THE ROBUST DESIGN OF MECHANICAL ELEMENTS AND SYSTEMS
CSME Forum 2002

Jorge Angeles
Department of Mechanical Engineering
& Centre for Intelligent Machines
McGill University, Montreal, Quebec, Canada
angeles@cim.mcgill.ca

Abstract

The concepts underlying Taguchi's *robust engineering* are revised with the purpose of applying them to the general area of mechanical design. We have found that Taguchi's is a black-box approach to design, which was the rule in the postwar years, when these concepts were developed. Not any more. Current design engineering is largely model-based, although the need for goods and services that are as insensitive as possible to environmental changes is growing, mostly because of the need to design reliable, accurate multi-degree-of-freedom machinery. Along the same lines, we revise here Suh's *axiomatic design*. We have found that the two paradigms, Taguchi's and Suh's, are compatible, although their formalisms may be disparate. In the end, we propose a framework intended for a sound formulation of robust engineering based on the underlying mathematical model on which every modern design job relies.

Introduction

Robust engineering is usually associated with Taguchi's *philosophy* pertaining to *off-line* quality control, as developed in the postwar years to respond to the needs of the Japanese industry of the time [1]. Indeed, Taguchi realized that industry in those days had adopted the practice of *on-line* quality control, whereby the *production process* is monitored at the end of the production line. What robust engineering propounds is quality control off-line, i.e., at the design stage, whereby extensive experiments can be conducted, in a *controlled environment*, to determine the sensitivity of the *performance* of the design product to variations in the *environment conditions*.

As a matter of fact, robust engineering is only an instance of *robust decision-making*, which encompasses three main activities: (*i*) management; (*ii*) design engineering; and (*iii*) control engineering. We will focus here on robust design engineering, while paying special attention to mechanical design. In this context, robust design has received increasing attention. We can cite here the work of Gadallah and ElMaraghy [2], who proposed a method aimed at robust mechanical design, which they exemplified with the design of a clutch. Zhu and Ting [3], in turn, reported on the robust design of linkages. Germane to robust design, *axiomatic design*, as propounded by Suh [4], aims at the optimum design of products, including software, based on two fundamental principles, which Suh calls *axioms*, as described below.

Robust Design vs. Axiomatic Design

While Suh developed his design principles independent of Taguchi's robust engineering, the two approaches are equivalent to some extent. What makes them look disparate is the lack of a theoretical framework encompassing both. We recall briefly these two approaches.

Robust Design

Taguchi's *philosophy* relies on two key concepts:

1. The *signal-to-noise (S/N) ratio*, that measures the sensitivity of a design to changes in environment conditions—the larger the S/N ratio, the more robust the design—and

2. The *loss function*, that measures the loss of society by virtue of a flawed design.

The signal-to-noise ratio thus gives a sense of how sensitive the design is to variations in the design environment, and hence, in robust design one should aim at the maximization of the S/N ratio. The loss function, in turn, plays a role in design similar to that of the *dissipation function* in mechanics, which stems, in turn, from the *Second Law of Thermodynamics*. Here, we cannot help but think of the Concorde catastrophe: On July 25th, 2000, an Air France Concorde killed 113 people upon taking off at Paris Charles-de-Gaulle Airport. According to the French government report, released on January 15th, 2002, the culprit was a piece of debris left on the runway by a Continental Airlines DC-10 that had taken off just before the Concorde. A piece of metal that fell from the DC-10 fuselage, over which the Concorde ran, blew its tires, sparks were generated by the friction between runway and metal, which thus burned the tires, thereby producing additional debris, which went into the engine intake. The ensuing flames that engulfed the airplane were a natural consequence of the foregoing series of fatal events. The actual culprit, however, was a flawed design: The main landing gear of the Concorde is placed just *before* the engine intake. In subsonic civilian airplanes, the landing gear is *after* the engine intake. The loss of life and property was enormous.

Taguchi's *methodology*, in turn, is based on a suitably defined series of experiments whose results are then stored in what Taguchi calls *orthogonal arrays*. These are tableaus relating recorded variations in environment conditions—here, contrary to the environment under which the designed object will operate, the designer *can control* the environment parameters—with the corresponding variations in the performance indicators. The procedure by which the designer decides on the values to be assigned to the design variables is well established within the practice of robust engineering, and consists in selecting the appropriate combination of values according to the orthogonal arrays.

Axiomatic Design

Suh's axiomatic design methodology relies on two *axioms*:

1. *The Independence Axiom*: The best design is one in which all *design functions* are independent.
2. *The Minimum-Information Axiom*: The best design is the one containing the minimum amount of information.

While Taguchi's concepts are transparent, in that they involve well-known engineering items, namely, signals, noise and loss functions, the foregoing axioms need further explanation. We start with the Independence Axiom, which is best explained with the design of a simple mechanical system, a pitch-roll robotic wrist. This is a two-degree-of-freedom mechanical system intended for two independent *functions*, pitch at a rate p about a horizontal axis, and roll at a rate r about the axis of symmetry of the wrist gripper. Two alternatives are considered, as shown in Figs. 1 and 2 .

