AIAA JOURNAL
Vol. 36, No. 1, January 1998

# New Approach to Constrained Shape Optimization Using Genetic Algorithms

M. C. Sharatchandra,* Mihir Sen,† and Mohamed Gad-el-Hak‡
*University of Notre Dame, Notre Dame, Indiana 46556*

**A robust genetic algorithm for constrained functional optimization is described. The function being sought is represented both in a piecewise-linear fashion and in two different types of orthogonal series representations, satisfying in each case specified end conditions of both Dirichlet and Neumann types. The search for the optimal function is translated to one of determining the coefficients of a series expansion, and a genetic algorithm is developed for this purpose. The method is validated in terms of test problems for which the global optimum solutions are known. The results indicate that, if the population size of the chromosome pool is held constant, the performance of the piecewise-linear-representation approach deteriorates considerably as the number of degrees of freedom increases. In contrast, the orthogonal series representations do not suffer from this drawback, and a significant reduction in the population size can be achieved. Therefore, the latter methodology offers a far more efficient approach to functional optimization than previously attempted. The developed methodology was applied to the determination of an optimal micropump shape. The genetic algorithm uncovered shapes that were nonintuitive but yielded vastly superior pump performance.**

## I. Introduction

THE basic underlying principle of genetic algorithms (GAs) is that of the Darwinian evolutionary principle of natural selection,[1] wherein the fittest members of a species survive and are favored to produce offspring. The first mathematically rigorous formalism of evolutionary algorithms is due to Holland[2] of the University of Michigan. GAs have since been extensively used to solve optimization problems where conventional methods are either inapplicable or inefficient. In the context of GAs, the members of a species may be regarded as candidate solutions to a problem under investigation. Note in GAs, the concept of a species is superficial, having no tangible context in nonbiological applications. The members are ranked according to how well they satisfy a certain criterion, and the fittest members are most favored to combine amongst themselves to form the next generation of members, which then replace the preceding generation. The idea behind this is that fitter members tend to produce even fitter offspring, which represent better solutions to the problem at hand. The literature on GAs is exhaustive; the books by Holland[3] and Goldberg[4] provide an excellent introduction to the subject. The reader is referred to Goldberg et al.[5] and to Chambers[6] for extensive bibliographies. This paper is not intended to provide a comprehensive introduction to the subject, and it is assumed that the reader has at least a cursory knowledge of the basic features of genetic algorithms. The primary objective of the present work is to introduce a new efficient approach to constrained shape optimization using a GA. Before going into some of the specifics of GAs, we will outline a class of problems to which they are particularly suited.

### A. Shape Optimization Problem

Many practical problems involve the determination of an optimal shape. The ballistic trajectories of missiles provide an example, as do the shapes of airfoils that maximize lift-to-drag ratios or vehicle configurations that minimize skin-friction and pressure drag. In many cases variational methods can be used to determine the optimum shape, provided an appropriate variational principle exists. The efficacy of variational methods has been demonstrated in the context of the well-known Brachistochrone problem, proposed by Johann Bernoulli in 1696. Most shape optimization problems of practical interest, however, are not amenable to the method of variational calculus, and recourse to methods involving a discrete representation of the shape to be determined is necessary. In such methods, the nature of the objective functional $I$ is changed from $I: C[a, b] \to \boldsymbol{R}$ to $I: \boldsymbol{R}^N \to \boldsymbol{R}$, so that the shape optimization problem is reduced to a finite dimensional one. Here, $C[a, b]$ is the space of all continuous functions in the interval $[a, b]$, and $\boldsymbol{R}^N$ is an $N$-dimensional Hilbert space.

Gradient-based optimization methods require computations of both the objective functional and its derivatives with respect to each of the degrees of freedom characterizing the shape. This is often not possible in practice and otherwise tends to be CPU intensive when the functional $I$ is expensive to compute. A deterioration in performance of gradient/Hessian-based algorithms is often evident as the number of degrees of freedom increases. Another limitation is that the gradient methods invariably yield a local and not a global optimum. Some of the difficulties of gradient-based optimization methods can be alleviated by using discrete sensitivity analysis. Recently, Burgreen and Baysal[7] used this approach in the context of aircraft wing design based on the Euler equations. In their analysis, the wing shape was approximated by a Bezier–Bernstein parameterization. The predicted wing shapes were nonintuitive but plausible, but may be regarded at best as local optimum solutions.

For shape optimization, the major advantages of GAs may be summarized as follows. 1) They search from a population and not from a single parameter set. 2) They are capable of searching for solutions from disjointed feasible domains. 3) They are capable of locating global, as opposed to local, optima. 4) They can operate on irregular functions and those that are nondifferentiable. 5) They do not require the computations of gradients and determinants of Hessian matrices. 6) They operate on an encoding of the parameter set and not on the parameters themselves. Thus, they are capable of handling Boolean variables and those that vary between a few discrete states.

### B. Review of Relevant Literature

Bechert[8] provides an interesting summary of the early use of evolutionary strategies for experimental shape optimization in the context of fluid dynamics. The approaches used were very similar to, and share many of the characteristics of, modern genetic algorithms. Using such a technique, Klockgether and Schwefel[9] attempted to find the optimal shape of a supersonic nozzle. The shape of the nozzle was varied in the experiment through a flexible assemblage of

Received Nov. 26, 1996; revision received Aug. 15, 1997; accepted for publication Sept. 24, 1997. 

