 ME 555

MONOTONICITY ANALYSIS AND MODEL REDUCTION

MONOTONITICITY ANALYSIS

Goals.

Identify models that are not well constrained.

Identify active constraints a priori.

Monotonicity Principles.

In many preliminary design models, due to simplification of trade-offs, the objective and constraint functions are *monotonic* with respect to the design variables.

A continuous differentiable function f(**x**) is strictly increasing (decreasing) with respect to (wrt) a design variable x_i, if $\partial f/\partial x_i > 0$ ($\partial f/\partial x_i < 0$). We say that f is coordinate-wise monotonic wrt x_i, or that x_i is a monotonic variable.

We assume all design variables are strictly positive.

Monotonicity analysis is based on two simple principles:

> *First Monotonicity Principle (MP1)*: In a well-constrained objective function every increasing (decreasing) variable is bounded below (above) by at least one active constraint.
>
> *Second Monotonicity Principle(MP2)*: Every monotonic variable not occurring in a well-constrained objective function is either irrelevant and can be deleted from the problem together with all constraints in which it occurs, or relevant and bounded by two active constraints, one from above and one from below.

Implicit Function Theorem.

Functional relations in a model usually must be rearranged in order to cast them into a standard form. In general, a given relation may be equivalent to several standard forms, some of which may appear monotonic and some not.

Monotonicity should be examined in the original formulation of the problem and if algebraic manipulations are employed, care should be taken not to disguise monotonic properties. However, once a relation has identified monotonicities, these are invariant in the sense that any standard form resulting in identified monotonicities will have exactly the same type of monotonicities.

The following implicit function theorem, specifically stated for monotone functions, is useful in studying how monotonicity properties may be inherited when equalities are eliminated through variable substitution.

Implicit Function Theorem. Let X_i, $i = 1, \ldots, n$, be n subsets (finite or infinite) of R and let $X = \{\mathbf{x}$, where $\mathbf{x} = (x_1, \ldots, x_n)^T$, $x_i \in X_i$, $i = 1, \ldots, n\}$. Let F: in X be a function coordinate-wise monotonic on X. Then, for each s in the range of F and for each $i = 1, \ldots,$ n, there exists a function $f(i,s; \mathbf{x}_i')$ of the variable vector $\mathbf{x}_i' = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ such that $f(i,s; \mathbf{x}_i')$ is (coordinate-wise) monotonic wrt x_i. Furthermore, for $1 \leq j \leq n$ and $i \neq j$, if F is monotonic in the same (opposite) sense wrt x_i and x_j, then $f(i,s; \mathbf{x}_i')$ is decreasing (increasing) wrt x_j.

Monotonicity Principles can be applied directly when the model contains only inequalities, i.e.,

minimize	$f(\mathbf{x})$	$\mathbf{x} \in X$
subject to	$g_j(\mathbf{x}) \leq 0$	$j = 1, 2, \ldots, m$

The procedure is as follows.

1. Examine all functions and identify variables that are monotonic.
2. Apply monotonicity principles to check for possible poor bounding.
3. If a variable is not bounded properly, re-examine the model to fix the problem. You may,
   (a) modify the objective and/or constraint functions to introduce better trade-offs;
   (b) add a new constraint that will provide the correct bound. This constraint will be subsequently identified as active.
4. If all variables are properly bounded, identify active constraints based on the monotonicity principles.
5. Eliminate at least one identified active constraint and a corresponding variable. Create a new model of reduced dimensionality.
6. Repeat the process (2) - (5) until no further reduction is possible.

Models with equalities that are not eliminated may be handled using the concept of "equality direction"—see the last section of this document.

Example A

Consider the problem

$$\begin{aligned}
\text{maximize} \quad & f = 0.0201\, d^4 w n^2 \\
\text{subject to} \quad & g_1 = d^2 w - 675 \leq 0 \\
& g_2 = d - 36 \leq 0 \\
& g_3 = n - 125 \leq 0 \\
& g_4 = n^2 d^2 - (0.419)(10^7) \leq 0 \\
& d, n, w \geq 0
\end{aligned} \tag{1}$$

This problem involves maximizing the stored energy in a flywheel and is reduced to the above mentioned form after some manipulations [Siddall 1972].

Note that the standard negative null form is obtained by using the objective {minimize -f}.

