curve to a point Q or R. We can evaluate the robustness of these sub-optimal designs with the robustness function, and we can choose the design to satisfice the performance and to maximize the robustness.

6.3 Information-gap Models of Uncertainty

We have used info-gap models throughout this paper. In this section we present a succint formal definition, in preparation for the proof of lemma 1 in section 6.4.

An info-gap model of uncertainty is a family of nested sets; eqs.(14), (41), (77) and (92) are examples. An info-gap model entails no measure functions (probability densities or membership functions). Instead, the limited knowledge about the uncertain entity is invested in the structure of the nested sets of events.

Mathematically, an info-gap model is a set-valued function. Let S be the space whose elements represent uncertain events. S may be a vector space or a function space. Let $\Re$ denote the set of non-negative real numbers. An info-gap model $\mathcal{U}(\alpha, u)$ is a function from $\Re \times S$ into the class of subsets of S. That is, each ordered pair (α, u), where $\alpha \geq 0$ and $u \in S$, is mapped to a set $\mathcal{U}(\alpha, u)$ which is a subset of S.

Two axioms are central to the definition of info-gap models of uncertainty:

Axiom 1: Nesting. *An info-gap model is a family of nested sets:*

$$\alpha \leq \alpha' \quad \Longrightarrow \quad \mathcal{U}(\alpha, u) \ \subseteq \ \mathcal{U}(\alpha', u) \tag{109}$$

We have already encountered this property of info-gap models in eqs.(9), (37) and (67) where we noted that this inclusion means that the uncertainty sets $\mathcal{U}(\alpha, u)$ become more inclusive as the 'uncertainty parameter' α becomes larger. Relation (109) means that α is an **horizon of uncertainty.** At any particular value of α, the corresponding uncertainty set defines the range of variation at that horizon of uncertainty. The value of α is unknown, so the family of nested sets is typically unbounded, and there is no 'worst case'.

All info-gap models of uncertainty share an additional fundamental property.

Axiom 2: Contraction. *The info-gap set, at zero horizon of uncertainty, contains only the center-point:*

$$\mathcal{U}(0, u) = \{u\} \tag{110}$$

For instance, the info-gap model in eq.(14) is a family of nested intervals, and the center-point is the nominal load, $\widetilde{\phi}$, which is the only element of $\mathcal{U}(0, \widetilde{\phi})$ and which belongs to the intervals at all positive values of α.

Combining axioms 1 and 2 we see that u belongs to the zero-horizon set, $\mathcal{U}(0, u)$, and to all 'larger' sets in the family.

Additional axioms are often used to define more specific structural features of the info-gap model, such as linear [2, 4] or non-linear [3] expansion of the sets as the horizon of uncertainty grows. We will not need these more specific axioms. Info-gap models of uncertainty are discussed extensively elsewhere [1, 4].

6.4 Proofs

Proof of lemma 1. By the contraction axiom of info-gap models, eq.(110), and from the choice of the value of $r_{\mathrm{c},i}$:

$$\max_{u \in \mathcal{U}(0, \widetilde{u})} R_i(q, u) = r_{\mathrm{c},i} \tag{111}$$

Hence 0 belongs to the set of α-values in eq.(101) whose least upper bound equals the robustness, so:

$$\widehat{\alpha}(q, r_{\mathrm{c}}) \geq 0 \tag{112}$$

Now consider a positive horizon of uncertainty: $\alpha > 0$. Since $R_i(q, r_{\mathrm{c}})$ is upper unsatiated at q:

$$\max_{u \in \mathcal{U}(0, \widetilde{u})} R_i(q, u) \ < \ \max_{u \in \mathcal{U}(\alpha, \widetilde{u})} R_i(q, u) \tag{113}$$

Together with eq.(111), this implies that this positive value of α does not belong to the set of α-values in eq.(101). Hence:

$$\alpha > \widehat{\alpha}(q, r_{\mathrm{c}}) \tag{114}$$

Combining relations (112) and (114) completes the proof. ■

Proof of theorem 1. Special case of lemma 1. ■

7 Conclusion: An Historical Perspective

We have focussed on the epistemic limitations — information gaps — which confront a designer in the search for reliable performance. The central idea has been that the functional performance of a system must be traded-off against the immunity of that system to info-gaps in the models and data underlying the system's design. A system which is designed for maximal performance will have no immunity to errors in the models and data underlying the design. Robustness can be obtained only by reducing performance-aspirations.