*Postdoctoral Research Associate, Department of Aerospace and Mechanical Engineering.
†Professor, Department of Aerospace and Mechanical Engineering.
‡Professor, Department of Aerospace and Mechanical Engineering. Associate Fellow AIAA.

rings of varying diameter. The shape that eventually resulted from the evolutionary experiment was nonintuitive and could only be explained in hindsight. Rechenberg[10] used a six-parameter model to determine the optimal shape of an elbow. Evolutionary strategies were also used in the experiments of Hillebrand et al.[11] to search for the optimal shape of a centrifugal fan casing. More recently, Davalos and Rubinsky[12] used a GA to solve the one- and two-dimensional transient diffusion equations. However, their crossover procedure exchanged entire alleles, and this appears to have necessitated inordinately large population sizes in relation to the simplicity of their test problems. Using an adaptive solution methodology based on GAs, Queipo et al.[13] determined the optimal configuration of electronic components in printed-circuit-board networks.

Applications of GAs in the context of the aerospace sciences have been the focus of considerable recent attention. The research of Gage and Kroo[14] applied GAs to the topological design of nonplanar wings. Anderson[15] used a penalty weight approach to consolidate several different subsonic-wing-design objectives into a single objective functional, but has noted that the solutions had an undesirably strong dependence on the weights. As an alternative, Anderson and Gebert[16] described a multiobjective approach to determining the optimal values of seven parameters characterizing a subsonic wing shape using a Pareto-GA with two different objective functions. Very recently, Doorly et al.[17] used a parallel GA in conjunction with a flow solver to determine optimal airfoil shapes approximated using a $B$-spline. The population size in the latter two studies was 80 and this number seems to be typical of research involving shape optimization using GAs. Other recent examples of the use of GAs in the context of airfoil design are the studies of Yamamoto and Inoue,[18] Obayashi and Oyama,[19] and Oesterle.[20]

### C. Objectives

Large population sizes (80–100) appear to have been necessary in the preponderant majority of shape optimization studies using GAs. The major contribution of the present study is a methodology wherein a significant reduction in the population size can be achieved regardless of the number of degrees of freedom characterizing an optimal shape. This is achieved by using a discrete series representation that exploits the exponential convergence characteristics of orthogonal basis functions. A GA is developed in conjunction with three types of discrete series representations of the function to be determined. The first is a piecewise-linear approximation to the function, the second a Fourier series representation, and the third type uses the Chandrasekhar-Reid orthogonal functions[21] satisfying four end conditions. Thus, the search for the optimal shape is translated to a search for the corresponding optimal vector of coefficients. It will be demonstrated that the major advantage of the latter two representations is the ability to shift the search space for the coefficients from the physical space to a smaller one, where a more efficient and focused search is possible. The present study also departs markedly from other functional optimization studies in that a very strong emphasis is placed on the nature of the discrete representation. In addition, the formulation is extended to accommodate various types of end constraints as well as a global equality constraint. Finally, an alternative to bit-string representation will be demonstrated. This alternative retains the essential features of crossover, and an algebraic equivalent of the mutation operator will be shown.

It will be shown through test problems that whereas the piecewise-linear approach deteriorates in performance as the number of degrees of freedom increases, the orthogonal series approach does not. The first test problem is that of finding the curve with least perimeter that yields a specified area in the Cartesian plane. The next example problem demonstrates how a GA can be used to solve heat diffusion problems by minimizing a variational integral to obtain the temperature profiles. This is believed to be the very first systematic approach to the problem of functional optimization using GAs. By way of a practical application, we use the method to determine the optimal design shapes of the viscous micropump proposed by Sen et al.[22] and shown schematically in Fig. 1.

A final note on the choice of the test problems is in order. They have not been chosen to demonstrate the superiority of GAs over traditional gradient-based optimization techniques. Such is not the intent of the present study. Thus, test problems 1 and 2 were deliberately kept simple so as to clearly illustrate the main aspect of this study: namely, the advantages of using orthogonal series representations of the shape (function) to be optimized, over piecewise-linear discrete ones. Test problem 3 is more challenging and is a topic that naturally arises in an area of current interest to the authors. Though it can certainly be solved using other optimization schemes, the micropump problem provides us with a good test bed for developing the present GA methodology.

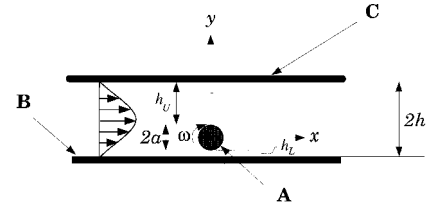



**Fig. 1 Schematic of the basic viscous micropump.**

## II. Problem Definition

In the present study, we consider the following specific functional optimization problem. Determine the shape function $f(x), x \in [a, b]$ that extremizes the functional $I\{f(x)\}$ under one or more equality constraints. We shall, without loss of generality, consider $a = 0$ and $b = 1$. The constraints can be at the ends and/or global. Possible end constraints are a mix of Dirichlet type and Neumann type end conditions as follows: 1) $f$ or its derivative specified at $x = a$ and at $x = b$ or 2) $f$ and its derivative specified at $x = a$ and at $x = b$. The global equality constraint is in the form

$$\sum_{j=0}^{N_d} \int_{p_j}^{q_j} g_j(x) \frac{\mathrm{d}^j}{\mathrm{d}x^j} f(x)\,\mathrm{d}x = C \tag{1}$$

where $N_d$ is the number of continuous derivatives with respect to $x$; $p_j$ and $q_j \in [0, 1], \forall j$; $g_j(x)$ is a weighting function, and $C$ is a constant.