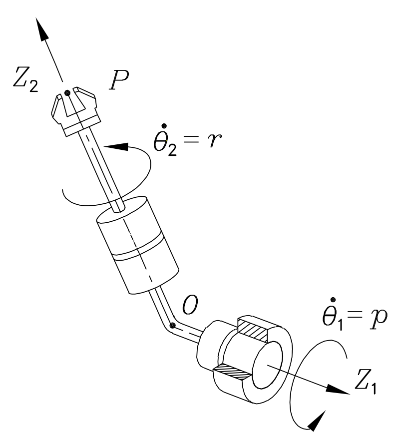

Figure 1: A directly-driven pitch-roll wrist

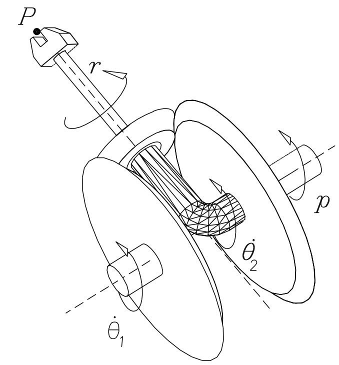

Figure 2: An indirectly-driven pitch-roll wrist

The first alternative is based on actuators collocated at the pitch and the roll axes, which is why we call this layout "directly-driven," or, simply, *design A*. The second alternative comprises a differential gear train between the actuators and the pitch and the roll axes, which is why we call this layout "indirectly-driven," or simply *design B*. In order to best describe the velocity relations in the foregoing system, we introduce the vector of joint rates $\dot{\boldsymbol{\theta}} = [\,\dot{\theta}_1,\, \dot{\theta}_2\,]^T$, which are produced by the actuators, and the vector of *Cartesian velocities* $\mathbf{v} = [\,p,\, r\,]^T$, which represent the functions for which the wrist is designed. Both vectors are linearly related by a "Jacobian" matrix $\mathbf{J}$, i.e.,

$$\dot{\boldsymbol{\theta}} = \mathbf{J}\mathbf{v} \tag{1a}$$

Let $\mathbf{J}_A$ and $\mathbf{J}_B$ denote the Jacobian matrices for case A and for case B, respectively. We thus have, after rather simple calculations,

$$\mathbf{J}_A = \mathbf{1}, \quad \mathbf{J}_B = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ \nu & -\nu \end{bmatrix} \tag{1b}$$

in which $\mathbf{1}$ denotes the 2×2 identity matrix and ν is the gear ratio between sun and planet gears. That is, if we let N_s and N_p denote the numbers of teeth of the sun and the planet gears, respectively, then

$$\nu = \frac{N_s}{N_p} \tag{2}$$

In Suh's formulation, the pitch and the roll rates represent the *functional requirements*, while $\dot{\theta}_1$ and $\dot{\theta}_2$ would play the role of the *design parameters*, the Jacobian matrix $\mathbf{J}$ thus being the *design matrix*. According to the Independence Axiom, then, design A is better than design B because $\mathbf{J}_A$ is diagonal, and hence, leads to a "decoupling of functions," thereby verifying the Independence Axiom. Design B is coupled, and hence, violates the Independence Axiom. Within the axiomatic design paradigm, the design matrix would better be square, although rectangular matrices can also be accommodated. Indeed, a design job in which the number of design parameters equals that of functional requirements is called "ideal"; one in which the former is smaller than the latter is called "coupled"; the opposite of a coupled design is called "redundant."

Two remarks follow:

- Design tasks in which the number of functional requirements and that of design parameters coincide are the exception more so than the rule. In the particular example at hand we have equal numbers, but this with the purpose of illustration only.

- In many instances the design matrix is diagonalizable, the decoupling of functions thus being possible, regardless of the quality of the design. However, the quality of a design should not be dependent on a change of variables, which is what diagonalization actually amounts to.

The Minimum Information Axiom presupposes a measure of the information contained in a design. Suh defines this measure based on the classical definition of information content [5], i.e., as the logarithm of the reciprocal of the probability of an event. In axiomatic design, the event is meeting a functional requirement. In this vein, the ith functional requirement is assigned, somehow, a probability p_i of being met, the information content I_i for this requirement being defined as

$$I_i = \log_2 \left(\frac{1}{p_i} \right) \tag{3}$$

where base-2, or binary, logarithms are used, and hence, the unit of information content is a *bit*. If natural logarithms are used instead, then the unit of information content is called a *nat*. Thus, the information content I of a given design is the binary logarithm of the reciprocal of the product of the m individual probabilities, under the assumption that the number of functional requirements is m, i.e.,

$$I = \log_2 \left(\frac{1}{p_1 p_2 \cdots p_m} \right) \tag{4}$$

In fact, this axiom is more elusive to grasp for the layman. However, notice that the common ground between Taguchi's robust engineering and Suh's axiomatic design lies in the parallelism between *energy*, as in the loss function and the Second Law of Thermodynamics, and *information*, as pertaining to its transmission. As information is transmitted, entropy is produced that leads to a loss of information, in the same way that entropy is the culprit for the loss of useful energy to unrecoverable heat.

A Theoretical Framework for Robust Design

In setting up a framework for robust design, we start by classifying the variables at play in the design task. We thus have:

- *Design Variables (DV)*: those variables to be decided on by the designer with the purpose of meeting performance specifications under given conditions. In our above example, the design variables are, for the layout of Fig. 1, the angle made by axes Z_1 and Z_2, which need not be 90°, and can be assigned freely by the designer. For the layout of design B, of Fig. 2, in turn, the design variables are N_s and N_p. Other design variables will arise in the sequel. We assume that the design task involves n such variables, which are thus grouped in the *design variable vector* $\mathbf{x}$, i.e.,

$$\mathbf{x} \equiv \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^T$$

- *Design Environment Parameters (DEP)*: Those variables over which the designer has no control, and that define the conditions of the environment under which the designed object will operate. These are, for example, ambient temperature and pressure, humidty level, and so on. These variables, moreover, are of a random nature. In the same example, the environment conditions can be the torques around the pitch and the roll axes exerted by the environment, e.g., as arising from contact forces in a debarring operation. We assume that, for the design task at hand, ν DEP describe the environment conditions, and group all these into the *design environment vector* $\mathbf{p}$, namely,

$$\mathbf{p} \equiv \begin{bmatrix} p_1 & p_2 & \cdots & p_\nu \end{bmatrix}^T$$

- *Performance Functions*: Functions that represent the performance of the design in terms of DV and DEP. In our example above, these functions are readily identified as the pitch and the roll. If we assume that the design task under discussion involves m such functions, then we group them all within the *performance vector* $\mathbf{f}(\mathbf{x};\mathbf{p})$:

$$\mathbf{f} = \mathbf{f}(\mathbf{x};\mathbf{p}) \equiv \begin{bmatrix} f_1 & f_2 & \cdots & f_m \end{bmatrix}^T \quad (5)$$