Our discussion is motivated by the recognition that the designer's understanding of the relevant processes is deficient, that the models representing those processes lack pertinent components, and that the available data are incomplete and inaccurate. This very broad conception of uncertainty — including model structure as well as more conventional data "noise" — has received extensive attention in some areas of engineering, most notably in robust control [9]. This scope of uncertainty is, however, a substantial deviation from the tradition of probabilistic analysis which dominates much contemporary thinking.

Least-squares estimation is a central paradigm of traditional uncertainty analysis. The least-squares method was developed around 1800, independently by Gauss (1794–5) and Legendre (1805–8), for estimation of celestial orbits [6, 10]. Newtonian mechanics, applied to the heavenly bodies, had established irrevocably that celestial orbits are elliptical, as Kepler had concluded experimentally. However, the data were noisy so it was necessary to extract the precise ellipse which was hidden under the noisy measurements. The data were corrupted, while the model — elliptical orbits — was unchallenged. Least-squares estimation obtained deep theoretical grounding with the proof of the central limit theorem by Laplace (1812), which established the least-squares estimate as the maximum-likelihood estimate of a normal distribution. The least-squares idea continues to play a major role in modern uncertainty analysis in such prevalent and powerful tools as Kalman filtering, Luenberger estimation [8] and the Taguchi method [11].

What is characteristic of the least-squares method is the localization of uncertainty exclusively on data which are exogenous to the model of the underlying process. The model is unblemished (Newtonian truth in the case of celestial orbits); only the measurements are corrupted. But the innovative designer, using new materials and exploiting newly discovered physical phenomena, stretches models and data to the limits of their validity. The designer faces a serious info-gap between partial, sometimes tentative, insights which guide much high-paced modern design, and solid complete knowledge. The present work is part of the growing trend to widen the range of uncertainty analysis to include the analyst's imperfect conceptions and representations. For all our sapience, we are after all only human!

Acknowledgement

The author acknowledges with pleasure the support of the Samuel Neaman Institute for Advanced Studies in Science and Technology. The author is indebted to Prof. Eli Altus and Mr. Yakov Saraf of the Technion for useful comments.

8 References

1. Yakov Ben-Haim, 1996, *Robust Reliability in the Mechanical Sciences,* Springer-Verlag, Berlin.

2. Yakov Ben-Haim, 1999, Set-models of information-gap uncertainty: Axioms and an inference scheme, *Journal of the Franklin Institute,* 336: 1093–1117.

3. Yakov Ben-Haim, 2000, Robustness of model-based fault diagnosis: Decisions with information-gap models of uncertainty. *International Journal of Systems Science,* 31: 1511–1518.

4. Yakov Ben-Haim, 2001, *Information-gap Decision Theory: Decisions Under Severe Uncertainty,* Academic Press, London.

5. Yakov Ben-Haim, Scott Cogan and Laëtitia Sanseigne, 1998, Usability of Mathematical Models in Mechanical Decision Processes, *Mechanical Systems and Signal Processing.* 12: 121–134.

6. Michiel Hazewinkel, managing editor, 1988, *Encyclopaedia of Mathematics: An Updated and Annotated Translation of The Soviet "Mathematical Encyclopaedia",* Kluwer Academic Publishers, Dordrecht, Holland.

7. J.M. Keynes, 1921, *Treatise on Probability,* Macmillan & Co. Ltd., London.

8. David G. Luenberger, 1969, *Optimization by Vector Space Methods,* Wiley, New York.

9. Zhihua Qu, 1998, *Robust Control of Nonlinear Uncertain Systems,* Wiley, New York.

10. Stephen M. Stigler, 1986, *The History of Statistics: The Measurement of Uncertainty before 1900.* The Belknap Press of Harvard University Press, Boston.

11. Genichi Taguchi, 1986, *Introduction to Quality Engineering: Designing Quality into Products and Processes,* translated and published by the Asian Productivity Organization, Tokyo.