## III. Function Discretization

Implementation of a functional optimization procedure on a digital computer usually requires a discrete representation of the function $f(x)$ to be determined. Accordingly, we represent $f(x)$ in a generalized series representation as follows:

$$f(x) = \mu(x) + \sum_{j=1}^{N_{df}} a_j\, \psi_j(x), \qquad 0 \le x \le 1 \tag{2}$$

where $N_{df}$ is the number of degrees of freedom. The basis functions $\psi_j(x)$ are linearly independent, but do not in general satisfy the end conditions. These are handled by an appropriate choice of the end-constraint function $\mu(x)$. The nature of the optimization problem has now been translated to one of determining the optimal vector of coefficients, $\boldsymbol{a} = \{a_1, a_2, \ldots, a_{N_{df}}\}$. We consider three types of basis functions, which are described in turn in the following subsections.

### A. Piecewise-Linear Approximation (PLA)

Here the function $f$ is represented by a series of line segments connected at the ordinates $a_k, k = 1, 2, \ldots, N_{df}$. The basis function is

$$\begin{aligned} \psi_j(x) &= \frac{x_{j+1} - x}{x_{j+1} - x_j}, && x \in [x_j, x_{j+1}), && j \in [1, N_{df} - 1] \\ &= \frac{x - x_{j-1}}{x_j - x_{j-1}}, && x \in [x_{j-1}, x_j), && j \in [2, N_{df}] \\ &= 0, && \text{otherwise} \end{aligned} \tag{3}$$

where $x_j \in [0, 1]$, such that $x_1 = 0$, and $x_{N_{df}} = 1$. If the end conditions are of the Dirichlet type, then $a_1 = f_0, a_{N_{df}} = f_1$, and $\mu(x) = 0$. If they are of the Neumann type, then $a_1$ and $a_{N_{df}}$ are related to $a_2$ and $a_{N_{df}-1}$, respectively, through the finite difference approximations

$$a_1 = a_2 - f_0' x_2 + \mathcal{O}[x_2] \tag{4}$$

**Table 1 Shape and end-constraint functions corresponding to various end conditions**

| Case | $x=0$ | $x=1$ | $\psi_j(x)$, $j=1,2,\ldots,N_{df}$ | $\mu(x)$ |
|---|---|---|---|---|
| 1 | $f(0)=f_0$ | $f(1)=f_1$ | $\sin(j\pi x)$ | $(1-x)f_0+xf_1$ |
| 2 | $f(0)=f_0$ | $f'(1)=f_1'$ | $\sin[(2j-1/2)\pi x]$ | $f_1'x+f_0$ |
| 3 | $f'(0)=f_0'$ | $f(1)=f_1$ | $\cos[(2j-1/2)\pi x]$ | $f_1-f_0'(1-x)$ |
| 4 | $f'(0)=f_0'$ | $f'(1)=f_1'$ | $\cos(j\pi x)$ | $f_0'x+\frac{1}{2}(f_1'-f_0')x^2$ |

and

$$a_{N_{df}} = a_{N_{df}-1} + f_1'\left(1-x_{N_{df}-1}\right) + \mathcal{O}\left[1-x_{N_{df}-1}\right] \tag{5}$$

where the prime superscript denotes derivative with respect to $x$. Here again $\mu(x)=0$.

The piecewise-linear approximation (PLA) has the following drawbacks.

1) The basis functions are not orthogonal. A large number of degrees of freedom may be required to capture intricate forms of $f(x)$.

2) Because of the $C^0$ continuity of the basis functions, difficulties could arise if higher derivatives of $f(x)$ are present in the global equality constraint. This point will be elaborated upon in Sec. V.B.

### B. Fourier Series Approximation (FSA)

To avoid the drawbacks of the PLA, we consider the use of basis functions that are orthogonal and form complete sets in $C^\infty[0,1]$. The corresponding representation is a Fourier sine or cosine series according to the types of end conditions that the function $f(x)$ is required to satisfy. Table 1 shows the appropriate choices of $\psi_j(x)$ and $\mu(x)$ when a single condition is imposed at each end. The forms of $\mu(x)$ in Table 1 are the lowest degree polynomial expressions satisfying the end conditions. If more than one condition is specified at either end, it must be treated as a local equality constraint. Note that if either or both end constraints are not specified in each case, then the corresponding parameter, $f$ or $f'$, is free to float, i.e., treated as an additional unknown.