Global vs. Local Robust Design

Robust design is too ample to be amenable to a thorough analysis with the few elements laid down in the above discussion. In a nutshell, what we need to solve a design problem with a robust solution is a model, which can be of many different kinds. The simplest models to handle are those described by simple *functional relations* of the form of eq.(5). Most practical decision-making problems, of wich design tasks are a subset, are not amenable to such a simple formulation. For example, the decision-making process underlying the creation of the *International Alphabet*, recorded in Fig. 3, is a paradigm of robust design. The International Alphabet, adopted, among others, by the *International Civil Aviation Organization* (ICAO), was created to allow the transmission of a spoken message with the minimum distortion, even in the presence of a variety of accents, dictions and educational backgrounds.

In transmitting an oral message via a noisy communication channel, which involves the vocal system of the speaker, an "A" as such can be mistaken with an "eight." The signal is strengthened by replacing "A" with "Alpha," such a mistake thus becoming much less likely.

Along the same lines, the Concorde catastrophe could have been avoided, had a ro-

A	Alpha		N	November
B	Bravo		O	Oscar
C	Charlie		P	Papa
D	Delta		Q	Quebec
E	Echo		R	Romeo
F	Foxtrot		S	Sierra
G	Golf		T	Tango
H	Hotel		U	Uniform
I	India		V	Victor
J	Juliett		W	Whiskey
K	Kilo		X	X-Ray
L	Lima		Y	Yankee
M	Mike		Z	Zulu

Figure 3: The International Alphabet

bust design been proposed in this engineering masterpiece. What is needed here is a suitable model that would make it apparent that placing the landing gear before the engine intake leads to a weak—as opposed to robust—design. How to formulate this model, however, is an open question. Design objects such as the International Alphabet or a civilian supersonic aircraft that is immune to debris on the runway will be termed here globally-robust designs. At the other end of the spectrum we have *locally-robust designs*, which are modellable in the form of eq.(5). This paper focuses on the latter.

The essence of locally-robust design is the assumption that the functional relation of eq.(5) is differentiable at least with respect to the DEP, and hence, for "small" variations $\Delta\mathbf{p}$ of $\mathbf{p}$, the performance vector $\mathbf{f}$ undergoes correspondingly "small" variations $\Delta\mathbf{f}$, the two variations being related by the *performance matrix.* More concretely, let

$\mathbf{p}_o$: Nominal value of vector $\mathbf{p}$, which thus represents the *nominal operation conditions*;

$\Delta\mathbf{p} \equiv \mathbf{p} - \mathbf{p}_o$: A small variation in the DEP, of a random nature; and

$\Delta\mathbf{f} \equiv \mathbf{f}(\mathbf{x}; \mathbf{p}_o + \Delta\mathbf{p}) - \mathbf{f}(\mathbf{x}; \mathbf{p}_o)$: The small variation in the design performance induced by the variation $\Delta\mathbf{p}$, while keeping the design variables unchanged.

We thus have, in light of eq.(5), and upon linearization around the nominal point $\mathbf{p}_o$, while keeping $\mathbf{x}$ fixed,

$$\Delta\mathbf{f} = \mathbf{F}\Delta\mathbf{p}$$

Notice that $\mathbf{F}$ is nothing but the Jacobian matrix of $\mathbf{f}$ with respect to $\mathbf{p}$, evaluated at the nominal environment conditions, i.e.,

$$\mathbf{F} \equiv \left.\frac{\partial\mathbf{f}}{\partial\mathbf{p}}\right|_{\mathbf{p}=\mathbf{p}_o}$$

Given the random nature of the DEP, we also need models for their randomness, i.e., their probability distributions. In this paper we assume that the variations of the DEP obey normal distributions with zero mean and identical standard deviations, the covariance matrix of vector $\mathbf{p}$ thus being *isotropic*, i.e., a multiple of the identity matrix, namely,

$$E[\Delta\mathbf{p}(\Delta\mathbf{p})^T] = \sigma^2\mathbf{1} \qquad (7)$$

where σ is the common standard deviation and $\mathbf{1}$ denotes the $\nu \times \nu$ identity matrix. Non-isotropic covariance matrices can also be handled within the framework proposed here, but these lead to additional algebraic complexities that are unnecessary for the purposes of this paper.

The Sensitivity Matrix

Properly speaking, "small" variations are meaningless when one refers to physical quantities that bear units. In the foregoing formulation we must assume that all variations are *relative*, and measured with respect to their nominal values. Within the realm of robust engineering, in fact, all variations are measured percentagewise. Hence we can safely assume that all variables involved in eq.(5) are dimensionless. Hence we can measure the magnitude of both $\Delta\mathbf{p}$ and $\Delta\mathbf{f}$ by their *Euclidean norms*, namely, as the square root of the sum of the squares of the components of each of these vectors. We are interested in assessing the magnitude of the variation $\Delta\mathbf{f}$ induced by a variation $\Delta\mathbf{p}$. To this end, we take the square of the Euclidean norm of the two sides of eq.(6), thus obtaining

$$||\Delta\mathbf{f}||^2 = \Delta\mathbf{p}^T\mathbf{F}^T\mathbf{F}\boldsymbol{\Delta}\mathbf{p}$$

The product $\mathbf{F}^T\mathbf{F}$ is termed henceforth the *sensitivity matrix*, and is represented by $\mathbf{S}$, namely,

$$\mathbf{S} \equiv \mathbf{F}^T\mathbf{F}$$

which is apparently a $\nu \times \nu$ positive-definite matrix.