Constraint g_1 is active by MP1 wrt w. Elimination of w from the objective gives f = (0.0201) (675)d^2n^2. Clearly g_4 will be active except if g_2, g_3 combined impose a stricter bound on d^2n^2. For the given numbers this is not the case, so there are infinite solutions

$$\max f^* = (5.68)(10^7)$$
$$w_* = 675\, d_*^{-2} \qquad n_* = 2047\, d_*^{-1} \qquad 16.4 \le d_* \le 36 \tag{2}$$

Example B

Consider now the problem

$$\begin{aligned} \text{minimize} \quad & f = x_1^{-2} + x_2^{-2} + x_3^{-2} \\ \text{subject to} \quad & g_1 = 1 - x_1 - x_2 - x_3 \le 0 \\ & g_2 = x_1^2 + x_2^2 - 2 \le 0 \\ & g_3 = 2 - x_1 x_2 x_3 \le 0 \\ & x_i \ge 0, \ \ i = 1, 2, 3 \end{aligned} \tag{3}$$

By MP1 constraint g_2 must be active providing upper bounds on x_1 and x_2.
However, x_3 is unbounded from above since f, g_1, and g_3 are all decreasing wrt x_3.

The problem has no solution unless the objective and/or constraint functions are appropriately modified by remodeling. Any nonredundant equality constraint would also serve.

The usual practice in such cases is to add a simple inequality constraint, e.g., $x_3 \le a$, $a > 0$. However, this constraint will be always active by construction, which means that the optimum is in fact determined by this artificial bound.

In real problems, such information must be consciously considered because the optimum is essentially arbitrarily fixed by the modeler.

LINEAR ACTUATOR EXAMPLE

In this extensive example we will go through various modeling steps. The procedure followed is typical for models with relatively small number of variables and explicit functions, but can be used to advantage in other situations as well.

Model Setup

A drive screw is part of a power module assembly that converts rotary to linear motion, so that a given load is linearly oscillated at a specified rate. Such a device is used in some household appliances. The assembly consists of an electric motor, drive gear, pinion (driven) gear, drive screw, load-carrying nut, and a chassis providing bearing surfaces and support. The present example addresses only the design of the drive screw, schematically shown in Figure 1.

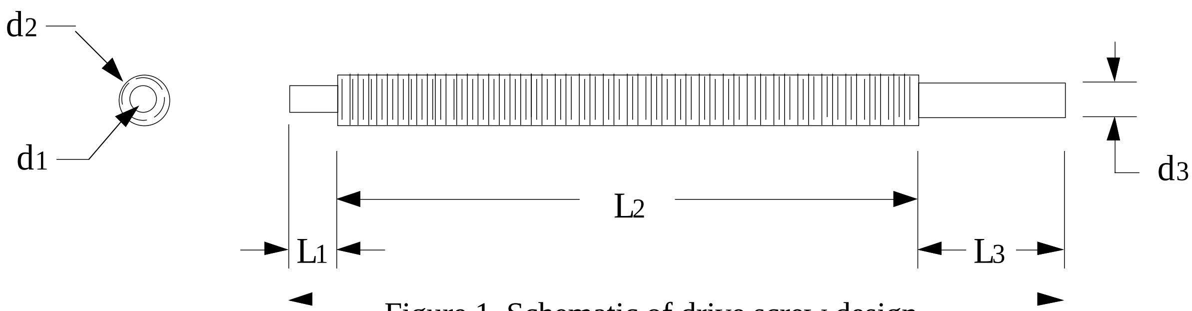

Figure 1. Schematic of drive screw design

The objective function in the design model is to minimize product cost consisting of material and manufacturing costs. Machining costs for a metal drive screw or injection molding costs for a plastic one are considered fixed for relatively small changes in the design, hence only material cost is taken as the objective to minimize, namely

$$f_o = (C_m \pi/r)(d_1^2 L_1 + d_2^2 L_2 + d_3^2 L_3). \tag{4}$$

Here C_m is the material cost ($/lb), d_1, d_2 and d_3 are the diameters of gear/drive screw interface, threaded and bearing surface segments, respectively, L_1, L_2 and L_3 being the respective segment lengths.

There are operational, assembly, and packaging constraints. For strength against bending during assembly we set

$$Mc_1/I \leq \sigma_{all}, \tag{5}$$

where the bending moment $M = F_a L/2$, F_a being the force required to snap the drive screw into the chassis during assembly, and L being the total length of the component

$$L = L_1 + L_2 + L_3. \tag{6}$$

Furthermore, $c_1 = d_2/2$, $I = \pi d_2^4/64$ is the moment of inertia, and σ_{all} is the maximum allowable bending stress for a given material.

During operation, a constraint against fatigue failure in shear must be imposed:

$$KTc_3/J \leq \tau_{all}. \tag{7}$$

Here, K is a stress concentration factor, T is the applied torque, $C_3 = d_1/2$, $J = \pi d_1^4/32$ is the polar moment of inertia, and τ_{all} is the maximum allowable shear stress. The torque is computed from the equation

$$T = T_m C_2 N_S/N_m, \tag{8}$$

where T_m is the motor torque, $c_2 = 1/16$ (lb/ounce) is a conversion factor, N_s and N_m is the number of teeth on the screw (driven) and motor (drive) gear, respectively.