### C. Chandrasekhar-Reid Series Approximation (CRA)

In many practical situations, it is quite conceivable that both $f$ and $f'$ are constrained at the endpoints of the interval. For such functions, the FSA is not as convenient, because both the basis functions and their derivatives do not vanish at the endpoints and two additional conditions have to be implemented as local equality constraints. As an alternative, therefore, we propose the following basis functions that vanish at the endpoints along with their first derivatives:

$$\begin{aligned} \psi_{2j}(x) &= C_j(x), \qquad j=1,2,\ldots,N_{df}/2 \\ \psi_{2j-1}(x) &= S_j(x), \qquad j=1,2,\ldots,N_{df}/2 \end{aligned} \tag{6}$$

where $C_j(x)$ and $S_j(x)$ are the Chandrasekhar-Reid functions[21] defined by

$$C_j(x) = \frac{\cosh\lambda_j(2x-1)}{\cosh\lambda_j} - \frac{\cos\lambda_j(2x-1)}{\cos\lambda_j} \tag{7}$$

and

$$S_j(x) = \frac{\sinh\mu_j(2x-1)}{\sinh\mu_j} - \frac{\sin\mu_j(2x-1)}{\sin\mu_j} \tag{8}$$

where $\lambda_j$ and $\mu_j$, $j=1,2,\ldots,N_{df}/2$, are, respectively, the first $N_{df}/2$ roots of

$$\frac{\tanh\lambda}{\tan\lambda} = -1 \tag{9}$$

and

$$\frac{\coth\mu}{\cot\mu} = 1 \tag{10}$$

Although not formally known as Chandrasekhar-Reid functions, these functions were first developed by Chandrasekhar and Reid[21] to solve eigenvalue problems arising in convective instability studies. In their original work, $C_j(x)$ and $S_j(x)$ are defined in the interval $[-\frac{1}{2},\frac{1}{2}]$ and are, thus, referred to as even and odd functions.

The Chandrasekhar-Reid functions are orthonormal in the interval [0, 1]. In addition, $C_j(x)$ and $S_j(x)$ are, respectively, symmetric and skew-symmetric about $x=\frac{1}{2}$. The appropriate form of the end-constraint function satisfying the specified values of $f(0)$, $f(1)$, $f'(0)$ and $f'(1)$ is

$$\begin{aligned} \mu(x) = f_0 + x\{f_0'\} - x^2\{3(f_0-f_1)+2f_0'+f_1'\} \\ + x^3\{2(f_0-f_1)+f_0'+f_1'\} \end{aligned} \tag{11}$$

If the problem calls for fewer than four end constraints, then the unspecified parameters have to be treated as additional unknowns.

In case $f(x)$ is symmetric about $x/2$, all of the coefficients of $S_j(x)$ will be zero. Therefore, including them in the expansion would result in a significant amount of wasteful computation, and Eqs. (6) and (11) should then be replaced by

$$\psi_j(x) = C_j(x), \qquad \forall j \tag{12}$$

and

$$\mu(x) = f_e + f_e'(x-x^2) \tag{13}$$

where $f_0 = f_1 = f_e$ and $f_0' = -f_1' = f_e'$. An analogous treatment holds for the case of a function that is skew-symmetric about $x/2$.

## IV. Binary Representation and Coefficient Extraction

We have proposed three different types of basis functions in each of which the optimization problem has been reduced to one of determining the vector of coefficients $\boldsymbol{a}$ of the resulting series expansion. This section will explain how a GA can be used for that purpose. In a GA, the vector of unknowns $\boldsymbol{a}$ is usually represented as a binary chromosome of length at least $N_{\text{bit}} \times N_{df}$, where $N_{\text{bit}}$ is the length of each substring of the chromosome. (Section VII demonstrates an alternative to binary representation.) The binary chromosome is longer if additional unknowns resulting from $\mu(x)$ are involved. Each chromosome is composed of $N_{df}$ substrings of 0s and 1s, each of length $N_{\text{bit}}$, which are used to encode the specific values of the elements of $\boldsymbol{a}$. The following schematic illustrates a binary chromosome:

$$\underbrace{101011}_{J_1=43}\overbrace{100000}^{J_2=32}\underbrace{010011}_{J_3=19}\overbrace{000111}^{J_4=7}$$

which is a chromosome with $N_{df}=4$ and $N_{\text{bit}}=6$.

Each binary substring of length $N_{\text{bit}}$ represents an integer $J_k$, where $0 \le J_k \le 2^{N_{\text{bit}}}-1$. Thus,

$$J_k = \sum_{j=1}^{N_{\text{bit}}} \delta_j^{(k)} 2^{N_{\text{bit}}-j+1}, \qquad k=1,N_{df} \tag{14}$$

where $\delta_j^{(k)}$ (which is either 0 or 1) represents the $j$th locus (or position) of the $k$th binary substring.

For each basis function, we shall choose a mapping that gives the value of $a_k$ for a given value of $J_k$. Specifically, the integer interval $[0, 2^{N_{\text{bit}}}-1]$ is mapped onto a set of real intervals $[p_k, q_k]$ for each of the $k$ modes. The rate of convergence of the GA is inversely proportional to the width of each interval $[p_k, q_k]$. The endpoints $p_k$ and $q_k$ are chosen differently for each type of basis function. In what follows, we discuss the method of coefficient extraction for each of them.

