EXPERIENCES WITH OPTIMIZERS IN STRUCTURAL DESIGN
Adaptive Computing in Engineering Design and Control - Pymouth, U.K., Sept. 1994

A.J. Keane

Lecturer in Structural Dynamics, Department of Engineering Science,
University of Oxford, Parks Road, Oxford, OX1 3PJ, U.K.
E-mail: andy.keane@eng.ox.ac.uk

ABSTRACT

The control of structural vibration in cars, aeroplanes, ships, etc., is of great importance in achieving low noise targets. Currently, such control is effected using viscoelastic coating materials although much current research is concerned with active, anti-noise based control measures. This paper is concerned with a third approach: that of using advanced structural analysis methods and modern optimization techniques to produce designs with inherently superior vibration performance. It gives a brief review of a ten year programme of work, illustrating some of the difficulties that have been encountered and the progress that has been made. Central to this process has been the availability of a library of optimization techniques that can be readily applied to the problems of interest. In recent years this suite has been expanded to include a number of evolutionary computing methods such as the genetic algorithm, simulated annealing, evolutionary programming, etc. These have shown significant advantages over more classical methods.

By way of example this paper considers three rather different problems: an extremely simple structure consisting of two rods and a spring, a moderately complex satellite boom and a computationally efficient but optimizationally hard, test problem that allows study and comparison of the various optimization methods available.

INTRODUCTION

The role of computational structural analysis is fundamental to the design of all modern cars, aircraft, satellites, ships, offshore marine installations, etc. It is now common-place to design and build highly complex structures that achieve startlingly good performance while withstanding the worst that nature can throw at them. However, it is inevitable that, as current problems are tackled and brought under control, higher standards are desired both by operators and regulators. Initially, structural integrity was the main design criterion, now this is considered alongside reduced through life cost, ease of maintenance, high standards of habitability, etc. One of the consequences of this shift in horizon is the need for new design and analysis tools to be brought on stream to deal with problems that have, until recently, not been seen as high priorities.

One area that is becoming increasingly important in all engineering spheres is that of noise and vibration control. Many engineering structures suffer from exposure to noise and vibration sources. These often excite unwanted structural vibrations causing damage or the transport of vibrational energy to distant parts of the structure where it cannot be tolerated. For example, the engines in ships and aircraft almost always vibrate and, despite isolation treatments, excite motions of their mounting points. Since most structures have inherently low damping characteristics such motions may well be large. This vibrational energy can then flow through the structure and cause significant motions in, say, cabin panels. These in turn radiate noise into the working environments of crew or passenger cabin spaces. Vibrations of this kind may also affect sensitive equipment or lead to fatigue failures in light-weight alloy structures. The most common treatment for such problems is to coat the structural elements with heavy viscoelastic damping materials with consequent weight and cost penalties. Moreover, the effectiveness of such treatments diminishes with the vibration levels which makes continuously improving noise and vibration targets difficult to meet. Clearly, if the vibrational energy could be contained near to the points of excitation there would be a reduced need for damping treatments and, additionally, they could be concentrated in regions where they were most effective. This is precisely the aim of the vibration isolators used between most pieces of equipment and their supporting structure. However, it is difficult and expensive to isolate large pieces of structure in this way, although this is done in modern ship design where sometimes entire deckhouses may be placed on resilient mountings. Of course, some pieces of structure, such as the wings on an aircraft, cannot be mounted resiliently. The upshot of this problem is the desire for some kind of structural filter design capability that could be used to design structures able to carry static loads but that blocked higher frequency motions.

This paper outlines some elements of a programme of work thas has been carried out over the last ten years addressing noise isolation problems. The programme has had two main threads: the first concerns the search for computationally efficient methods for analysing the structures of interest, using the technique known as Statistical Energy Analysis (SEA)[1,2,3,4,5,6,7,8] while the second has examined ways of utilizing such methods during the design process, principally by the use of optimizers to improve designs[9,10,11,12,13,14].

STATISTICAL ENERGY ANALYSIS

As has already been stated, noise standards are becoming increasingly stringent and structural vibration can lead to rapid fatigue failures, particularly in light weight structures. Unfortunately, most of the structures used in modern engineering design are quite complex and, particularly at low to medium frequencies (50-500Hz., say), their acoustic behaviour cannot be divorced from their structural dynamics. These frequency ranges are not amenable to analysis using finite elements (FE) or standard acoustics methods. The FE method, for instance, depends heavily upon numerical procedures which demand large, fast computational facilities in order to deal with mathematical models representing very detailed idealisations of the physical structures. The computational demands increase with structural and material complexity, and with analysis frequency range. Even today, when computational methods are highly developed and optimized, it is not generally practicable to predict the detailed vibrational behaviour of such structures at frequencies beyond the first twenty or so vibrational modes. Conversely, acoustics methods rely heavily on statistical assumptions that may well not be valid until rather high frequencies are encountered where very many modes of the structure are excited. In such circumstances other methods may be more appropriate.

One such is Statistical Energy Analysis (SEA) which dates from the early 1960's[15], when engineers sought new analytical methods for dealing with the problem of predicting the response of launcher and payload structures to rocket noise at launch. Statistical concepts and models of dynamic behaviour, which had been exploited for many years in the analysis of sound fields, were borrowed and adapted to structural systems. System parameters were expressed in probabilistic terms, and the objective of an analysis was seen to be the prediction of the ensemble-average behaviour of sets of grossly similar realisations of an archetypal system (such as the products of an industrial production line). System response to vibrational inputs was characterised by time-average vibrational energy; energy flow between coupled sub-systems was expressed in terms of energy transfer coefficients; and vibration distribution was determined from power balance equations.

Since that time, SEA has undergone relatively little rigorous theoretical development, but the basic results for simple idealised systems have been tested against those from alternative, deterministic approaches which may be implemented because of the simplicity of the models. SEA predictions have also been tested against experimental results, but, because of the large amount of poorly, and even misleadingly, presented information, it is extremely difficult to draw reliable conclusions from some of these. However, it can be confidently stated that when properly applied, SEA concepts have provided a very useful framework on which to design experiments and generate empirical data. In fact, on the basis of such data, both the European Space Agency and NASA routinely use SEA methods to predict the response of spacecraft to inflight vibrational inputs.

The correct application of SEA to any problem requires that the structure of interest be broken down into a number of interacting sub-systems. It then requires an understanding of the couplings between these subsystems and the losses within them. Figure 1 illustrates a set of subsystems that might be used to model a typical offshore rig. This figure also illustrates how energy might flow around this structure for a typical source of vibration. Although this figure is quite simple the underlying problem is extremely complex, requiring details of the behaviour of each of the constituent subsystems together with knowledge of their interactions. The study of the noise performance of each subsystem would represent a considerable feat of analysis; optimization of the whole assembly currently lies beyond the scope of existing computational facilities. No doubt, as computational capabilities increase such things will become feasible, in the meantime attention has been focussed on simpler models.

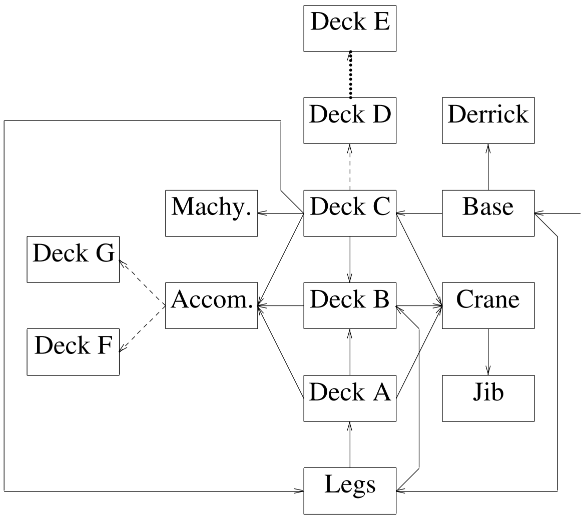

Figure 1 - a subsystem breakdown for a typical offshore rig with energy flows indicated by arrows.