To meet the specified linear cycle rate of oscillation, a speed constraint is imposed:

$$c_4 N_m S_m / N_s N_T \leq S \tag{9}$$

where $c_4 = 60^{-1}$ (no. of threads/rev)(min/sec) is a conversion factor, S_m is the motor speed (rpm), N_T is the number of threads per inch, and S is the specified linear cycle rate (in/sec).
In order for the screw to operate in a drive mode the following constraint must be satisfied [Juvinall 1983]

$$(Wd_2/2)\left[\frac{\pi f d_2 + N_T^{-1}\cos a_n}{\pi d_2 \cos a_n - N_T^{-1} f}\right] \leq T \tag{10}$$

Here W is the drive screw load, f is the friction coefficient, N_T^{-1} is the lead of screw threads, and α_n is the thread angle measured in the normal plane. There is also an upper bound on the number of threads per inch imposed by mass production considerations,

$$N_T \leq 24. \tag{11}$$

From gear design considerations, particularly avoidance of interference, limits on the numbers of gear teeth are imposed.

$$N_m \geq 8 \qquad\qquad N_S \leq 52 \tag{12}$$

Finally, there are some packaging and geometric considerations that impose constraints:

$$8.75 \leq L_1 + L_2 + L_3 \leq 10.0 \tag{13}$$

$$7.023 \leq L_2 \leq 7.523 \tag{14}$$

$$1.1525 \leq L_3 \leq 1.6525 \tag{15}$$

$$d_1 \leq d_2,\ d_3 \leq d_2,\ d_2 \leq 0.625. \tag{16}$$

Note that several assumptions were invoked in the model above: manufacturing costs remain fixed; a high volume production is planned; Standard Unified Threads are used; the assembly force for the drive screw is concentrated at the midpoint; frictional forces are considered only between the threads and the load nut, and all other frictional forces are assumed negligible.

Model Validity Constraints

During the early stages of developing a mathematical optimization model many assumptions are made in order to obtain reasonably simple expressions for the objective and constraint functions. One must always check whether subsequent results from optimization conform to these assumptions, lest they are violated. This would indicate that the model used is inappropriate for the optimal design obtained and the optimization results are at least suspect and possibly erroneous. The remedy is usually a more accurate, probably also more complicated, mathematical model of the phenomenon under question. For example, Equation (5) is valid only if the length/diameter ratio is more than ten.

Material Choice as a Parameter

A final observation on the initial drive screw model is that a significant trade-off exists on the choice of material: stainless steel vs. plastic. A steel screw will have higher strength and be smaller in size but will require secondary processing, such as rolling of threads and finishing of bearing surfaces. A plastic screw would be made by injection molding in a one-step process that is cheaper but more material would be used due to lower strength, plastic having a higher cost per pound than steel. Specialty plastics with high strength would be even more expensive and less moldable. Thus the choice of material must be based on a model that contains more information than the current one. The constant term representing manufacturing costs should be included in the objective. Indeed a more accurate cost objective should include capital investment costs for manufacturing.

Nevertheless, substantial insight can be gained from the present model if we include material as a *parameter*; in fact each material is represented by four parameters C_m, σ_{all}, τ_{all}, f. In the model analysis that follows we keep these parameters in the model with their symbols, rather than giving numerical values in so far as possible. The goal is to derive as many additional results as possible independently of the material used. This will substantially facilitate a post-optimal parametric study on the material. It would be much more difficult to treat material as a variable, because then we would have four additional variables with discrete values and implicitly linked, perhaps through a table of material properties. As mentioned earlier, this would destroy the continuity assumed in nonlinear programming formulations.

Standard Null Form

The model is now summarized in the negative null form, all parameters represented by their numerical values (Table 1), except for material parameters. All "intermediate" variables defined through equalities are eliminated together with the associated equality constraints by direct substitution. This elimination should be always done when possible, in order to arrive at a model with only inequality constraints, thus facilitating subsequent monotonicity analysis. It is a *model reduction* step, since the number of design variables is reduced.

Table 1

List of Parameters (Material values for stainless steel)

C_m	material cost ($/lb)
f	friction coefficient (0.35 for steel on plastic)
F_a	force required to snap drive screw into chassis during assembly (6 lb)
K	stress concentration factor (3)
L_1	length of gear/drive screw interface segment (0.405 in)
S	linear cycle rate (0.0583 in/sec)
S_m	motor speed (300 rpm)
T_m	motor torque (2 in-ounces)
W	drive screw load (3 lb)
α_n	thread angle in normal plane (60°)
σ_{all}	maximum allowable bending stress (20,000 psi)
τ_{all}	maximum allowable shear stress (22,000 psi)

Note however that this may not be always a *model simplification* step, as the resulting expressions may become more complex with undetermined monotonicities. Some judgement must be exercised here. Occasionally, "directing" an equality may be useful (see further below) in avoiding direct elimination.

In the model below the intermediate variables M, L, T, I, c_1, c_3, and J together with the corresponding defining equalities have been eliminated. The variables are d_1, d_2, d_3, L_2, L_3, N_m, N_S, N_T.