In order to gain insight into the significance of the sensitivity matrix, let us denote with S the *sensitivity* of the design performance to changes in the DEP, i.e.,

$$S \equiv \frac{||\Delta\mathbf{f}||}{||\Delta\mathbf{p}||}$$

Now, since $\mathbf{S}$ is symmetric and positive-definite, its eigenvalues are all real and positive, its eigenvectors being mutually orthogonal. Thus, without loss of generality, assume that $\mathbf{S}$ is in diagonal form:

$$\mathbf{S} = \mathrm{diag}(\lambda_1^2, \cdots, \lambda_\nu^2)$$

where $\lambda_1^2 \leq \lambda_2^2 \leq \cdots \leq \lambda_\nu^2$ are the eigenvalues of the design sensitivity matrix $\mathbf{S}$.

Then,

$$S^2 = \frac{\lambda_1^2\Delta p_1^2 + \cdots + \lambda_\nu^2\Delta p_\nu^2}{\Delta p_1^2 + \cdots + \Delta p_\nu^2}$$

and hence, if we let $||\Delta\mathbf{f}|| = S_o^2$, with S_o a known constant, the above relation defines a ν-axis ellipsoid, namely,

$$\frac{\Delta\mathbf{p}^T\mathbf{S}\Delta\mathbf{p}}{S_o^2} = 1 \qquad (8)$$

which is called the *sensitivity ellipsoid*. Upon expansion of the numerator of eq.(8), and after simple rearrangements, we obtain the *canonical form* of the ellipsoid equation:

$$\frac{||\Delta p_1||^2}{S_o^2/\lambda_1^2} + \frac{||\Delta p_2||^2}{S_o^2/\lambda_2^2} + \cdots + \frac{||\Delta p_\nu||^2}{S_o^2/\lambda_\nu^2} = 1 \quad (9)$$

It is thus apparent that the semi-axes of the sensitivity ellipsoid are inversely proportional to the square roots of the eigenvalues of $\mathbf{S}$, as depicted in Fig. 4 for $\nu = 3$. The eigenvalues of $\mathbf{S}$ can thus be fairly termed the *principal sensitivities* of the design task, the ellipsoid semiaxes indicating the *principal sensitivity directions*, pointing in the directions of the eigenvectors of the design-sensitivity matrix $\mathbf{S}$.

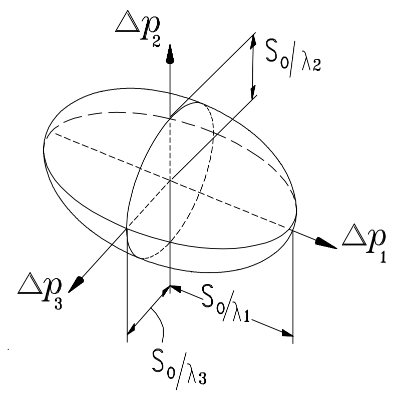

Figure 4: A three-dimensional sensitivity ellipsoid with $\lambda_1 < \lambda_2 < \lambda_3$

If we realize that the position vector of any point on the surface of the sensitivity ellipsoid is $\Delta\mathbf{p}$, it is then apparent that a *large variation* of $\Delta\mathbf{p}$ in the direction of Δp_1 produces the same variation S_o in the performance as a *small variation* of $\Delta\mathbf{p}$ in the direction of Δp_ν. What this means is that the sensitivity of the design performance to variations in the design-environment parameters is more disparate the more widespread the eigenvalues of $\mathbf{S}$ are. In designing a mechanical element or a mechanical system robustly, we should thus aim at maximizing the shortest semiaxis of the sensitivity ellipsoid. Equivalently, we should aim at minimizing the maximum eigenvalue λ_ν of the sensitivity matrix $\mathbf{S}$. Now, since design tasks are always subject to constraints, we cannot render λ_ν a minimum without affecting the other eigenvalues, i.e., without increasing them. The outcome is that, as we attempt to minimize λ_ν, λ_1, the smallest eigenvalue, will tend to increase. The least sensitive design is, then, one with an "ellipsoid" of identical axis lengths, i.e., a sphere. Such a design will be henceforth termed *isotropic*. Notice that the positive square roots of the eigenvalues of $\mathbf{S}$ are the *singular values* of the design matrix $\mathbf{F}$.

Thus, an indicator of the sensitivity of a design to changes in the environment is the *eccentricity* of the sensitivity ellipsoid, one measure of which can be the ratio of the longest to the shortest semiaxes, i.e., $\lambda_{\max}/\lambda_{\min}$, which happens to be one form of defining the *condition number* $\kappa_2(\mathbf{F})$ of the performance matrix $\mathbf{F}$, i.e., the condition number of this matrix based on the matrix 2-norm [6]:

$$\kappa_2(\mathbf{F}) = \frac{\lambda_{\max}}{\lambda_{\min}} \tag{10}$$

Simple as it is, the above definition of the matrix condition number poses challenging problems to the designer when, for example, trying to minimize it over certain parameters. Indeed, the maximum or, for that matter, the minimum eigenvalue of a square matrix is not an *analytic function* of the matrix entries. This means that explicit, simple formulas for its time derivatives are elusive. The good news is that the matrix condition number lends itself to alternative definitions, the one suiting best our aims being the general definition, if based on the *matrix Frobenius norm*. The general definition of the condition number of an $n\times n$ matrix $\mathbf{A}$ is recalled below:

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\| \tag{11}$$

where $\|\cdot\|$ denotes *any* norm of the matrix argument $(\cdot)$. There are various possibilities to compute a norm of $\mathbf{A}$. In particular, the Frobenius norm is defined as

$$\|\mathbf{A}\|_F = \sqrt{\frac{1}{n}\mathrm{tr}(\mathbf{A}\mathbf{A}^T)} \tag{12}$$

which thus assigns unity to the norm of the $n \times n$ identity matrix. If the Frobenius norm is used in the definition of the condition number, then we have, for an $n \times n$ matrix $\mathbf{A}$,

$$\kappa_F = \|\mathbf{A}\|_F\|\mathbf{A}^{-1}\|_F \tag{13}$$

Bandwidth of the Sensitivity Matrix

The matrix condition number is an alien concept to most design engineers. For this reason, we prefer to work with an equivalent concept, that will lead to an engineering measure of the robustness of a given design. The *bandwidth* of a matrix $\mathbf{A}$, measured in decades, is first defined as

$$b = \log_{10}\left(\frac{\sigma_{max}}{\sigma_{min}}\right)$$

where σ_{max} and σ_{min} are the maximum and minimum singular values of the given matrix, respectively. Since the eigenvalues of the sensitivity matrix are identical to the singular values of the performance matrix, we can also write, with the notation introduced above,

$$b = \log_{10}\left(\frac{\lambda_{max}}{\lambda_{min}}\right) \tag{14a}$$

which is the square root of the ratio of the largest to the smallest eigenvalues of the sensitivity matrix. The bandwidth of the performance matrix thus represents an index for evaluating the robustness of a design, and hence, b can also be termed the *design bandwidth.* It is noteworthy that the bandwidth of an isotropic matrix is zero, while that of a singular matrix tends to infinity, i.e.,

$$0 \leq b < \infty$$

Notice that the design bandwidth can be a measure of the information content of the design, and hence, minimizing the bandwidth of the performance matrix is tantamount to minimizing the information content of the design.

Therefore, the spectrum of an isotropic performance matrix is a single harmonic; that of a rank-deficient performance matrix, which has at least one vanishing singular value, is infinity. In the presence of an infinite design bandwidth we have at least one DEP that does not influence the design performance. Also notice that, while the performance matrix need not be square, the sensitivity matrix is necessarily square, symmetric and positive-definite.

In closing this section, an important remark is in order: The sensitivity matrix being symmetric and positive-definite, it need not be associated with a Hessian. Furthermore, the sensitivity matrix is always diagonalizable and hence, always verifies the Independence Axiom. What we have shown is that the diagonality of Suh's design matrix is not at stake in the foregoing axiom, but rather the bandwidth of the performance matrix $\mathbf{F}$ introduced above. However, in an ideal design, the sensitivity matrix is proportional to the identity matrix. The Independence Axiom then should be rephrased to reflect this property. By the same token, the design performance matrix of an ideal design is isotropic, which means that it leads to a sensitivity matrix not only diagonal, but with identical diagonal entries. Moreover, the same design performance matrix is of minimum bandwidth, the ideal design thus satisfying both the Independence Axiom *and* the Minimum Information Axiom.

A Case Study: The Design of a Pitch-Roll Wrist

We compare here the two alternative designs of a pitch-roll wrist, as displayed in Figs. 1 and 2 . We do this for three distinct performances: (a) kinetostatic; (b) elastostatic; and (c) elastodynamic.

Kinetostatic Performance

In our robust design framework, the performance matrix for kinetostatic performance is the Jacobian matrix $\mathbf{J}_A$ for design A and $\mathbf{J}_B$ for design B. The sensitivity matrix for each design is correspondingly $\mathbf{S}_A$ or $\mathbf{S}_B$. These matrices are readily derived below:

$$\mathbf{S}_A = \mathbf{J}_A\mathbf{J}_A^T = \mathbf{1}$$
$$\mathbf{S}_B = \mathbf{J}_B\mathbf{J}_B^T = \frac{1}{4}\begin{bmatrix} 2 & 0 \\ 0 & 2\nu^2 \end{bmatrix}$$

In quantifying the robustness of each design via its bandwidth, we will need the condition number of each Jacobian matrix or, correspondingly, of each sensitivity matrix. Notice that the condition number of the sensitivity matrix is the square of the condition number of the Jacobian. The Frobenius-norm condition numbers $\kappa_F(\mathbf{J}_A)$ and $\kappa_F(\mathbf{J}_B)$ are derived from

$$\kappa_F^2(\mathbf{J}_A) = ||\mathbf{J}_A||_F^2 \, ||\mathbf{J}_A^{-1}||_F^2 \qquad (16a)$$
$$\kappa_F^2(\mathbf{J}_B) = ||\mathbf{J}_B||_F^2 \, ||\mathbf{J}_B^{-1}||_F^2 \qquad (16b)$$

The foregoing matrix norms are calculated below:

$$\|\mathbf{J}_A\|_F^2 = \mathrm{tr}\left(\frac{1}{2}\mathbf{J}_A\mathbf{J}A^T\right) = 1$$
$$\|\mathbf{J}_A^{-1}\|_F^2 = \mathrm{tr}\left(\frac{1}{2}\mathbf{J}_A^{-1}\mathbf{J}_A^{-T}\right) = 1$$
$$\|\mathbf{J}_B\|_F^2 = \mathrm{tr}\left(\frac{1}{2}\mathbf{J}_B\mathbf{J}_B^T\right) = \frac{1}{4}(1+\nu^2)$$
$$\|\mathbf{J}_B^{-1}\|_F^2 = \mathrm{tr}\left(\frac{1}{2}\mathbf{J}_B^{-1}\mathbf{J}B^{-T}\right) = \frac{1+\nu^2}{\nu^2}$$

Hence,

$$\kappa_F(\mathbf{J}_A) = 1, \quad \kappa_F(\mathbf{J}_B) = \frac{1+\nu^2}{2\nu} \qquad (18)$$

Therefore, design A, the directly-driven wrist, is kinetostatically isotropic, and hence, optimum in this regard, while design B depends on ν, the gear ratio between sun and planet gears. Notice that, usually, $N_s > N_p$, and hence, usually $\nu > 1$. However, in this analysis we do not assume the foregoing relation and plot $1/\kappa_F(\mathbf{J}_B)$ for positive values of ν—notice that the condition number being unbounded from above, it is not advisable to attempt a plot of it—for $0 \le \nu \le 50$ in Fig. 5, with a zoom-in for $0 \le \nu \le 10$ in Fig. 6.