A SIMPLIFIED MODEL

Perhaps the simplest system that can be formulated when considering the flow of vibrational energies between connected structures consists of two, axially vibrating rods which are mutually coupled at a point through a spring. To simulate the noise control design problem these rods may then be assumed to have varying material and/or geometrical properties which must be selected to provide the desired characteristics. Here, the mass properties of the individual rods are modelled as unknown, but bounded variables that vary as functions of position. The response variables of interest are taken to be the steady state energy levels in the two systems, these variables in turn being described in terms of various system transfer functions. Of particular interest is the coupling power receptance function, i.e., it is this quantity that must be minimized if energy is to be prevented from flowing from one structure to the other. The energy flow characteristics of this type of system has been studied by several authors[16, 17, 6] and the expression for the desired receptance function is reproduced here without detailed derivation. Thus, when two rods are coupled at $x_i = a_i$ and are excited by point forces $F_i(t)$ acting at $x_i = b_i$, with frequency spectra $S_{F_i F_i}$, $i = 1,2$, the coupling power receptance for flows from the first to the second is given by

$$H_{12}(\omega) = \frac{\omega^2 k_c^2 c_2}{m_1^2 m_2 |\Delta|^2} \sum_{r=1}^{\infty} (\psi_r^2(a_2) / |\phi_r|^2) | \qquad (1)$$

$$\times \sum_{i=1}^{\infty} (\psi_i(b_1)\psi_i(a_1)/\phi_i) |^2 \qquad .$$

In the above equation the summations over the indices i and r respectively denote summations over the modes of the first and second rods. The quantities ϕ_i and Δ are given by

$$\phi_i = \omega_i^2 - \omega^2 + i c_1 \omega \qquad (2)$$

and

$$\Delta = 1 + (k_c/m_1) \sum_{i=1}^{\infty} (\psi_i^2(a_1)/\phi_i) \qquad (3)$$

$$+ (k_c/m_2) \sum_{r=1}^{\infty} (\psi_r^2(a_2)/\phi_r) \qquad .$$

Here, ω_i and ψ_i are respectively the natural frequencies and mode shapes of the rods. The quantities c_i,m_i and ρ_i respectively denote the coefficient of viscous damping, total mass and mass per unit length of the ith system. k_c is the coupling spring constant. The forces F_1 and F_2 are assumed to be statistically independent.

When the masses of the individual rods are modelled as varying quantities, the natural frequencies and mode shapes vary in a very complex fashion. This results in the receptance function having properties that are extremely difficult to predict. Here, the aim of the optimization process is to obtain a minimum value of $H_{12}(\omega)$ as a function of the mass profiles of the two rods. A general analytical solution to this problem is currently not possible. However, for every realization of the pairs of rods, the natural frequencies and mode shapes can be calculated for the individual rods and this information incorporated into equations (1-3) to generate the receptance function. A crucial step in this calculation involves obtaining the natural frequencies and mode shapes of the rods. When the mass properties vary along the lengths of the rod special means are needed to derive these natural frequencies and mode shapes[6]. In short, even this simple case leads to a rather complex and time consuming optimization problem.

Initial System

All design problems start from some simple initial system which is refined as the design process progresses. Here the initial values of the system properties are taken to be ρ=4.156 kg/m, L=5.182 m, AE=0.1785 MN and c=80 s^{-1}. The coupling spring constant is taken to be k_c=0.5x10^7 N/m and the points of coupling fixed at a_1=2.3m (0.44L) and a_2=3.3 m (0.64L). The forces are assumed to act on only the first subsystem with the second subsystem remaining externally unforced. The point of forcing is taken to be b_1=1.192 m (0.23L) and the driving frequency range of interest is taken to be 5,000 rad/s to 25,000 rad/s. This represents a very simple configuration with two similar systems having uniform properties and arbitrarily chosen drive and coupling points.

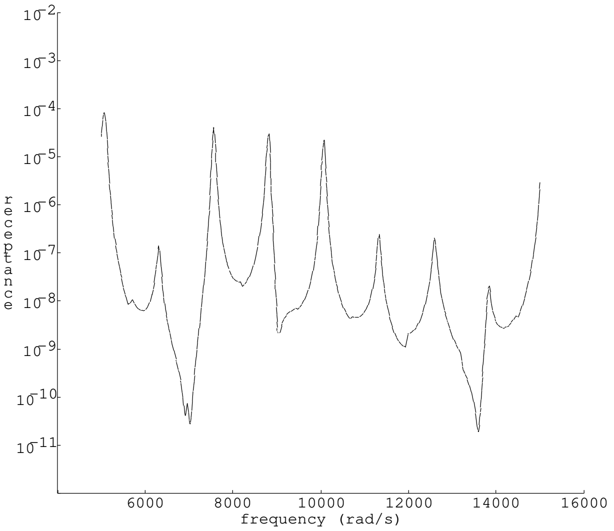

Figure 2 - power cross receptance function, $H_{12}(\omega)$, for the two rod model.

To evaluate this configuration in terms of energy flows, $H_{12}(\omega)$ must be calculated and here the infinite summations appearing in equations (1-3) have been carried out over modes occurring in a frequency band of width 8,796 rad/s centred at the driving frequency ω. Figure 2 shows the resulting power cross receptance function and it can be seen to be dominated by the natural frequencies of the two systems. Notice the logarithmic scales used to plot the function; it shows massive variations in behaviour from frequency to frequency. The behaviour of systems with non-uniform properties is similarly complex and even harder to predict *a priori*. It is the goal of optimization methods to be able to rapidly explore such problems without prior knowledge, finding useful configurations as they go.

Optimization

When formulating this problem as an optimization study, a decision must be made on how many sections to discretize the rods into, i.e., on how many variables will be used to describe the problem. Clearly, if few are used the optimization problem will be simplified but the range of options studied will be restricted and good, novel configurations may not be found. Conversely, too fine a discretization may waste computing time leading to the search being stopped too early and again the possibility of missing good designs. As in all things a compromise must be made and in order to see the effect of this decision two levels of discretization have been studied, with the rods being broken into 2 or 20 piecewise uniform elements of equal length, such that $\rho_{im}, i=1,2, m=1,2(20)$, is the mass per unit length of the ith rod in the mth section. This variation also allows study of the ability of the optimizers to handle different orders of problem, i.e., four or forty variables. The optimizers are then required to determine values of $\rho_{im}, i=1,2, m=1,2(20)$, which lead to the minimum of the function $H_{12}(\omega)$ at ω=10,000 rad/s. The decision to minimize energy flow at a single frequency considerably reduces the length of time taken to study this problem, as in reality the response would be integrated over a frequency range and the resulting integral would be minimized. Such integration requires many more function evaluations but it does smooth the function being examined and this somewhat simplifies the problem, i.e., here a quicker but harder problem has been chosen in the interests of research into the applicability of optimization methods in this area. The results of applying the best method found here to the integral of $H_{12}(\omega)$ are given later to demonstrate the great potential benefits optimization may bring in producing enhanced noise performance.

All real design problems are bounded in some sense; in structural dynamics such bounds often reflect the need for structures to carry static or low frequency loads, also real designs may need to use stock material sizes and so on. It is quite common for optima to be created by the intersections of these constraints in some way or other. Such constraints are often difficult for optimizers to handle since they represent massive local distortions in the optimization surface. Here the design is arbitrarily constrained by ensuring that the total mass of the individual rods given by

$$m_i = \sum_m l_m \rho_{im} \; ; \; i=1,2 \tag{4}$$

take values in the range $3 < m_i < 50$ kg ; $i=1,2$ along with the restriction on each individual subsection that $0.5 < \rho_{im} < 10.0$ kg/m.