MODEL 1

$$\text{minimize } f_o = (C_m\pi/4)\,(0.405\, d_1^2 + L_2\, d_2^2 + L_3\, d_3^2)$$

subject to

$$g_1 = 38.88 + 96\, L_2 + 96\, L_3 - \pi\sigma_{all}\, d_2^3 \le 0$$

$$g_2 = 6\,(N_s/N_m) - \pi\,\tau_{all}\, d_1^3 \le 0$$

$$g_3 = 8.345 - L_2 - L_3 \le 0 \qquad g_4 = -9.595 + L_2 + L_3 \le 0$$

$$g_5 = L_2 - 7.523 \le 0 \qquad g_6 = 7.023 - L_2 \le 0$$

$$g_7 = L_3 - 1.6525 \le 0 \qquad g_8 = 1.1525 - L_3 \le 0$$

$$g_9 = d_2 - 0.625 \le 0 \qquad g_{10} = d_3 - d_2 \le 0$$

$$g_{11} = d_1 - d_2 \le 0 \qquad g_{12} = 5\,(N_m/N_S) - 0.0583\, N_T \le 0$$

$$g_{13} = 1.5\, d_2 \left(\frac{f\pi\, d_2 + 0.5\, N_T^{-1}}{0.5\,\pi\, d_2 - f\, N_T^{-1}} \right) - 0.125\,(N_S/N_m) \le 0$$

$$g_{14} = N_T - 24 \le 0 \qquad g_{15} = 8 - N_m \le 0$$

$$g_{16} = N_S - 52 \le 0 \qquad (17)$$

As there are no equality constraints, there are eight degrees of freedom corresponding to the eight design variables. Three of these variables, N_m, N_T and N_S, must take integer values, so this problem is in fact a mixed continuous-integer variable nonlinear programming problem and standard numerical NLP methods will not work. We will see later how this is dealt with in the particular example.

Feasibility Checking

Before embarking on analyzing the model, it is a good idea to check that the feasible domain is not empty, i.e., there exists at least one proven feasible point. From a mathematical viewpoint this may be a hard problem (possibly as hard as the optimization itself), but from an engineering viewpoint past experience can be a guide. In the drive screw example, an existing design using stainless steel has the following values: d_1 = 0.1875 in, d_2 = 0.3125 in, d_3 = 0.2443 in, L_2 = 7.273 in, L_3 = 1.4025 in, N_m = 8, N_S = 48, N_T = 18. The design is feasible with an objective function value of f_o = 0.635 C_m. Now we can proceed knowing that an optimization attempt is possible.

Monotonicity

Looking at Model 1 one notes that most constraints were written in a form that requires no divisions. This is always advisable, since in subsequent numerical processing a denominator may become zero and cause an abrupt termination by overflow error. This can happen even if the imposed constraints exclude the relevant variable values, because many numerical algorithms will temporarily operate in the infeasible domain. In Model 1, constraint g_{13} has not been rewritten yet because of concern that this might obscure its monotonicity wrt d_2 and N_T. Let us examine this more carefully. Assuming a strictly positive denominator of the first tern in g_{13} we multiply both sides by it and collect terms:

$$g_{13}: \ 1.5 f\pi d_2^2 + 0.75 N_T^{-1} d_2 - 0.0625 \pi (N_S/N_m) d_2 + 0.125 f (N_S/N_m) N_T^{-1} \leq 0 \quad (18)$$

Clearly g_{13} decreases wrt N_T but is not monotonic wrt d_2. In fact,

$$\partial g_{13}/\partial d_2 = 3 \pi f d_2 + 0.75 N_T^{-1} - 0.0625 \pi (N_S/N_m) \quad (19)$$

that can be positive or negative depending on the variable values. All remaining functions in the constraints have obvious monotonicities.

Variable Transformation

The model can be further simplified by a variable transformation. We observe that the two variables N_S, N_m appear together as a ratio everywhere except in the simple bounds g_{15}, g_{16}. We can define a new variable R,

$$R = N_m/N_S, \quad (20)$$

which indeed is the reduction ratio of the gear drive, and eliminate variable N_m using $N_m = RN_S$.

The new model, including the reformulated constraint g_{13}, is as follows.