### A. PLA

In the PLA, $a_k = f(x_k)$ so that $p_k$ and $q_k$ are the (presumably known) lower and upper bounds $f_{\min}$ and $f_{\max}$, respectively, of the function $f(x)$ in the interval [0, 1]. Consequently, $a_k$ is recovered from $J_k$ according to

$$a_k = f_{\min} + \left[\frac{2J_k}{2^{N_{\text{bit}}}-1} - 1\right](f_{\max}-f_{\min}), \qquad k=1,N_{df} \tag{15}$$

### B. FSA

In the context of the FSA, the $a_k$ are values of the Fourier coefficients, which constitute a monotonically decreasing sequence at least when the function being represented is continuous. We will exploit this decay property to limit the search space for the coefficients. Following Titchmarsh[23] and Jeffreys and Jeffreys,[24] $a_k$ are extracted as

$$a_k = \frac{A}{k^{\alpha_0+\alpha_1}}\left(\frac{2J_k}{2^{N_{\text{bit}}}-1}-1\right), \qquad k = 1, N_{df} \tag{16}$$

where $\alpha_0 \in [0, 1]$ is the order of the Lipschitz condition

$$\lim_{h\to 0}[f(x+h)-f(x)] = A\,|h|^{\alpha_0}, \qquad \forall x \in [0, 1] \tag{17}$$

and $A$ is some positive constant of $\mathcal{O}[1]$. Furthermore, as shown by Jeffreys and Jeffreys,[24] $\alpha_1 = 0$ if $\alpha_0 < 1$, and $\alpha_1 = 1$ if $\alpha_0 = 1$ and $f'(x)$ is of bounded variation in $[0, 1]$. The existence and boundedness of $f'(x)$ everywhere in the interval are sufficient but not necessary conditions for $\alpha_1 = 1$ and for the Fourier coefficients to decay at least as fast as $k^{-2}$. Note that the additional unknowns resulting from $\mu(x)$ are extracted according to Eq. (15) and not according to Eq. (16). In the context of functional optimization, it is obviously not known a priori what order of Lipschitz condition the resulting function $f(x)$ will satisfy. Thus, the function that evolves out of the GA is also governed, to some extent, by the assumptions of its continuity.

### C. CRA

The first few coefficients of a Chandrasekhar-Reid series expansion decay relatively slowly. However, the coefficients of the higher modes tend to decay almost exponentially. These observations are heuristic in nature being based on the coefficients of the $C_j(x)$ and $S_j(x)$ series approximations of test functions. For the case of a general function (neither symmetric nor skew-symmetric about $x/2$), the coefficients are recovered according to the criteria

$$\begin{aligned} a_{2k-1} &= \frac{A}{k}\left(\frac{2J_{2k-1}}{2^{N_{\text{bit}}}-1}-1\right), && k = 1, 2, \ldots, N_{\text{co}} \\ &= \frac{A}{\lambda_k^2}\left(\frac{2J_{2k-1}}{2^{N_{\text{bit}}}-1}-1\right), && k = N_{\text{co}}, \ldots, \frac{N_{df}}{2} \end{aligned} \tag{18}$$

for the coefficients of $C_j(x)$ with a similar expression for the coefficients of $S_j(x)$. In Eq. (18), $N_{\text{co}}$ is the cutoff mode number at which the decay criterion changes, typically about 5.

## V. Global Equality Constraints

Most optimization problems involve one or more external conditions that the optimal solution must satisfy exactly. In this study, we restrict the analysis to the case of a single linear equality constraint, which can be expressed in the general form of Eq. (1). If this global constraint is specified as a linear combination of $f(x)$ and its derivatives at discrete points in $[0, 1]$, then the weighting functions take on the form of weighted Dirac-delta distributions.

A vector of coefficients $\boldsymbol{a}^\star = \{a_1^\star, a_2^\star, \ldots, a_{N_{df}}^\star\}$ extracted from a randomly selected binary sequence will not, in general, satisfy Eq. (1). Therefore, some type of scaling of the vector $\boldsymbol{a}^\star$ is needed. Consider first the case

$$\tilde{C} = C - \sum_{j=0}^{N_d}\frac{\mathrm{d}^j\mu(x)}{\mathrm{d}x^j} \neq 0 \tag{19}$$

Then $\boldsymbol{a}$ actually satisfying the constraint is $\boldsymbol{a} = \beta\boldsymbol{a}^\star$, where the scale factor $\beta$ is

$$\beta = \tilde{C}\Big/\sum_{k=1}^{N_{df}} a_k^\star \sum_{j=0}^{N_d}\int_{p_j}^{q_j} g_j(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}\psi_k(x)\,\mathrm{d}x \tag{20}$$