Optimization Methods

Having decided on the function to be optimized, the number of variables involved and the constraints to be applied, attention must then be focussed on the method to be used. Optimization has been studied for a great many years, both as an abstract mathematical problem and also from the viewpoint of engineering design. A great many methods have evolved which are detailed in a sizable literature with each method having its champions and detractors. A good survey of what may be called ‘classical’ and ‘heuristic’ methods, presented in an engineering context, is given by Siddall[18]. A common feature of these methods is that they date from before the advent of really cheap powerful computers. They have thus tended to be developed with low order problems in mind, since the computational power to test them on high dimensional problems was not available at their inception. They commonly make use of derivatives and hill climbing methods. Optimization methods can also be found in most libraries of commercially produced numerical software and that produced by NAg Ltd. is typical[19]. Finally, in recent years two competing, modern methods have received much publicity, i.e., the Genetic Algorithm (GA) and Simulated Annealing (SA, sometimes known as stochastic relaxation). Both these methods are acquiring a significant literature detailing their performance and variants on the basic themes[20, 21].

method	no. vars.	range consts.	minimum	feasible	no. trials
GA	4	yes	0.6021E-8	yes	500
SA	4	yes	0.9805E-8	yes	500
Linear	4	yes	0.7787E-8	yes	54
H&J	4	yes	0.2619E-5	yes	132
E04UCF	4	yes	0.6538E-6	yes	65
Linear	4	no	0.1682E-10	ranges	259
H&J	4	no	0.1613E-10	ranges	988
GA	40	yes	0.6887E-8	yes	500
SA	40	yes	0.4119E-7	yes	500
Linear	40	yes	0.7834E-6	yes	336
H&J	40	yes	0.8639E-8	yes	3,895
E04UCF	40	yes	0.2619E-5	yes	228
Linear	40	no	0.1725E-10	ranges	3,403
H&J	40	no	0.2185E-8	yes	7,426

Table 1 - comparison of results for the two rod model.

Since the problem being studied here is known to be complex and in one form, of high order, a representative selection has been taken across these methods and the results compared. The methods tried were

(1) Genetic Algorithm (GA)$,

(2) Simulated Annealing (SA)*,

(3) Linear Approximation with Linear Programming†,

(4) Hooke and Jeeves Direct Search‡ and

(5) NAg Routine E04UCF+.

In all these cases the underlying methods are detailed in the references with the following variations/options being adopted here:

(1) the GA uses binary encoding, a sixteen bit word length, survival proportional to fitness, elitism, single point crossover, inversion, low level mutation and optional niche forming;

(2) the SA uses an annealing schedule controlled by three parameters with one governing the number of temperatures in the schedule, one their spread and the third their absolute values with the temperatures being divided logarithmically within the desired range;

(3) the linear method is used as given except that the variable bounds are strictly enforced by using them as constraints;

$ The GA works by maintaining a pool or population of competing designs which are combined to find improved solutions. In their basic form, each member of the population is represented by a binary string that encodes the variables characterizing the design. The search progresses by manipulating the strings in the pool to provide new generations of designs, hopefully with better properties on average than their predecessors. The processes that are used to seek these improved designs are set-up to mimic those of natural selection: hence the method's name.

* SA is based on the kinetics of freezing crystals, where it is observed that minimum energy states are reached for sufficiently slow cooling. Essentially, the method makes random small changes to the design and these are accepted if they improve it and, occasionally, even if they worsen it. The likelyhood of changes that worsen the solution being accepted is controlled by a Boltzman probability function which is dependent on the so called annealing temperature.

† This approach uses small explorations in the vicinity of the current design to establish the local slope. The problem is then assumed to be linear near this point and classical linear programming methods used to find the best point in this sub-space. The process is then repeated around the new point and so on, with the size of the linear region being reduced at each new base point. This method is very rapid for smoothly varying problems with few variables but tends to find local rather than global solutions.

‡ This approach is a classical heuristic search that wanders around the design space following a pre-programmed pattern of moves, seeking the steepest ascent to the maximum. When used with a one pass penalty function it also uses random point generation to test its final solution for better points in the vicinity of the current design. When such an improvement is found this is used to restart the search.

+ E04UCF is a classical gradient based method that uses small steps in each direction to evaluate the gradient if such information cannot be supplied directly.

(4) the Hooke and Jeeves method is used as given except that the variable bounds are strictly enforced by using them as constraints and

(5) the NAg routine is used as given except that no derivative information is supplied to the routine.

The first four of these are unconstrained methods and when using these optimizers the constraints have been enforced by applying a simple one pass external penalty function of the type described by Siddall[18].

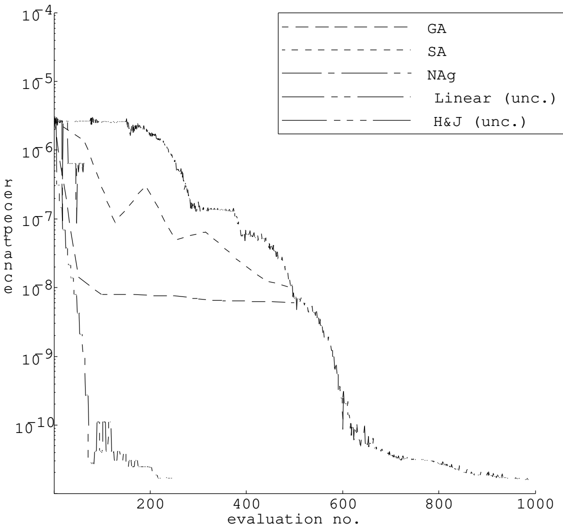

Figure 3 - objective function $H_{12}(\omega)$ versus evaluation number for four variables; all methods, linear approximation and Hooke and Jeeves search without variable bounds constraints.

Four Variable Problem

To begin with, each method was applied to the problem with an allocation of up to 500 evaluations of the objective function, using four variables to describe the rods. The first five lines of Table 1 summarize the results obtained and it may be seen that the best result is gained with the GA although the linear approximation reaches nearly the same result in around 12% of the time. The Hooke and Jeeves search did not make significant changes to the start point while the NAg routine made only modest improvements. The SA method performs nearly as well as the GA although showing more violent fluctuations around its steady downward progress.

Two features common to both the linear and Hooke and Jeeves methods, as implemented by Siddall[18], are that, by default, they both take bounds on the design variables as being advisory and that they often use more function evaluations than desired if significant improvements are being achieved. Lines six and seven of Table 1 summarize the effects of running these two methods with the explicit variable bounds constraints removed. In both cases significant improvements in the objective function are obtained, with the Hooke and Jeeves search using nearly 1,000 steps. This indicates that the additional freedom made available to the optimizers allows them to drive a notch into the curve at the frequency chosen for optimization, although the rates of convergence are no faster. Note also, that the final designs violate the restrictions on the design variable ranges. The progress of the various optimizers is illustrated in figure 3 where the curves of $H_{12}(\omega)$ versus evaluation number for all the methods used are plotted on a common base.

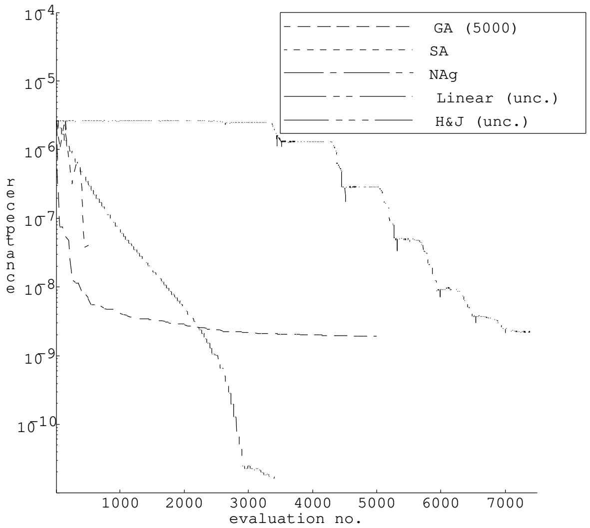

Figure 4 - objective function $H_{12}(\omega)$ versus evaluation number for 40 variables; all methods, linear approximation and Hooke and Jeeves search without variable bounds constraints.