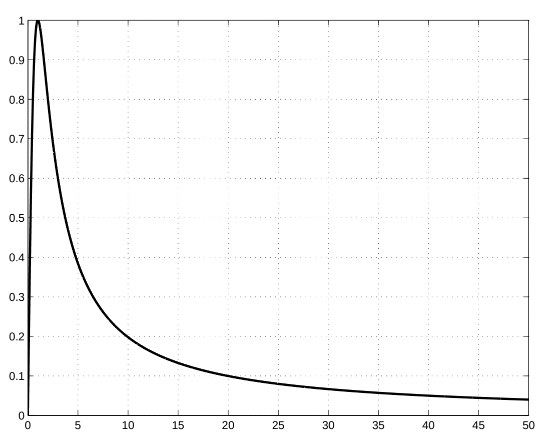

Figure 5: $1/\kappa_F(\mathbf{J}_B)$ vs. ν, $0 \le \nu \le 50$

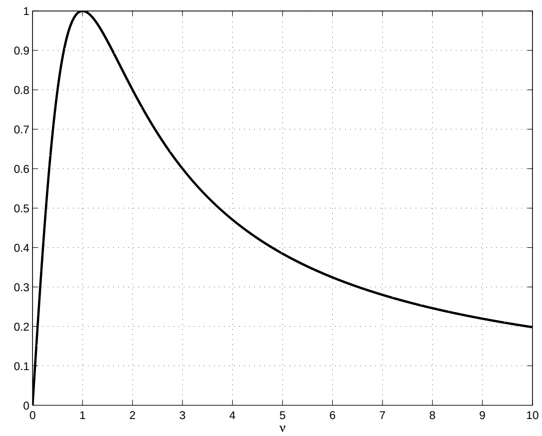

Figure 6: $1/\kappa_F(\mathbf{J}_b)$ vs. $0 \le \nu \le 10$

It is apparent that, if we make $\nu = 1$, then design B, the indirectly-driven wrist, becomes kinetostatically isotropic. Notice that design A corresponds to a *serial* layout, while design B to a *parallel* layout. Moreover, many industrial robots are designed with a wrist of type B in which the sun and the planet gears are identical, which means that robotics engineers are intuitively designing a kinestotatically isotropic, and hence, optimally robust wrist.

Elastostatic Performance

The elastostatic performance of the wrist is of interest when the wrist is subject to environment forces and torques while the actuators are locked at one given posture—given values of θ_1 and θ_2—which is either constant or variable with time, as when performing a path-tracking task. Under these conditions, the system becomes a linearly elastic structure. For brevity, we shall assume that only the joints of the system are flexible. For design A, moreover, we assume that joint i has a torsional stiffness of k_i, for $i = 1, 2$, where, in view of the cascaded array, most likely the first joint is stiffer than the second one, for it is carrying a bigger load and must fight against gravity. For design B we can safely assume that the two joints are identical and hence, have identical stiffnesses k.

The elastostatic relations now follow from Hooke's Law: Let $\mathbf{K}_A$ and $\mathbf{K}_B$ be the 2×2 stiffness matrices for designs A and B, respectively, and $\boldsymbol{\tau}$ the 2-dimensional vector of moments exerted by the environment on the wrist. Furthermore, we assume that, by virtue of the elasticity of the joints, the joint variables θ_1 and θ_2 undergo angular deflections u_1 and u_2, which are grouped within the 2-dimensional vector $\mathbf{u}$. We thus have

$$\mathbf{K}_A \mathbf{u} = \boldsymbol{\tau}, \quad \mathbf{K}_B \mathbf{u} = \boldsymbol{\tau} \tag{19a}$$

where

$$\mathbf{K}_A = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}, \quad \mathbf{K}_B = k\mathbf{1} \tag{19b}$$

Moreover, if we let τ_p and τ_r be the environment-exerted moments in the directions of the pitch and the roll axes, respectively, then

$$\boldsymbol{\tau} = [\, \tau_p \quad \tau_r \,]^T \tag{19c}$$

At this stage of the design, the design variables are k, k_1 and k_2, to be assigned values so as to render the design as robust as possible to disturbances in the form of static loads on the gripper.

After some routine calculations, we obtain, as values of the condition number of each stiffness matrix,

$$\kappa_F(\mathbf{K}_A) = \frac{k_1^2 + k_2^2}{2k_1 k_2}, \quad \kappa_F(\mathbf{K}_B) = \frac{1 + \nu^2}{2\nu} \tag{20}$$

If we assume that k_2 is a fraction $\alpha_k < 1$ of k_1, then $\kappa_F(\mathbf{K}_A)$ can be expressed in terms of parameter α, namely,

$$\kappa_F(\mathbf{K}_A) = \frac{1 + \alpha_k^2}{2\alpha_k} \tag{21}$$

which turns out to be of the same form as $\kappa_F(\mathbf{K}_B)$. Not only this. The expressions for the condition numbers of the two stiffness matrices are identical to those of $\kappa_F(\mathbf{J}_B)$. As a consequence, then, the two wrists can be designed with elastostatic isotropy if $\nu = 1$ or, correspondingly, $\alpha_k = 1$. However, while $\nu = 1$ is rather common in industrial robots, $\alpha_k = 1$ is impossible for design A because of the disparate values of the loads at which the pitch and the roll axes are subjected. As a consequence, then, design B is more robust than design A from the elastostatic performance viewpoint.