### A. (Near-) Homogeneity in the Basis Functions

It is obvious from Eq. (20) that, when $|\tilde{C}|$ is very close to or equal to zero, the multiplicative scaling procedure described earlier will yield the null vector. Therefore, for homogeneous or near-homogeneous constraints, we propose an additive scaling procedure as follows:

$$a_k = a_k^\star + \beta R_k \tag{21}$$

where $R_k = 1, j \in [2, N_{df} - 1]$, for the PLA, and

$$R_k = \left|\frac{a_k^\star + \epsilon}{a_1^\star + \epsilon}\right| \tag{22}$$

for the FSA and CRA. Equation (22) preserves the decay of $a_k$ determined by Eqs. (16) and (18). Further, $\epsilon$ is some arbitrarily small number introduced so that $R_k \to 1$ as $a_k^\star \to 0, \forall k$. Accordingly, $\beta$ is determined as

$$\beta = \frac{1}{Z}\left[\tilde{C} - \sum_{k=1}^{N_{df}} a_k^\star \sum_{j=0}^{N_d}\int_{p_j}^{q_j} g_j(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}\psi_k(x)\,\mathrm{d}x\right] \tag{23}$$

where $Z$ is given by

$$Z = \sum_{k=1}^{N_{df}} R_k \sum_{j=0}^{N_d}\int_{p_j}^{q_j} g_j(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}\psi_k(x)\,\mathrm{d}x \tag{24}$$

We recommend the multiplicative scaling procedure when $|\tilde{C}| = \mathcal{O}[0.1]$ or greater, to retain the affine nature of the relation between the post- and prescaling shapes.

### B. Special Forms of the Integral in Equations (20) and (23)

In the case of the PLA, $\psi_k(x) \in C^0[0, 1]$, which leads to the question of interpretation of the integrand for $j > 0$ in Eq. (1). The expression for the integral corresponding to the $j$th derivative is

$$I_{j,k} = \int_{p_j}^{q_j} g_j(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}\psi_k(x)\,\mathrm{d}x = \int_{p_j}^{q_j} g_j(x)\,\mathrm{d}\psi_k^{(j-1)}(x) \tag{25}$$

where parenthesized superscripts denote derivatives with respect to $x$. Owing to the discontinuities at the breakpoints, the integral in Eq. (25) does not exist in the Riemann sense for $j \geq 2$. However, existence of the Stieltjes integral requires that $\psi_k^{(j-1)}(x)$ be of bounded variation in $[p_j, q_j]$. This condition holds only for $j \leq 2$ and accordingly places the restriction $N_d \leq 2$ in Eq. (1). For $j = 2$, integration by parts twice reduces Eq. (25) to

$$I_{2,k} = g_2\,\psi_k^{(1)}(x)\Big|_{p_{2+}}^{q_{2-}} - g_2^{(1)}\,\psi_k(x)\Big|_{p_2}^{q_2} + \int_{p_2}^{q_2} g_2^{(2)}\psi_k(x)\,\mathrm{d}x \tag{26}$$

where the third term of the right-hand side (RHS) of Eq. (26) is a Riemann integral, provided the weighting function $g_2(x)$ has continuous second derivatives with respect to $x$.

In summary, the following observations may be made regarding the implementation of equality constraints in the context of the PLA. 1) Third and higher derivatives cannot be handled. 2) Second derivatives cannot be evaluated locally. 3) First derivatives can be evaluated locally, if the points of evaluation do not coincide with breakpoints.

## VI. Fitness Evaluation and Selection Probability

The first step in the implementation of a GA is to set up a parent population of $N_{\text{pop}}$ randomly generated binary chromosomes. Once the coefficients corresponding to a particular chromosome have been recovered and scaled to satisfy the imposed constraints, the objective functional $I(\boldsymbol{a})$ can be evaluated for each of the $N_{\text{pop}}$ chromosomes. The values of $I$ are referred to as the raw fitnesses of the chromosomes. The probability of selection, however, is based

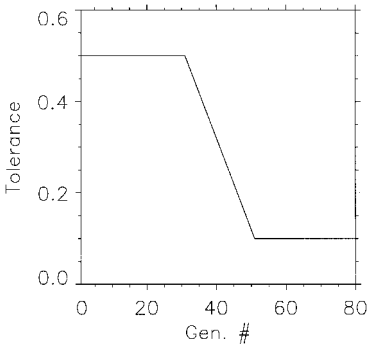



**Fig. 2 Variation of the tolerance schedule, following de la Maza and Tidor.[25]**

on the values of the fitness function $F[I(\boldsymbol{a})]$, which in this study is evaluated from the Boltzmann selection procedure[25] as

$$F(\boldsymbol{a}) = \exp[I(\boldsymbol{a})/T] \tag{27}$$

where $T$ is referred to as a tolerance parameter. It has been shown that, whereas the Boltzmann selection procedure is invariant to translational shifts in the optimization surface, the proportional selection procedure, where $F(\boldsymbol{a}) = I(\boldsymbol{a})$, is not. Such a translational shift is required, for example, to accommodate negative values of the fitness function. A very high value of $T$ can result in the same undesirable consequences associated with the dependence of proportional selection to translational shifts. However, this parameter can be controlled at will by the user, and an effective variation of $T$ with generation number can actually be used to control the performance of the GA, as shown in Fig. 2. Initially, a higher tolerance is used to maintain the diversity of the chromosome pool. If the number and depth of local optima are large, then the initial tolerance should be relatively high and the variation of $T$ should be relatively small. The tolerance is almost always decreased as the tendency toward convergence becomes apparent in later generations.