Forty Variable Problem

Turning to the case of forty variables, the first five lines of the second half of table 1 shows the results of these calculations. Again some of the methods aborted the search before using the entire allocation of 500 trials, while the Hooke and Jeeves search continues until nearly 4,000 evaluations have been used. It may be seen that the best result is once more obtained with the GA, although it is slightly worse than for the four variable case using the same number of function evaluations. Here, the linear approximation does significantly worse than before, with the Hooke and Jeeves search coming second, albeit at the cost of around eight times as many evaluations as the GA. The NAg routine this time failed to achieved significant improvements on the start point. The SA method performed quite well but with less rapid convergence than the GA.

Relaxing the variable bound constraints once more gives the linear and Hooke and Jeeves methods more room for manoeuvre. The final two lines of Table 1 give details of these last two calculations and in this case the Hooke and Jeeves search produces a feasible result despite the relaxed constraints. Moreover, the final result represents the best achieved throughout this initial survey, being some 64% less than that achieved by the GA with four variables. It does however require nearly 7,500 evaluations to achieve this. These features may be seen in figure 4 where the curves of $H_{12}(\omega)$ versus evaluation number for all the methods used are again plotted on a common base, c.f., figure 3. Figures 3 and 4 allow an important conclusion to be drawn; viz., the rates of convergence of the classical methods are all affected, to a greater or lesser extent, by the number of variables used in the problem. In particular, the rapid convergence of the linear method when dealing with low order problems is lost when tackling high order cases. Figure 4 also shows the effect of allowing the GA to proceed for 5,000 evaluations to see if it can better the Hooke and Jeeves method and this is seen to be the case. Its rate of convergence is also vastly better.

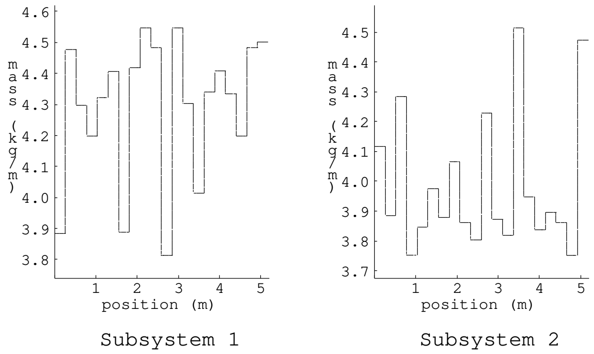

Figure 5 - mass profiles ρ_{im} for the system range optimized using the GA with 40 variables.

Minimization of Energy Flow over a Range of Frequencies

As has previously been mentioned, to be useful in producing enhanced noise performance the integral of the energy flow with respect to drive frequency must be minimized. This is a very time consuming calculation and so has not been used as the basis of the previous rather wide ranging survey. Nonetheless, to illustrate the power of the GA with default parameters it has been applied to this problem with an integration width equal to 8,796 rads/s centred at the previous frequency of interest of 10,000 rad/s. The resulting mass profiles and curve of $H_{12}(\omega)$ are shown as figures 5 and 6. It is clear from these figures that the optimizer has been able to identify a configuration that reduces the energy flows over a range of frequencies, reducing the three main response peaks in the range considered by more than 50dB. Moreover, although the resulting mass profiles shows some tendencies towards geometric repetitions they show no precise patterns. As such, they can confidently be expected to give similar reductions in noise transmission when slightly modified, either deliberately or as a result of manufacturing inaccuracies. This rather random nature of the mass profiles is, of course, inherent in the GA approach to design since each new variant is the result of essentially random changes from a previous design as it is the selection process that the GA governs rather than the synthesis of new designs.

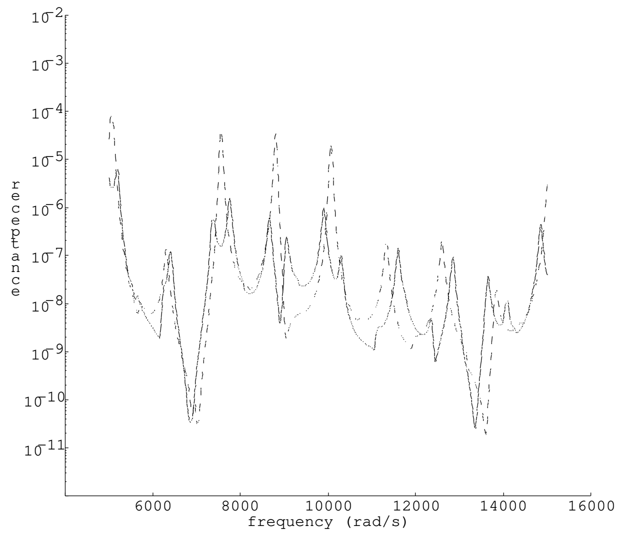

Figure 6 - $H_{12}(\omega)$ versus ω for the initial system and that range optimized using the GA with 40 variables: range optimized system - solid line; original system - dotted line.

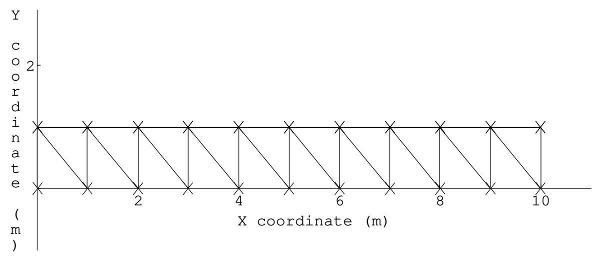

Figure 7 - a simplified, two-dimensional satellite boom.

A SATELLITE BOOM

Figure 7 indicates the second structure to be considered here. This consists of a simplified, two-dimensional satellite boom made up of forty Euler-Bernoulli beams all having the same properties per unit length. Here, *EA* is taken as 12.86 kN, *EI* as 69.8MNm^2 and mass per unit length as 2.74 kg/m. The beams are all either 1 m or 1.414 m long and the joints at (0,0) and (0,1) are taken to be pinned to ground; all other joints are free to move. The structure has been chosen to be two-dimensional for simplicity. It is excited by a point transverse force half way between (0,0) and (1,0) and, for this study, the aim is set as the minimization of the vibrational energy level in the right-hand end vertical beam (which in practice might carry an instrumentation package or similar). The damping of the structure is fixed so that the normal modes of the uncoupled beam elements all have a constant bandwidth of 20 s^{-1}.

Figure 8 shows how the vibrational energy level of the end beam varies with frequency for this base design. As can be seen from the figure, the energy level exhibits the many peaks and troughs characteristic of a complex, lightly damped structure. Moreover, there is a large response in the region 100-200 Hz. The calculation of a single frequency point in this plot requires the solution of some 260 complex simultaneous equations and this takes around 20 s on a Sun Sparc 10 workstation; i.e., the calculation of a frequency band average takes around seven minutes using the 21 point rule to be adopted here. Even so, this is considerably faster than can be achieved using a finite element approach with the same level of accuracy. Nonetheless, it remains a very hard problem for optimizers to deal with in a realistic time span.

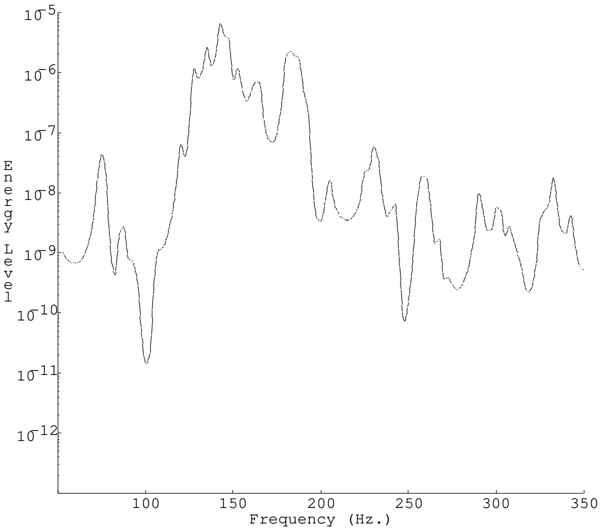

Figure 8 - initial frequency response of the end beam of the boom.