MODEL 2

$\min f_o = (C_m\pi/4)(0.405 d_1^2 + L_2 d_2^2 + L_3 d_3^2)$

subject to:

$g_1 = 38.88 + 96 L_2 + 96 L_3 - \pi\sigma_{all} d_2^3 \leq 0$

$g_2 = 6 - \pi \tau_{all} R d_1^3 \leq 0$

$g_3 = 8.345 - L_2 - L_3 \leq 0$ $\quad g_4 = -9.595 + L_2 + L_3 \leq 0$

$g_5 = L_2 - 7.523 \leq 0$ $\quad g_6 = 7.023 - L_2 \leq 0$

$g_7 = L_3 - 1.6525 \leq 0$ $\quad g_8 = 1.1525 - L_3 \leq 0$

$g_9 = d_2 - 0.625 \leq 0$ $\quad g_{10} = d_3 - d_2 \leq 0$

$g_{11} = d_1 - d_2 \leq 0$ $\quad g_{12} = 5 R - 0.0583 N_T \leq 0$

$g_{13} = 1.5 \pi f R d_2^2 + 0.75 R N_T^{-1} d_2 - 0.0625 \pi d_2 + 0.125 f N_T^{-1} \leq 0$

$g_{14} = N_T - 24 \leq 0$ $\quad g_{15} = 8 - R N_S \leq 0$

$g_{16} = N_S - 52 \leq 0$ $\quad (21)$

Note that the requirement of integer values for N_m is now converted to one of rational values for R.

Table 2
Monotonicity Table for Model 2 (with Model Repairs in Parentheses)

Variables Functions	d_1	d_2	d_3	L_2	L_3	R	N_S	N_T
f_o	+	+	+	+	+			
g_1		-		+	+			
g_2	-					-		
g_3				-	-			
g_4				+	+			
g_5				+				
g_6				-				
g_7					+			
g_8					-			
g_9		+						
(g_{10})		-	+					
g_{11}	+	-						
g_{12}						+		-
g_{13}		**U**				+		-
g_{14}								+
g_{15}						-	-	
g_{16}							+	
(g_{17})			-					
(g_{18})		-	+					-

Repairing A Model

Model 2 is now used to perform the first cycle of monotonicity analysis. The Monotonicity Table is a convenient tool to do this, see Table 2. The columns are the design variables and the rows are the objective and constraint functions, the entries in the table being the monotonicities of each function with respect to each variable. Positive (negative) sign indicates increasing (decreasing) function, U indicates undetermined or unknown monotonicity. An *empty* entry indicates that the function does not depend on the respective variable, so the table acts also as an incidence table. (Items in parentheses will be explained in the next subsection.)

Monotonicity Principles can be quickly applied by inspection using the Monotonicity Table. Looking at Table 2, by MP1 wrt d_3 we see that Model 2 is not well constrained because no lower bound exists for d_3. Note that $d_3 > 0$ is not an appropriate bound because of the *strict* inequality. If the model were treated numerically as is, no convergence would occur if the algorithm was successful or an erroneous result would be found if the algorithm was lead astray. Examining the engineering meaning of this model deficiency we see that an adequate thrust surface must be provided to keep the shaft from wearing into the bearing support, so we accept the simple remedy of adding a new constraint

$$g_{17} = 0.1875 - d_3 \leq 0 \tag{22}$$

Poor boundedness is cause for concern, so the above deficiency triggers also a closer examination of constraint g_{10}, the only other one containing d_3: $d_3 \leq d_2$. Upon reflection, constraint activity for g_{10} should not be allowed, i.e., $d_2 > d_3$. Examining the geometry of the screw more closely a new constraint is discovered

$$g_{18}:\ d_2 \geq d_3 + 1.2990/N_T \qquad (23)$$

Optimality Rules

At this point we decide to add constraints $g_{17}(d_3^-)$ and $g_{18}(d_2^-, d_3^+, N_T^-)$ to the model and delete g_{10} as redundant. These changes are shown in parentheses in Table 2. Applying monotonicity analysis to this model we now obtain the following results, which represent necessary rules for optimality.

(R1) By MP1 wrt d_1, g_2 is active.

(R2) By MP1 wrt d_2, at least one constraint from the set $\{g_1, g_{11}, g_{13}\}$ is active.

(R3) By MP1 wrt L_2, at least one constraint from the set $\{g_3, g_6\}$ is active.

(R4) By MP1 wrt L_3, at least one constraint from the set $\{g_3, g_8\}$ is active.

(R5) By MP2 wrt R, either all constraints g_2, g_{12}, g_{13} and g_{15} are inactive or at least one from each of the sets $\{g_2, g_{15}\}$, $\{g_{12}, g_{13}\}$ is active.

(R6) By MP2 wrt N_S, either g_{15} and g_{16} are both active or they are both inactive.

(R7) By MP2 wrt R, if g_2 is active then at least one of $\{g_{12}, g_{13}\}$ is active. Then, by MP2 wrt N_T and (R1), g_{14} is active.

The original eight degrees of freedom have now been reduced to five, because of the identified active constraints:

$$g_2:\ \pi\,\tau_{all}\,R\,d_1^3 = 6 \qquad g_{14}:\ N_T = 24$$
$$g_{17}:\ d_3 = 0.1875 \qquad (24)$$

The remaining rules above give only *conditional* activity and in order to identify a single constraint as active *dominance* arguments are required. We will proceed with these later below. Thus, due to (R1), (R5) must be modified as:

(R5') By MP2 wrt R, at least one constraint from the set $\{g_{12}, g_{13}\}$ is active.