Once $F_k(\boldsymbol{a})$ has been evaluated for each of the $N_{\text{pop}}$ chromosomes, an estimate of the probability of selection is

$$P_k = F_k \Big/ \sum_{j=1}^{N_{\text{pop}}} F_j, \qquad k = 1, N_{\text{pop}} \tag{28}$$

Note that the $P_k$ constitute a convex set on the interval $[0, 1]$. (Note de la Maza and Tidor[25] use the average of the $F_j$ rather than their sum. Consequently, their $P_k$ are nonconvex and have to be scaled before incorporation into the roulette-wheel procedure.) Clearly, chromosomes with larger fitnesses have a greater probability of selection. This is the algorithmic analog of the Darwinian principle of natural selection. In the context of a roulette wheel, the slot allocated to the $k$th chromosome would subtend an arc of $360P_k$ deg.

## VII. Reproductive Mechanisms

Crossover and mutation are the two dominant mechanisms of reproduction in a GA. We discuss in some detail how crossover is adapted to the present shape optimization problem. An important offshoot of this section will be to demonstrate that bit-string encoding is not an essential component of a GA.

### A. Local Crossover

To achieve a more efficient search for the optimum solution vector, it is necessary that crossover be applied to each binary substring of the chromosome pair selected for reproduction. A vector of crossover points $\boldsymbol{k}$ is generated at random according to

$$k_i = 1 + x\,[N_{\text{bit}} - 1] \tag{29}$$

where $x \in (0, 1)$ is a random number, and roundoff to the closest integer does not hold in Eq. (29), so that $k_i \in [1, N_{\text{bit}}), \forall i$.

Now let $\delta_{1j}^{(i)}$ and $\delta_{2j}^{(i)}$ represent the $j$th loci of the $i$th binary substrings of the first and second parent chromosomes. Likewise, let $\epsilon_{1j}^{(i)}$ and $\epsilon_{2j}^{(i)}$ be the corresponding representation of the first and second offspring chromosomes. Local crossover then implies that the binary substrings of offspring chromosomes are obtained from the parent chromosomes in conjunction with the vector $\boldsymbol{k}$ according to

$$\begin{aligned} \epsilon_{1j}^{(i)} &= \delta_{1j}^{(i)}, \qquad j = 1, k_i \\ &= \delta_{2j}^{(i)}, \qquad j = k_i + 1, N_{\text{bit}} \\ \epsilon_{2j}^{(i)} &= \delta_{2j}^{(i)}, \qquad j = 1, k_i \\ &= \delta_{1j}^{(i)}, \qquad j = k_i + 1, N_{\text{bit}} \end{aligned} \tag{30}$$

The following schematic shows how Eq. (30) is used to produce offspring. For simplicity, $N_{df} = 4$ and $N_{\text{bit}} = 6$. The vector of crossover points is chosen at random to be $\boldsymbol{k} = \{2, 1, 3, 5\}$:

| | | | | | |
|---|---|---|---|---|---|
| Parent Chromosome 1 : | 110111 | 001110 | 001001 | 010010 | |
| Parent Chromosome 2 : | 100101 | 010100 | 001011 | 101010 | |
| | \| | \| | \| | \| | <- -Locations of Shift Bits |
| Offspring Chromosome 1 : | 110101 | 010100 | 001011 | 010010 | |
| Offspring Chromosome 2 : | 100111 | 001110 | 001001 | 101010 | |

which is a multiple-point crossover.

Now according to Eq. (14), the integer numbers corresponding to the $i$th binary substrings of the offspring chromosomes are

$$L_{1i} = \sum_{j=1}^{k_i} \delta_{1j}^{(i)} 2^{N_{\text{bit}} - j} + \sum_{j=k_i+1}^{N_{\text{bit}}} \delta_{2j}^{(i)} 2^{N_{\text{bit}} - j} \tag{31}$$

and

$$L_{2i} = \sum_{j=1}^{k_i} \delta_{2j}^{(i)} 2^{N_{\text{bit}} - j} + \sum_{j=k_i+1}^{N_{\text{bit}}} \delta_{1j}^{(i)} 2^{N_{\text{bit}} - j} \tag{32}$$

Adding Eqs. (31) and (32) yields

$$L_{1i} + L_{2i} = J_{1i} + J_{2i} \tag{33}$$

where $J_{1i}$ and $J_{2i}$ are the integers corresponding to the $i$th parent binary substrings, given by Eq. (14). It also follows from Eq. (33) that, if $N_{\text{bit}}$ bits are used to encode each coefficient, then

$$\max\{L_{1i}, L_{2i}\} \le \min\{2^{N_{\text{bit}}} - 1, J_{1i} + J_{2i}\} \tag{34}$$

and further that $\exists\, \alpha$ such that

$$L_{1i} = \alpha J_{1i} + (1 - \alpha) J_{2i} \tag{35}$$

and

$$L_{2i} = (1 - \alpha) J_{1i} + \alpha J_{2i} \tag{36}$$

Because of the linearity of the $a_i$ vs $J_i$ relationships exemplified in Eqs. (15), (16), and (18), Eqs. (35) and (36) also imply the following statement in the present context of shape optimization using GAs: The pairs of offspring coefficient values resulting from crossover are random complementary linear combinations of the corresponding pairs of parent coefficient values.