Optimization Problem

The aim of this second study is set as the reduction of the frequency averaged response of the end beam in the range 150-250 Hz by the use of optimization, i.e., to trim down the upper half of the major response peak. In practice, such a requirement would reflect known sensitivities of the payload or known excitation frequencies. To meet other design requirements the optimization is constrained to keep the end beam unchanged in length and position with respect to the fixed points. Further, to meet structural requirements, all of the joints within the structure must be kept within fixed distances from their original positions. This ensures that no beam is too long or too short and also restricts the overall envelope of the structure. The free variables in the problem are thus set as the x and y coordinates of the 18 mid-span joints, i.e., 36 variables in all. In this case only the GA is used to tackle the optimization problem.

Results

To begin the optimization process relatively tight limits of ±5% are placed on the joint positions and 1,000 trials used, spread over 5 generations (giving 200 members per generation so as to ensure reasonable coverage over the rather large domain being investigated here, at the expense of final convergence). This run gives rise to the configuration shown in figure 9. With this configuration, the total energy flow in the 100Hz band investigated has been reduced from 0.33×10^{-6} to 0.21×10^{-7} (for unit forcing), i.e., a reduction of 23dB, indicating how purely periodic structures can give rise to very significant noise transmission problems. Next, the limits on the joint positions are relaxed to ±25% but only six of the midspan joints are allowed to move, again using 1,000 trials and 5 generations. This leads to the configuration of figure 10. It gives a slightly lesser level of isolation with the total energy flow now being 0.31×10^{-7}, i.e., a reduction of 21dB over the base design. As can be seen from figure 10, a kind of notch has been driven into the structure and this clearly causes reflections of energy carrying waves in the desired frequency ranges.

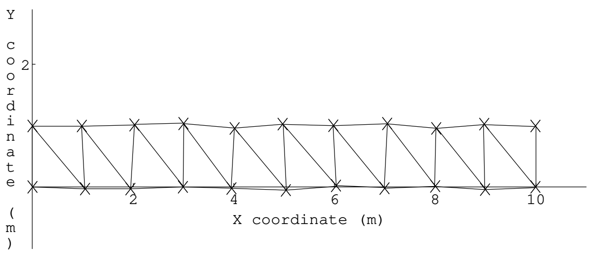

Figure 9 - optimized design with limits of ±5% on all 18 joints, 1,000 evaluations over five generations.

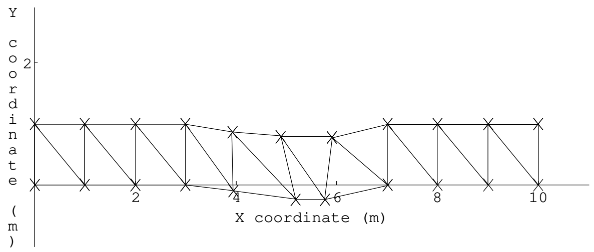

Figure 10 - optimized design with limits of ±25% on six mid-span joints, 1,000 evaluations over five generations.

Having gained some initial impressions, the GA was next applied with a ±25% constraint on the joint positions applied to the whole structure, but now using around 4,500 evaluations with 15 generations of 300 members resulting in the structure of figure 11. Here, the energy flow has been reduced to 0.13×10^{-9}, i.e., a reduction of 68dB. The final structure is particularly interesting, showing a strong overall pattern where, the depth of the composite cantilever is increased at both ends and reduced mid-span. Such a design is clearly workable and demonstrates the power of the technique being used. Figure 12 further illustrates this point, showing the full frequency response of this design compared to the base structure, and illustrating that gains over the 150-250 Hz band have not been achieved at the expense of significantly worsened behaviour at other frequencies.

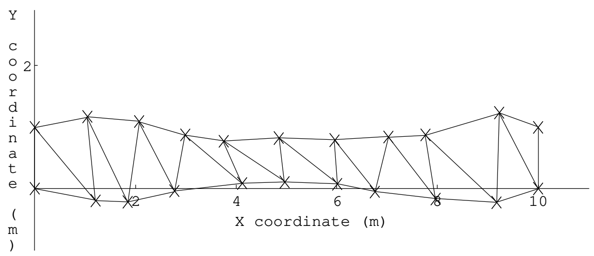

Figure 11 - optimized design with limits of ±25% on all joints, 4,500 evaluations over 15 generations.

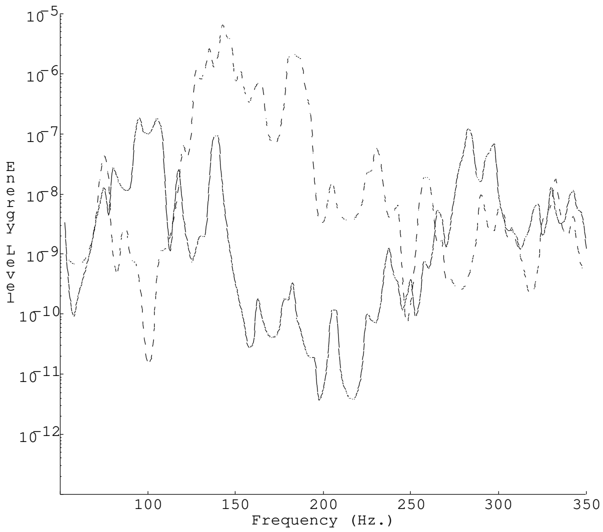

Figure 12 - frequency response of optimized design with limits of ±25% on all joints, 4,500 evaluations over 15 generations with response for initial design shown dotted.

A TEST PROBLEM

It is well known, and the previous examples have shown, that Genetic Algorithms are good at handling cases where the objective function of the problem is characterized by a number of sharp peaks, which lead classical, slope driven, methods to terminate at false optima. A comprehensive survey of this work is provided by Goldberg[20]. However, there is relatively less work in the literature dealing with the constraint surfaces which are commonly found in engineering design problems. Moreover, it is often the case that such constraints define the true global optimum solution to a design problem. They typically arise from material limitations, such as maximum stress or temperature levels. The calculation of these quantities is often complex and the designer has little *a priori* knowledge of how they will limit the choice of the free variables in the problem under consideration. The third example considered here is concerned with the application of optimizers, and the GA in particular, to such an extremely hard, constrained, multi-peaked function optimization.

GA's, along with many other techniques, have a number of difficulties in dealing with this kind of task; the two most severe appear to be:

(1) convergence to sub-optimal solutions and

(2) robustness of the convergence for repeated attempts using different random number sequences within the routines.

In an attempt to overcome these problems, a program of work has been carried out looking at chosing the best set of optimizer control parameters for a test problem. The choice of control parameters forms one of the major difficulties in using optimization software: ideally an optimizer should require few such controls and more importantly, the results in terms of accuracy and speed, should be relatively insensitive to this choice. It is usually the case that, on any given problem, one particular optimizer can be tuned to give startlingly good performance; it is much more difficult to choose a method and select its control parameters so as to get good results at the first attempt. Since optimization is only really difficult where computation times for achieving an optimum are long, no *real* problem allows the luxury of such fine tuning. Nonetheless, researchers working on optimizers must consider tuning aspects so as to be able to assess competing methods. When an optimizer has many, inter-related parameters, this forms an optimization task of its own, i.e., to chose a set of control parameters that will, on average, give good results in the minimum time when applied to a particular problem.