This interaction was also used in deriving (R7).

Active Constraint Elimination

The three active constraints, Eq. (24), are used to eliminate three variables from Model 2, namely d_1, d_3 and N_T. There are two reasons for this. One is that monotonicity analysis on the new reduced model may reveal additional activity requirements. Another is that dominance arguments will be simpler in the reduced model. Which variables to eliminate

Table 3
Monotonicity Table for Model 3

Variables / Functions	d_2	L_2	L_3	R	N_S
f_o	+	+	+	-	
g_1	-	+	+		
g_3		-	-		
g_4		+	+		
g_5		+			
g_6		-			
g_7			+		
g_8			-		
g_9	+				
g_{11}	-			-	
g_{12}				+	
g_{13}	U			+	
g_{15}				-	-
g_{16}					+
g_{18}	-				

is a judicious choice, based on what may be algebraically simpler and what may be a desirable form of the reduced model.

The new model is as follows:

MODEL 3

minimize $f_o = 0.25\,\pi\,C_m\,[(0.405\,(6/\pi\,\tau_{all}\,R)^{2/3} + L_2\,d_2^2 + L_3\,(0.1875)^2)]$

subject to:

$g_1 = 38.88 + 96\,L_2 + 96\,L_3 - \pi\sigma_{all}\,d_2^3 \le 0$

$g_3 = 8.345 - L_2 - L_3 \le 0$ $\qquad g_4 = -9.595 + L_2 + L_3 \le 0$

$g_5 = L_2 - 7.523 \le 0$ $\qquad g_6 = 7.023 - L_2 \le 0$

$g_7 = L_3 - 1.6525 \le 0$ $\qquad g_8 = 1.1525 - L_3 \le 0$

$g_9 = d_2 - 0.625 \le 0$ $\qquad g_{11} = (6/\pi\,\tau_{all}\,R)^{1/3} - d_2 \le 0$

$g_{12} = R - 0.2798 \le 0$

$g_{13} = 1.5\,\pi\,f\,R\,d_2^2 + 0.0313\,R\,d_2 - 0.0625\,\pi\,d_2 + 0.0052\,f \le 0$

$g_{15} = 8 - R\,N_S \le 0$ $\qquad g_{16} = N_S - 52 \le 0$

$g_{18} = 0.2416 - d_2 \le 0$ (25)

The Monotonicity Table for Model 3 is shown in Table 3. No new results are obtained from this table. So dominance arguments must be sought in order to clarify the previously stated conditional activities and obtain further model reduction.

Activity Map and Dominance

Consider constraints g_3, g_6 and g_8 and the derived rules (R3) and (R4). An *activity map*, as shown in Figure 2, can assist in dominance analysis. In this map all possible activity combinations for the three constraints are examined. Only constraint numbers are shown for simplicity, the overbar on a number indicating *not* active constraint and a plain number indicating an active one. The cross-hatched areas indicate combinations that are not possible.

In Figure 2, three combinations are excluded since they violate rules (R3) and/or (R4), as indicated. The combination 368 is excluded by the Maximal Activity Principle [Papalambros and Wilde 1988],which basically says that the number of active constraints cannot exceed the number of variables in them. Finally, if g_3 is inactive then both g_6 and g_8 must be active giving $L_2 = 7.023$, $L_3 = 1.1525$, and $L_2 + L_3 = 8.1755$, which violates g_3. Hence this case is infeasible and excluded.

From Figure 2, it is now obvious that g_3 must be active irrespectively of the activity of g_6 and g_8. This also makes g_4 and at least three constraints from the set $\{g_5, g_6, g_7, g_8\}$ be inactive. The conditional inactivity can be resolved easily as L_2, L_3 appear only in a small part of the model.

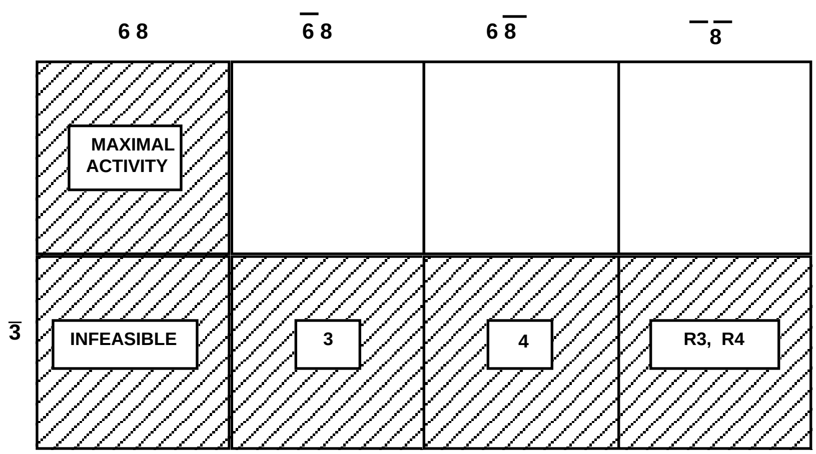

Figure 2. Activity map for constraints g_3, g_6 and g_8.