Elastodynamic Performance

Now we account for the behaviour of the wrist under "small" perturbations from the equilibrium state. Due to the inertia of the moving parts, and the elastic deformations experienced by the joints, as discussed above, the system undergoes parasitical vibrations that we want to keep at their lowest possible levels. In order to analyze the elastodynamic performance of the wrist, we resort to the usual *linarized* mathematical model, where we neglect damping. In this regard, we recall vector $\mathbf{u}$, as defined in the previous subsection , and denote with $\mathbf{M}_A$ and $\mathbf{M}_B$ the mass matrices of wrists A and B, respectively. The mathematical model at hand then becomes

$$\mathbf{M}_i \ddot{\mathbf{u}} + \mathbf{K}_i \mathbf{u} = \mathbf{0}, \quad i = A, B \tag{22}$$

In the foregoing model we assume, moreover, that the system is perturbed by a "small" displacement $\mathbf{u}(0)$ at an instant that is arbitrarily assigned the value $t = 0$. That is, $\mathbf{u}(0)$ denotes the initial condition of $\mathbf{u}(t)$.

The stiffness matrices $\mathbf{K}_A$ and $\mathbf{K}_B$ were derived in the previous subsection. In order to derive the mass matrices, we resort to some plausible assumptions:

For wrist A, we have only two moving rigid links, the first one rotating about the pitch axis, which is fixed to an inertial frame. Its moment of inertia about this axis will be labelled I_p; the second link rotates about the pitch and the roll axes, which we assume mutually perpendicular. We then assume that these two orthogonal axes are principal axes of inertia of the link. The principal moments of inertia of the second link about the pitch and the roll axes are labelled J_p and J_r, respectively. Routine calculations lead to

$$\mathbf{M}_A = \begin{bmatrix} I_p + J_p & 0 \\ 0 & J_r \end{bmatrix}$$

Furthermore, we let

$$J_p = \alpha_A I_p, \quad J_r = \beta_A I_p \tag{23a}$$

and hence,

$$\mathbf{M}_A = I_p \begin{bmatrix} 1+\alpha_A & 0 \\ 0 & \beta_A \end{bmatrix} \tag{23b}$$

Here, I_p, J_p and J_r are the design variables, to be determined so as to render the design as robust as possible for an elastodynamic performance.

For wrist B, in turn, we assume that the moments of inertia of the two sun gears are identical and label them I. We further assume that the pitch and the roll axes are principal axes of inertia of the body composed of the planet gear and the gripper, its moments of inertia about the pitch and the roll axes being labelled P and R, respectively. Likewise, routine calculations lead to

$$\mathbf{M}_B = \begin{bmatrix} I+P & 0 \\ 0 & I/\nu + R \end{bmatrix}$$

Furthermore, we let

$$P = \alpha_B I, \quad R = \beta_B I \tag{24}$$

Therefore,

$$\mathbf{M}_B = I \begin{bmatrix} 1+\alpha_B & 0 \\ 0 & 1/\nu + \beta_B \end{bmatrix} \tag{25}$$

The design variables for this wrist are I, P and R.

Now we cast the mathematical model of eq.(22) in a more suitable form. To this end, we first define

$$\mathbf{N}_i \equiv \sqrt{\mathbf{M}_i}, \quad i = A, B \tag{26}$$

Next, we introduce the change of variable

$$\mathbf{v} = \mathbf{N}_i \mathbf{u}, \quad \boldsymbol{\phi} = \mathbf{N}_i^{-1} \boldsymbol{\tau}, \quad i = A, B \tag{27}$$

the mathematical model thus reducing to a simpler form:

$$\ddot{\mathbf{v}} + \boldsymbol{\Omega}_i^2 \mathbf{v} = \mathbf{0}, \quad i = A, B \tag{28}$$

where matrix $\boldsymbol{\Omega}_i$ has units of frequency and is, hence, termed the *frequency matrix*. Upon rearranging the above mathematical model, we obtain

$$\ddot{\mathbf{v}} = -\boldsymbol{\Omega}_i^2 \mathbf{v}, \quad i = A, B \tag{29}$$

where we can readily identify, within our framework, $\mathbf{v}$ as the DEP vector and $\ddot{\mathbf{v}}$ as the performance function vector, matrix $-\boldsymbol{\Omega}_i^2$ thus being the performance matrix.

In the sequel, we shall assume that the design variables have been chosen so as to render matrices $\mathbf{J}_A$, $\mathbf{J}_B$ and $\mathbf{K}_B$ isotropic, although $\mathbf{K}_A$ cannot be assumed so because, as we discussed previously, this matrix cannot be rendered isotropic. After a few algebraic calculations, we obtain

$$\boldsymbol{\Omega}_A = \omega_A \begin{bmatrix} 1/\sqrt{1+\alpha_A} & 0 \\ 0 & \sqrt{\alpha_k/\beta_A} \end{bmatrix} \tag{30a}$$

with

$$\omega_A \equiv \sqrt{\frac{k_1}{I_p}}$$

and

$$\boldsymbol{\Omega}_B = \omega_B \begin{bmatrix} 1/\sqrt{1+\alpha_B} & 0 \\ 0 & \sqrt{\nu/(1+\nu\beta_B)} \end{bmatrix} \tag{30b}$$

with

$$\omega_B \equiv \sqrt{\frac{2k}{I}}$$

We can now obtain the condition number of the above frequency matrices. We thus have

$$\kappa_F(\mathbf{\Omega}_A^2) = \frac{1}{4}\frac{[\alpha_k(1+\alpha_A)+\beta_A]^2}{\alpha_k(1+\alpha_A)\beta_B}$$

and, if we introduce a new parameter γ_A as defined below:

$$\gamma_A \equiv \frac{\beta_A}{\alpha_k(1+\alpha_A)}$$

then the above expression for $\kappa_F^2(\mathbf{\Omega}_A^2)$ leads to

$$\kappa_F(\mathbf{\Omega}_A) = \frac{1+\gamma_A}{2\sqrt{\gamma_A}} \tag{31}$$

Likewise,

$$\kappa_F(\mathbf{\Omega}_B^2) = \frac{1}{4}\left(\frac{1}{1+\alpha_B}+\frac{1}{1+\beta_B}\right) + [(1+\alpha_B)(1+\beta_B)]$$