Details of the Genetic Algorithm

The GA used throughout this paper is fairly typical of those discussed by Goldberg[20]. As has already been noted, such methods work by maintaining a pool or population of competing designs which are combined to find improved solutions. The most common genetic operations cited in the literature as being used in this process are currently:

(1) selection according to fitness, i.e., the most promising designs are given a bigger share of the next generation;

(2) crossover, where portions of two good designs, chosen at random, are used to form a new design, i.e., two parents 'breed' an 'offspring';

(3) inversion, whereby the genetic encoding of a design is modified so that subsequent crossover operations affect different aspects of the design and

(4) mutation, where small but random changes are arbitrarily introduced into a design.

In addition, the number of generations and their sizes must be chosen, as must a method for dealing with constraints (usually by application of a penalty function).

The algorithm used here works with up to sixteen bit binary encoding (although parameters that are selected from a number of fixed possibilities use only the minimum necessary number of bits and the number of bits can be reduced by the user for continuous variables). It uses an elitist survival strategy which ensures that the best of each generation always enters the next generation and has optional niche forming to prevent dominance by a few moderately successful designs preventing wide ranging searches. Two penalty functions are available. The main parameters used to control the method may be summarized as:

N_{gen} the number of generations allowed (default 10);

N_{pop} the population size or number of trials used per generation which is therefore inversely related to the number of generations given a fixed number of trials in total (default 50);

P[best] the proportion of the population that survive to the next generation (default 0.8);

P[cross] the proportion of the surviving population that are allowed to breed (default 0.8);

P[invert] the proportion of the surviving population that have their genetic material re-ordered (default 0.5);

P[mutation] the proportion of the new generation's genetic material that is randomly changed (default 0.005);

a Proportionality Flag
which selects whether the new generation is biased in favour of the most successful members of the previous generation or alternatively if all P[best] survivors are propagated equally (default TRUE) and

the Penalty Function choice.

When using the GA to explore large design spaces with many variables, it has also been found that the method must be prevented from being dominated by a few moderately good designs which prevent further innovation. A number of methods have been proposed to deal with this problem; that used here is based on MacQueen's Adaptive KMEAN algorithm[22] which has recently been applied with some success to multi-peak problems[23]. This algorithm subdivides the population into clusters that have similar properties. The members of each cluster are then penalized according to how many members the cluster has and how far it lies from the cluster centre. It also, optionally, restricts the crossover process that forms the heart of the GA, so that large successful clusters mix solely with themselves. This aids convergence of the method, since radical new ideas are prevented from contaminating such sub-pools. The version of the algorithm used here is controlled by:

D_{min} the minimum non-dimensional Euclidean distance between cluster centres, with clusters closer than this being collapsed (default 0.05);

D_{max} the maximum non-dimensional Euclidean radius of a cluster, beyond which clusters sub-divide (default 0.2);

N_{clust} the initial number of clusters into which a generation is divided (default $0.2 \times N_{pop}$);

N_{breed} the minimum number of members in a cluster before exclusive inbreeding within the cluster takes place (default $0.1 \times N_{pop}$) and

α the penalising index for cluster members, which determines how severely members sharing an over-crowded niche will suffer, with small numbers giving less penalty (default 0.2), i.e., the objective functions of members of a cluster of m solutions are scaled by $m^{\min(\alpha,1)}[1-(E/D_{max})^{\alpha}] + (E/D_{max})^{\alpha}$, where E is the Euclidean distance of the member from its cluster centre (which is always less than D_{max}).

In this form the GA has 13 control parameters that the user may alter before applying the method.

Penalty Functions

As has already been mentioned, in most work on GA's, constraints are dealt with by the use of penalty functions: such functions are used to distort the objective function in order to force the search towards feasibility, when a constraint is, or is about to be, violated. Typically, a one pass external function is used so that, whenever a design violates a constraint, a large penalty is applied, effectively ruling it out of further consideration. This is very severe and sometimes the search stalls, particularly if it must follow along a constraint line to reach the optimum. Despite this disadvantage, the strategy is often quite successful, and requires a minimum of computer time when it works.

Other penalty functions soften this severe distortion and to try to achieve less risk of premature stalling of the search. One such is the Fiacco and McCormick function which was originally conceived for use with multiple passes of classical, slope driven optimizers. In this case the function is set up to fit naturally within the GA by being applied with increasing severity to each new generation. It modifies the objective function to be maximized, f, as follows:

$$f' = f - \sum_i \frac{(\phi_i^v)^2}{p^{\max(1, n/N_{pop}-2)}} \qquad (5)$$

$$- \sum_i \frac{p^{2.\max(1, n/N_{pop}-2)}}{\phi_i^s} \quad .$$

Here ϕ_i^v are the violated constraint conditions and ϕ_i^s those that are satisfied (in both cases being normalised so that violated constraints are negative and satisfied ones positive). p is a penalty quantity, less than one (default 0.5), chosen to suit the problem in hand, n the current function evaluation number and N_{pop} the number of members in each generation. Thus, to begin with, when n/N_{pop} is small, violated constraints are only slightly penalized along with satisfied ones that are near the constraint boundaries, effectively warning the optimizer of the presence of boundaries while allowing their exploration. As the optimization proceeds and the number of evaluations increases, n/N_{pop} grows causing $p^{\max(1, n/N_{pop}-2)}$ to become exponentially small, severely penalizing the violated constraints and removing the effect of the term involving the satisfied ones, causing the optimizer to chose feasible solutions where f' reverts to f. Notice that here the division n/N_{pop} is carried out in integer arithmetic causing this quantity to be constant over members of the same generation. By also subtracting two, but preventing the result from being less than unity, the penalty is deliberately held low during the first three generations of the method, giving the GA chance to get started before the penalizing function begins to take effect (see Siddall[18] for further details of the original usage of this function). The use of an evolving constraint penalty fits particularly naturally within the general scheme of the GA yet does not seem to have been much used before, perhaps because most applications of the GA to date have been to unconstrained problems. Certainly, the Fiacco and McCormick function has been widely discussed in the literature of 'classical' optimization methods.

A Bumpy Equation

To simulate a multi-peak optimization problem the following objective function can be defined :

$$f(x_i) = \left\{ \mathrm{abs}\left(\sum_{i=1}^{m} \cos^4(x_i) - 2\prod_{i=1}^{m} \cos^2(x_i)\right) \right\} \qquad (6)$$

$$\Big/ \left\{ \sqrt{\sum_{i=1}^{m} i x_i^2} \right\}$$

where the $x_i, i = 1, \cdots, m$ are the variables (expressed in radians) and m is the number of dimensions. This function produces a series of peaks that get smaller with distance from the origin and which are nearly symmetrical about $x_i = x_j, i, j = 1, \cdots, m$. The optimization problem is then defined as find x_i for $0 < x_i < 10, i = 1, \cdots, m$ starting from the point $x_i = 5, i = 1, \cdots, m$ to maximize the function subject to $\prod_{i=1}^{m} x_i > 0.75$ and $\sum_{i=1}^{m} x_i < 300$. The case for $m = 2$, which may be readily plotted, is illustrated in figure 13 along with the traces of the various optimizers discussed later. This two-dimensional problem has a number of features that make optimization difficult:

(1) the surface is nearly, but not quite symmetrical in $x_1 = x_2$, so that peaks always occur in pairs but with one always bigger than its sibling;

(2) the true maximum is 0.365 at (1.593, 0.471) which is defined by a constraint boundary;

(3) there is another similar peak of height 0.274 at (0.475, 1.578);

(4) the major maximum within the boundaries has a height of 0.263 at (3.087, 1.517) which is thus quite competitive with solutions on the slopes of the true, constrained maximal peak (where the slope is much steeper and the base area much smaller), causing many solutions to be trapped on this sub-optimal peak and

(5) the starting point lies on the line $x_1 = x_2$ where the function is always zero so that explorations carried out by making small equal changes in x_1 and x_2 show the function as invariant (affecting many hueristic methods).

When the problem is generalised for m greater than two it becomes even more demanding with families of similar peaks occurring with a highly complex constraint surface. The two-dimensional problem is studied initially since this aids display of the function and speed of calculation.