Consider eliminating L_2 $(= 8.345 - L_3)$ from the model. The submodel containing L_2 and L_3 becomes now:

$$\text{minimize } f_o = 0.25\pi C_m [0.405\,(6/\pi\tau_{all} R)^{2/3} + 8.345\, d_2^2 + L_3\{(0.1875)^2 - d_2^2\}]$$

$$g_1 = 38.88 + 96\,(8.345) - \pi\sigma_{all}\, d_2^3 \le 0$$

$$g_5 = 0.822 - L_3 \le 0 \qquad g_6 = L_3 - 1.322 \le 0$$

$$g_7 = L_3 - 1.6525 \le 0 \qquad g_8 = 1.1525 - L_3 \le 0 \qquad (26)$$

In the objective, the monotonicity wrt L_3 is determined by the sign of the quantity

$$(0.1875)^2 - d_2^2 \;, \qquad (27)$$

Table 4
Monotonicity Table for Model 4

Variables / Functions	d_2	R	N_S
f_o	+	-	
g_1	-		
g_9	+		
g_{11}	-	-	
g_{12}		+	
g_{13}	U	+	
g_{15}		-	-
g_{16}			+
g_{18}	-		

which is negative since $d_2 \geq 0.2416$ from g_{18}. Hence, the objective is decreasing wrt L_3 and an upper bound is required by MP1. Constraint g_6 is obviously the dominant and active one, so

$$L_{3*} = 1.322, \quad L_{2*} = 7.023 \tag{28}$$

and g_5, g_7, g_8 are inactive.

The above results lead to yet another further reduced model with only three degrees of freedom:

MODEL 4

minimize $f_o = 0.25\pi\, C_m[(0.405(6/\pi\, \tau_{all}\, R)^{2/3} + 7.023\, d_2^2 + 1.322(0.1875)^2)]$

subject to:

$g_1 = 840 - \pi\sigma_{all}\, d_2^3 \leq 0$ $\qquad g_9 = d_2 - 0.625 \leq 0$

$g_{11} = (6/\pi\, \tau_{all}\, R)^{1/3} - d_2 \leq 0$ $\qquad g_{12} = R - 0.2798 \leq 0$

$g_{13} = 1.5\, \pi\, f\, R\, d_2^2 + 0.0313\, R\, d_2 - 0.0625\, \pi\, d_2 + 0.0052\, f \leq 0$

$g_{15} = 8 - R\, N_S \leq 0$ $\qquad g_{16} = N_S - 52 \leq 0$

$g_{18} = 0.2416 - d_2 \leq 0$ (29)

The monotonicity table for this model is shown in Table 4. Rules (R2), (R5') and (R6) are still the only results derived from monotonicity principles.

At this point, the unknown monotonicity of g_{13} wrt d_2 prevents us from continuing the reduction process. Indeed,

$$\partial g_{13}/\partial d_2 = 3\pi f\, R d_2 + 0.0313\, R - 0.0625\, \pi \tag{30}$$

and for g_{13} to be increasing wrt d_2 we would need to have (for $f = 0.34$)

$$d_2 \geq 0.0595\, R^{-1} - 0.0033 \triangleq D_2\,(R^-) \tag{31}$$

for the entire feasible range of d_2, R. We note that

$$\begin{aligned} \max D_2\,(R^-) &= 0.0595\, R^{-1}_{min} - 0.0033 = (0.0595\, N_{Smax}/8) - 0.0033 \\ &= (0.0595)(52)/8 - 0.0033 = 0.3983 \end{aligned} \tag{32}$$

This last number falls within the known feasible range of d_2, $0.2416 \leq d_2 \leq 0.625$, so g_{13} appears really nonmonotonic in the feasible domain.

Parametric Models and Case Decomposition

The idea of *regional monotonicity* could be used here, i.e., try to identify in what interval values of d_2, g_{13} is monotonic and examine each case separately, comparing the results at the end. This would be unnecessarily complicated here. Instead, a simple problem decomposition can be applied: Case A with g_{13} inactive, and Case B with g_{13} active. We can examine these two cases separately and compare the results.

Before we proceed with the cases, it is instructive to recast Model 4 in a simplified parametric form, by introducing the parameters K_o', K_i, i = 0, 1, ..., 4, and rearranging (Table 5).

Table 5
Parameter Definitions for Model 5

K_o'	$= (0.25\,\pi\, C_m)$
K_o	$= K_o'(0.405)(6/\pi\, \tau_{all})^{2/3}$
K_1	$= (840/\pi\, \sigma_{all})^{1/3}$
K_2	$= (6/\pi\, \tau_{all})^{1/3}$
K_3	$= 1.5\,\pi\, f$
K_4	$= 0.0052\, f$

The revised model is as follows. Note that all parameters K_i relate to material properties.