Let us introduce parameter γ_B in the form

$$\gamma_B \equiv \frac{1+\nu\beta_B}{1+\alpha_B}$$

the above expression thus reducing to

$$\kappa_F(\mathbf{\Omega}_B) = \frac{1+\gamma_B/\nu}{2\sqrt{\gamma_B/\nu}} \tag{32}$$

Note that both $\kappa_F(\mathbf{\Omega}_A^2)$ and $\kappa_F(\mathbf{\Omega}_B^2)$ bear the same gestalt as $\kappa_F(\mathbf{J}_B)$, $\kappa_F(\mathbf{K}_A)$ and $\kappa_F(\mathbf{K}_B)$. In this vein, to render $\mathbf{\Omega}_A^2$, and hence $\mathbf{\Omega}_A$, isotropic, we must have $\gamma_A = 1$, which implies

$$\beta_A = \alpha_k(1+\alpha_A)$$

If we recall the definitions of the above parameters, we obtain

$$\frac{J_r}{I_p} = \frac{k_2}{k_1}\left(1+\frac{J_p}{I_p}\right)$$

which can be rearranged to yield

$$\frac{k_1}{I_p+J_p} = \frac{k_2}{J_r}$$

It is now apparent that an isotropic frequency matrix for design A is possible in light of the assumptions and the physical constraints. Indeed, while $k_1 > k_2$, $I_p + J_p$ is also greater than J_r.

Likewise, the condition for isotropy of $\mathbf{\Omega}_B$ is $\gamma_B = 1$, which then leads to

$$1+\beta_B = 1+\alpha_B$$

i.e., $\beta_B = \alpha_B$, and hence, $P = R$ will make the frequency matrix of design B isotropic. Notice that this condition is possible to verify, for what it implies is that the composite body planet-gear + gripper must have its moments of inertia with respect to the centre O of the mechanism about the pitch and roll axes identical. Usually, the differential gear train of Fig. 2 has a symmetric layout, with a second planet diametrically located with respect to the one shown in the figure.

Epilogue

Apparently, the only difference between the two alternative designs lies in the impossibility of rendering the stiffness matrix of design A elastostatically isotropic. Otherwise, the two designs can be rendered both kinetostatically and elastodynamically sotropic. What the foregoing analysis does not show is a major difference between the two designs: The actuator of the first joint having to carry the mass of the second actuator, in addition to the payload, must be much bigger than the second actuator. This is a known fact affecting all serial manipulators. Design B is, in fact, a parallel manipulator, which is, additionally, rather common in industry. There is yet another aspect of robustness that has not been included in the foregoing analysis. The decision-making process leading to the two alternatives studied here, which pertains rather to globally robust design. Notice that design B is bound to be more robust than design A because, in the former, the two actuators can be chosen as identical; both are used to fight against gravity, and hence, for the same payload, design B requires smaller actuators than design A.

One more essential difference in the two designs lies in their kinematics: The directly-driven design reproduces, at the pitch axis, the joint rate of the first actuator, if attenuated by the speed reduction of the transmission mechanism, in case there is one. Likewise, the roll rate is a replica of the joint rate of the second actuator, if attenuated by the speed reduction, in case a transmission is used. Therefore, any errors in the joint rates will be directly transmitted to the pitch and the roll rates. On the contrary, in the indirectly-driven design, the pitch rate is the mean value of the joint rates of the two actuators, while the roll rate is proportional to the difference of the joint rates of the two actuators. Both pitch and roll rates thus benefit of the filtering effect of averaging. The conclusion is that design B is less sensitive to control errors in the actuators than design A.

Finally, while design A at a first glance verifies the Independence Axiom when kinetostatic performance is at stake, while design B does not, we showed that design B can be rendered optimally robust for this performance by a proper choice of parameter ν.

Conclusions

Design principles based on the concept of robustness were proposed, as pertaining to the design of mechanical elements and systems. We revised Taguchi's and Suh's principles and found that the two are reconcilable within a general framework, as proposed here. The concepts were illustrated with the aid of an example: the design of a pitch-roll wrist. The design of this wrist was studied for three performances: kinetostatic, elastostatic and elastodynamic.

Acknowledgements

The work reported here is supported by NSERC (Canada's Natural Sciences and Engineering Research Council) under Strategic Projects 215729-98 and 246488-01.

References

1 Taguchi, G., *Taguchi on Robust Technology Development: Bringing Quality Engineering Upstream*, ASME Press, NY, 1993.

2 Gadallah, M. and El-Maraghy, H., "A statistical optimization approach to robust design," in Utpal, R., Usher, J., Hamid, R. (editors), *Simultaneous Engineering*, Gordon and Breach Science Publishers, Singapore, pp. 135–158, 1999.

3 Zhu, J., and Ting, K.L. 2001, "Performance Distribution Analysis and Robust Design", *ASME Journal of Mechanical Design*, Vol. 123, pp. 11–17.

4 Suh, N.P., *Axiomatic Design. Advances and Applications*, Oxford University Press, Oxford, 2001

5 Wiener, N., 1948, *Cybernetics*, John Wiley & Sons, New York.

6 Golub, G. and Van Loan, C., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1989.