Initial Optimization

To gain some initial insights into the two-dimensional problem, a number of different optimization methods were again deployed using the appropriate default control parameters, these were:

(1) the GA;

(2) simulated annealing (SA), using 10 temperatures;

(3) linear approximation with linear programming and

(4) the Hooke and Jeeves direct search followed by a local random search.

In all cases 1,000 function evaluations was set as the suggested maximum and a simple one pass external penalty function used to deal with constraint violations. As has already been noted, this penalty function causes all solutions that do not meet the constraint requirements to be considered as totally unworkable and consequently assigns them a very low, -10^{20}, objective function value. The resulting optimization traces are shown in figure 13. Here, only the best of generation/annealing temperature of the GA and SA are shown, together with all points from their final generation/annealing temperature. The linear approximation and Hooke and Jeeves methods used many fewer than 1,000 points since they rapidly converged to the wrong peak.

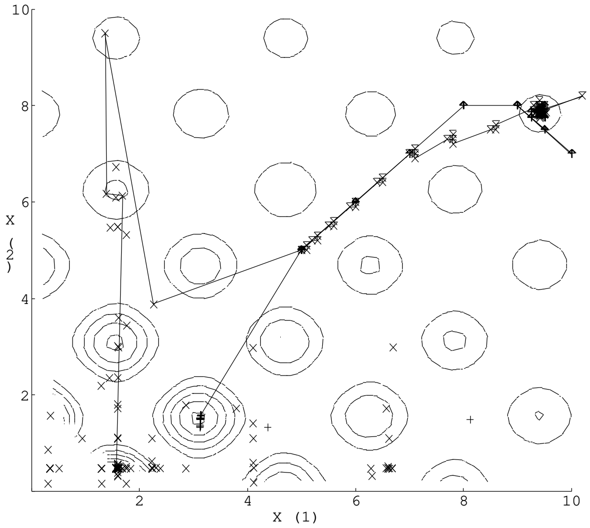

Figure 13 - contour map of two-dimensional test function with optimization traces of various methods, showing the best of generation/annealing temperature of the GA(+) and SA(×) runs plus the final generation/annealing temperature points. The coverage of the contours indicates the extent of the feasible region.

method	maximum (average)	feasible	no. trials
GA	0.7905	yes	140,000
SA	0.7422	yes	140,000
Linear	0.2951	yes	1,529
H&J	0.2850	yes	1,967

Table 2 - comparison of results for the fifty-dimensional problem, averages taken over five independent runs.

Figure 14 shows the results of fifty different runs of both GA and the SA, indicating their final results in each case. These runs differ only in the random number sequences used during the optimization process and indicate the variations in results that can be seen when using stochastic optimization methods: clearly, when dealing with hard problems with many peaks it is more useful to use multiple runs rather than a single longer run. Experience shows that five such parallel runs are usually sufficient to give confidence in the final result.

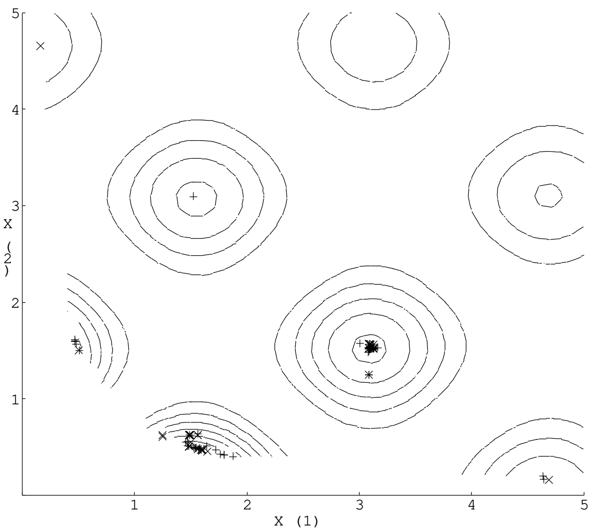

Figure 14 - detailed contour map of two-dimensional test function with final solutions for 50 independent runs of the GA(+) and SA(×).

Fifty-Dimensional Problem

To illustrate some of the potential pitfalls that can be experienced when using modern optimization methods the final sections of this paper examine the test function for $m=50$. In this cases it is obviously not possible to plot out the objective function. Moreover, the optimization runs need to be considerably longer to give good results on this extremely hard problem. Here 140,000 function evaluations has been chosen as the upper limit and in each case five different runs have been used to calculate the average performance of the methods. For the stochastic methods these involve using different random number sequences while for the other two different intial points have been used. Additionally, for this test the GA uses the modified Fiacco and McCormick penalty function mentioned earlier with $p=0.98$. Table 2 compares the results obtained using the four methods. As can be seen the GA and SA dramatically outperform the non-stochastic methods. However, in all cases the non-stochastic methods terminated after less than 2,600 evaluations, which makes their performance seem quite efficient until it is realised that all four methods (and a pure random search) get to an objective function value of around 0.28 in this many steps. What is hard with this problem is finding one of the few high peaks on a function which has tens of thousands of mediocre ones.

The principal difference between the two stochastic methods on this problem was found to be that the GA performed fairly well irrespective of the control parameters used, always turning in results over 0.7 . Conversely, the SA needed much more care in its use and, in particular, the degree of random mutation allowed between successive trials. Here quite low values proved most effective, given the very many function evaluations available. To further indicate the subtlety of this problem figure 15 shows a cross section of the fifty-dimensional surface taken at the final point reached by the best of the five GA runs. Here, variables one and 20 have been chosen arbitrarily as those to be varied. The solution is seen to lie on the constraint boundary and the complex nature of the function is apparent.

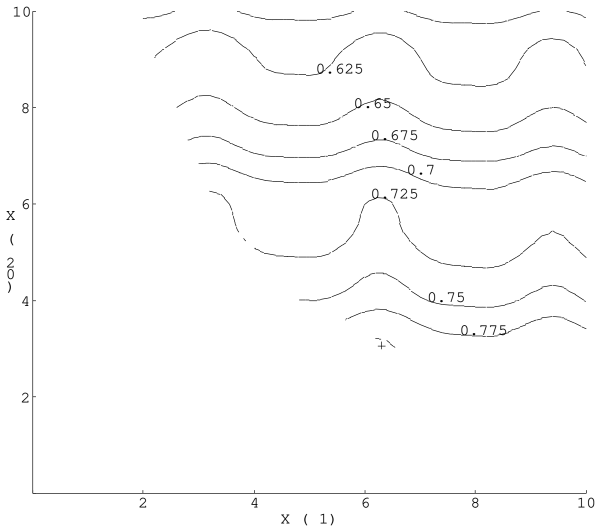

Figure 15 - contour map of fifty-dimensional test function plotted for $x(1)$ and $x(20)$ with the other variables fixed at the best values obtained from the GA searches - peak of 0.8031 marked (+).

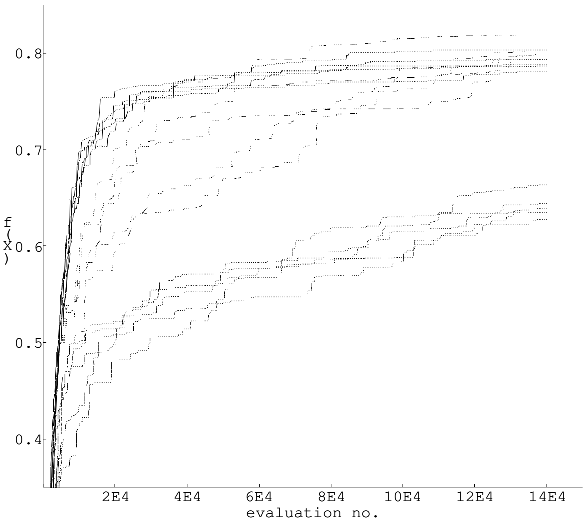

Figure 16 - objective function versus generation number for five independent runs of each of the three GA schemes mentioned in the text with; GA developed by the author - solid lines, GENESIS V5.0 - dashed lines, CHC - dotted lines.