MODEL 5

minimize $f_o = K_o R^{2/3} + K_o'(7.023\, d_2^2 + 1.322\,(0.1875)^2)$

subject to:

$$\begin{aligned}
&g_1 = K_1 - d_2 \leq 0 && g_9 = d_2 - 0.625 \leq 0 \\
&g_{11} = K_2 R^{1/3} - d_2 \leq 0 && g_{12} = R - 0.2798 \leq 0 \\
&g_{13} = K_3 R d_2^2 + 0.0313\, R d_2 - 0.1963\, d_2 + K_4 \leq 0 && \\
&g_{15} = 8 - R N_S \leq 0 && g_{16} = N_S - 52 \leq 0 \\
&g_{18} = 0.2416 - d_2 \leq 0 &&
\end{aligned} \tag{33}$$

Consider now Case A with g_{13} inactive. Then g_{12} is active from rule (R5') and $R_* = 0.2798$. From rule (R2) we now have

$$d_{2*} = \max \{(K_1, K_2 R_*^{1/3}, 0.2416)\}, \tag{34}$$

and from rule (R6) we have

$$8/R_* \leq N_{S*} \leq 52 \tag{35}$$

Note that one degree of freedom remains at the optimum, as N_m. N_S can be selected to give the largest rational number not exceeding 0.2798 and satisfying the range in Eq.(35). Any solution thus obtained must be checked that it satisfies the remaining inactive constraints. This *feasibility* check is frequently overlooked in the application of monotonicity analysis leading to erroneous conclusions.

Next, consider Case B with g_{13} active. Rules (R2) and (R5') are satisfied and no new activity results can be obtained. Note that now, *locally*, g_{13} must be decreasing wrt d_2, i.e., $g_{13}(d_2^-, R^+) \leqq 0$, while $f_o(d_2^+, R^-)$. An implicit solution of g_{13} gives $d_2 = g_{13}(R^+)$ and substitution in the objective gives $f_o(d_2^+, R^-) = f_o(g_{13}^+(R^+), R^-) = f_o(R)$, with R having unclear monotonicity. There are two degrees of freedom left, but a one-dimensional search in d_2 would suffice if g_{13} is solved explicitly for R and the objective is expressed as a function of d_2 only. Constraints g_{15} and g_{16} could be replaced by

$$R \geq 8 N_S^{-1} \geq 8/52 = 0.1538 \tag{36}$$

The results obtained numerically using the SQP code NLPQL [Schittkowski 1984] with $C_m = 10.0$ and continuous values for the integer variables indicate constraints $\{g_2, g_3, g_6, g_{14}, g_{17}\}$ and $\{g_{12}, g_{18}\}$ are active. The activity of the first five was the one discovered a priori by analysis. Constraints $\{g_1, g_4, g_5, g_7, g_8, g_9, g_{10}, g_{11}, g_{13}, g_{15}, g_{16}\}$ are inactive.

Directing An Equality

It is interesting to note that g_{18} is active in the final solution. Indeed according to Spotts [1985] the required relation is indeed an equality

$$h_{10}: \; d_3 = d_2 - 1.2990/N_T \tag{37}$$

rather than the inequality Eq. (23). Instead of using g_{18} in the model we could have used h_{10}, which would prevent direct application of MP1.

However, an equality constraint can be viewed as an active inequality that has been "properly directed" One way to determine such a direction is to use the monotonicity principles. Ignoring Equation (22) for the moment, and with Equation (37) replacing g_{10}, we see that by MP1 wrt d_3 a lower bound is required. Then h_{10} can provide this bound if directed as

$$d_3 \geqq d_2 - 1.2990\, N_T^{-1} \quad \text{or} \quad d_2 - 1.2990\, N_T^{-1} - d_3 \leqq 0. \tag{38}$$

Not much inference is possible when the directed h_{10} is used in the model instead of g_{18}. One could proceed by assuming h_{10} inactive and include g_{17}, which would then be active. This would take us essentially through the same steps as before, checking h_{10} for violation in the final solution. Interestingly, numerical results obtained for such a scenario using NLPQL indicate h_{10} is satisfied in the final solution with a zero multiplier value.

FINAL CAUTION:
There is a lingering concern regarding the physical meaning of g_{13} being active. Essentially, friction forces and applied forces on an equivalent inclined plane would be equal and motion would be impending. This would not represent a stable design for the lead screw, albeit possibly an optimal one. The designer must then examine the implications on the appropriateness of the model and/or the parameter values selected.

Also, satisfaction of model validity for the beam stress formula should be checked.

THIS MATERIAL WAS ADAPTED FROM
Papalambros, P., "Model Reduction and Verification Techniques", in *Advances in Design Optimization*, H. Adeli (ed.), Chapman and Hall, 1994.