It should not be concluded from this work that all GA's are the same, however. Figure 16 shows the results obtained on this problem when using the GA developed by the author, Grefenstette's GENESIS V5.0[24], which is a basic public domain code developed some years ago and Eshelman's CHC using a population of 30 and many restarts[25, 26]. Clearly, not all GA's are the same. Moreover, even though CHC gets to the best result of all it does so at the expense of initial speed of convergence. The results for GENESIS V5.0 demonstrate how research in this area has progressed over the years since it was last updated (1990). One pays one's money and takes one's pick !

CONCLUSIONS

This overview of some previous work by the author in the field of structural optimization has shown that significant noise isolation characteristics can be introduced into structures by modifying them in a controlled way. Moreover, the structures produced are not unworkable and do not suffer from degraded performance at frequencies away from those that have been the subject of optimization. It should be noted that the improved vibration isolation characteristics gained arise from the constructive reflections of travelling waves caused by the discontinuities introduced. Part of the improvement is seen to be the result of moving away from uniform designs. However, even more significant gains arise from the unusual geometries adopted when significant changes are made. Therefore, a simple small-scale randomization, although often beneficial when dealing with a structure that originally exhibits precise geometric periodicity, does not realize all the gains that can be made. It is, however, much harder to select a particular set of large-scale modifications that control the noise performance in a particular way, while still maintaining a reasonable overall shape. Moreover, it should be noted that, even using a four processor work-station, some of the runs discused here take more than five days of computer time to carry out. Clearly, larger structures would be even more expensive to deal with. However, when considering the cost of satellite or aircraft design and the penalties associated with redundant weight, such calculations should be well worth while.

These studies have also shown that the Genetic Algorithm (GA) performs very well on such tasks when compared to a range of other methods, especially if large numbers of variables are used to describe the designs. This arises because the flow of vibrational energy around a complex structure is dominated by the many resonances exhibited by such structures and also the large number of physical parameters needed to specify typical designs (in the cases studied here up to forty, many hundreds in a full ship or aircraft structural model); both features which the GA is seen to deal with well. Nonetheless, caution must still be exercised when using methods like the GA. Although beguilingly simple to describe they are, in fact, quite difficult to operate to their best effect. They are not unlike many other new and advanced techniques in this respect.

These points may be summarized as follows:

(1) the highly complex nature of the eigenproblems found in structural dynamics leads to difficult optimization domains;

(2) the more complex the design being considered the greater the scope for optimization;

(3) when few variables are used to describe a problem classical, slope driven or linearized methods are very rapid and powerful despite the complex relationships;

(4) when many variables are used to describe the problem the classical techniques may still work but they become rather expensive to use and tend to stall;

(5) Genetic Algorithms and Simulated Annealing methods are then more useful, particularly if false optima are to be avoided;

(6) these methods must be tried over several different sets of random numbers to gain confidence in their results;

(7) GA's seem to be less prone to the exact details of their control parameters but;

(8) they are not all the same and

(9) it is still relatively easy to construct problems of modest size that even the best available optimizers require hundreds of thousands of trials to deal with and even then cannot guarantee to solve.

In the context of this last point it is worth noting that speed of convergence to a 'good' sub-optimal peak may be more important than finding the best peak of all when dealing with engineering problems. However, the designer must be aware of the quality and capabilities of the tools he is using if he is to be confident that the optimized design really is a good one.

REFERENCES

1. A. J. Keane and W. G. Price 1986 *Proceedings of the International Conference on Vibration Problems in Engineering*, *Xi'an*, 881-889. An appraisal of the fundamentals of Statistical Energy Analysis.

2. A. J. Keane and W. G. Price 1987 *J. Sound Vib.* **117**(2), 363-386. Statistical Energy Analysis of Strongly Coupled Systems.

3. A. J. Keane and W. G. Price 1989 *J. Sound Vib.* **128**(3), 423-450. On the Vibrations of Mono-Coupled Periodic and Near Periodic Structures.

4. A. J. Keane and W. G. Price 1989 *Proc. R. Soc. Lond.* A**423**, 331-360. Statistical Energy Analysis of Periodic Structures.

5. A. J. Keane 1992 *Proc. R. Soc. Lond.* A**436**, 537-568. Energy Flows Between Arbitrary Configurations of Conservatively Coupled Multi-Modal Elastic Subsystems.

6. A. J. Keane and C. S. Manohar 1993 *J. Sound Vib.* **168**(2), 253-284. Energy Flow Variability in a Pair of Coupled Stochastic Rods.

7. A. J. Keane 1994 *Engineering Structures* **16**(2), 145-157. Statistical Energy Analysis of Offshore Structures.

8. C. S. Manohar and A. J. Keane 1994 *Phil. Trans. R. Soc. Lond.* **346**(A), 525-542. Statistics of energy flows in spring-coupled one-dimensional subsystems.

9. A. J. Keane 1981 Ship Concept Design System. M.Sc. Thesis, Department of Mechanical Engineering, University College London.

10. A. J. Keane, W. G. Price, and R. D. Schachter 1990 *Trans. R.I.N.A.* **132**(A), . Optimization Techniques in Ship Concept Design.

11. A. J. Keane 1993 *Proceedings of the International Conference on Artificial Neural Nets and Genetic Algorithms*, *ed. R. F. Albrecht, C. R. Reeves and N. C. Steele*, *Innsbruck*, *Springer-Verlag*, 536-543. Structural design for enhanced noise performance using Genetic Algorithm and other optimization techniques.

12. A. J. Keane to appear *Artificial Intelligence in Engineering*, , . Genetic algorithm optimization of multi-peak problems: studies in convergence and robustness.

13. A. J. Keane 1994 *Proceedings of the International Congress on Noise Control Engineering - Inter-noise 94*, *Yokohama*, The design of vibration isolating structures using Genetic Algorithm methods.

14. A. J. Keane to appear *J. Sound Vib.*, , . Passive vibration control via unusual geometries: the application of genetic algorithm optimization to structural design.

15. R. H. Lyon 1975 *Statistical Energy Analysis of Dynamical Systems : Theory and Applications.* MIT Press.

16. H. G. Davies 1973 *J. Acoust. Soc. Am.* **54**(2), 507-515. Random vibration of distributed systems strongly coupled at discrete points.

17. P. J. Remington and J. E. Manning 1975 *J. Acoust. Soc. Am.* **57**(2), 374-379. Comparison of Statistical Energy Analysis power flow predictions with an 'exact' calculation.

18. J. N. Siddall 1982 *Optimal Engineering Design: Principles and Applications.* , New York: Marcel Dekker, Inc..

19. NAg E04UCF 1990 *NAg Mark 14 Reference Manual.* , Oxford: Numerical Algorithms Group Ltd..

20. D. E. Goldberg 1989 *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley.

21. S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi 1983 *Science* **220**(4598), 671-680. Optimization by simulated annealing.

22. M.R. Anderberg 1975 *Cluster Analysis for Applications.* Academic Press.

23. X. Yin and N. Germay 1993 *Proceedings of the International Conference on Artificial Neural Nets and Genetic Algorithms*, *ed. R. F. Albrecht, C. R. Reeves and N. C. Steele*, *Innsbruck*, *Springer-Verlag*, 450-457. A Fast Genetic Algorithm with Sharing Scheme Using Cluster Methods in Multimodal Function Optimization.

24. J. J. Grefenstette 1984 A user's guide to GENESIS. Technical Report CS-83-11. , Vanderbilt Univ..

25. L. J. Eshelman 1991 *Foundations of Genetic Algorithms (FOGA-90)*, *ed. G. J. E. Rawlins*, *Bloomington, IN*, *MKAUF*, 265-283. The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination.

26. J. D. Schaffer 1994 *private communication*, , . The CHC genetic algorithm pushed harder.