A very significant implication of the preceding statement is that the need for bit-string encoding can be obviated, if the weighting factor $\alpha$ in Eqs. (35) and (36) is chosen such that the crossover mechanism described earlier is replicated. This basically involves the choice of an appropriate range of $\alpha$. Details are omitted for the sake of brevity, but the preceding analysis clearly demonstrates how an algebraic representation of crossover is possible. It was found in trials that the choice of convex linear combinations to create offspring resulted in a GA with superior convergence characteristics to those using binary representation.

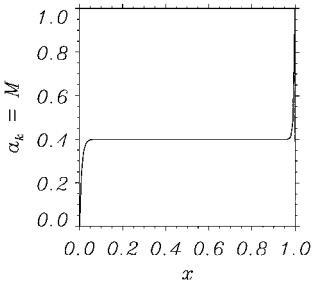



**Fig. 3 Mutation function $M(0.4, 0, 1, 0.01, x)$ plotted against the outcome of the random number generator.**

### B. Mutation

In the context of binary representation, mutation is the random switching of bits (from 0 to 1, or vice versa) in the offspring chromosomes with the mutation probability $p_m$, which is typically $\mathcal{O}[0.01]$. If binary representation is not used, then mutation has to be simulated by an equivalent algebraic representation. We propose a mutation function $a_k = M(z, b, u, p_m, x)$, which acts on an offspring coefficient $z = a_k^\star$ with probability $p_m$ to yield a mutated coefficient $a_k$. This mutated coefficient is bounded below by $b = a_{k,\min}$ and above by $u = a_{k,\max}$, the values of $b$ and $u$ corresponding to the appropriate decay criterion established in Sec. IV.C. The specific form of $M$ is

$$M = b + (z - b)\left\{1 - \exp\left(\frac{x}{p_m}\right)\right\}, \qquad x \le \frac{1}{2}$$

$$= z + (u - z)\exp\left(\frac{-2\{1 - x\}}{p_m}\right), \qquad x > \frac{1}{2} \tag{37}$$

Figure 3 shows a plot of $M(0.4, 0, 1, 0.01, x)$ where $x$ is the outcome of the uniform distribution-based random number generator. Clearly, mutation is restricted to very narrow bands of $x$, which is consistent with the observations yielded by random bit switching in binary representation.

### C. Final Remarks

It is important that mutation precede the implementation of the equality constraints. Once $N_{\text{pop}}$ offspring chromosomes have been generated and mutated, they replace the preceding generation. The so-called elitist strategy is employed, however, in that the fittest parent chromosome from each generation is automatically passed on to the next generation. To maintain the number of the population constant, one of the generated offspring is discarded at random. The process of reproduction is repeated for several generations or until a distinct tendency towards convergence to the global optimum is discerned.

## VIII. Test Problems

By way of two test problems, the optimal solutions of which are known beforehand, the methodology developed in the preceding sections will be validated.

### A. Minimum Arc-Length-to-Area-Ratio Problem

The problem under investigation is that of finding the curve whose arc length is the least of all curves enclosing a certain specified area. Consider the function $f(x)$ in the interval $[0, 1]$, such that $f(1) = 0$ and the area under the curve

$$\int_0^1 f(x)\,dx = A \tag{38}$$

where $A$ is some fixed real number. We pose the shape-optimization problem as the minimization of the arc length of $f(x)$, $x \in [0, 1]$,

$$L\{f(x)\} = \int_0^1 \sqrt{1 + \{f'(x)\}^2}\,dx \tag{39}$$

subject to the equality constraint Eq. (38) and the end constraint $f(1) = 0$.

In the context of the PLA, the functional has to be modified as follows:

$$L\{\boldsymbol{a}\} = \sum_{k=1}^{N_{df}-1} \sqrt{(a_{k+1} - a_k)^2 + \Delta^2} \tag{40}$$

where $\Delta = 1/(N_{df} - 1)$ is the distance between the equally spaced breakpoints $z_k = (k - 1)\Delta$, $k = 1, N_{df}$ on $[0, 1]$. Note that here $a_{N_{df}} = 0$ owing to the end constraint but $a_1$ is to be determined along with the rest of the $a_k$.

It is more convenient to maximize the functional

$$I(\boldsymbol{a}) = A/L\{\boldsymbol{a}\} \tag{41}$$

When $A = \pi/4$, the optimal shape of $f(x)$ is known to be a circular arc.

The fixed control parameters for the GA are chosen as follows: population size = 24, number of bits/substring = 10, crossover probability = 0.75, and mutation probability = 0.05. The computations involved both binary and nonbinary representations. No perceptible differences were observed, confirming the validity of the alternative approach laid forth in Sec. VII.A.

#### 1. *Piecewise-Linear Approach*

Figure 4 shows the evolution of the optimal shape for $N_{df} = 5$. The plotted shapes are the fittest of the 24 members in the population at each generation and represent the average of 10 runs of the GA. By the 25th generation, a distinct resemblance to a circular arc is obtained and only marginal improvement is obtained beyond this point. At $x = 1$, the actual curve has an infinite slope and with five equally spaced points, it is not possible to capture this behavior with any reasonable degree of accuracy. Nevertheless, the rather remarkable closeness to the exact shape is consistent with the observations of Davalos and Rubinsky,[12] who used a GA with a five-node discrete representation to solve the one-dimensional heat diffusion equation with Dirichlet boundary conditions. Their computed temperature

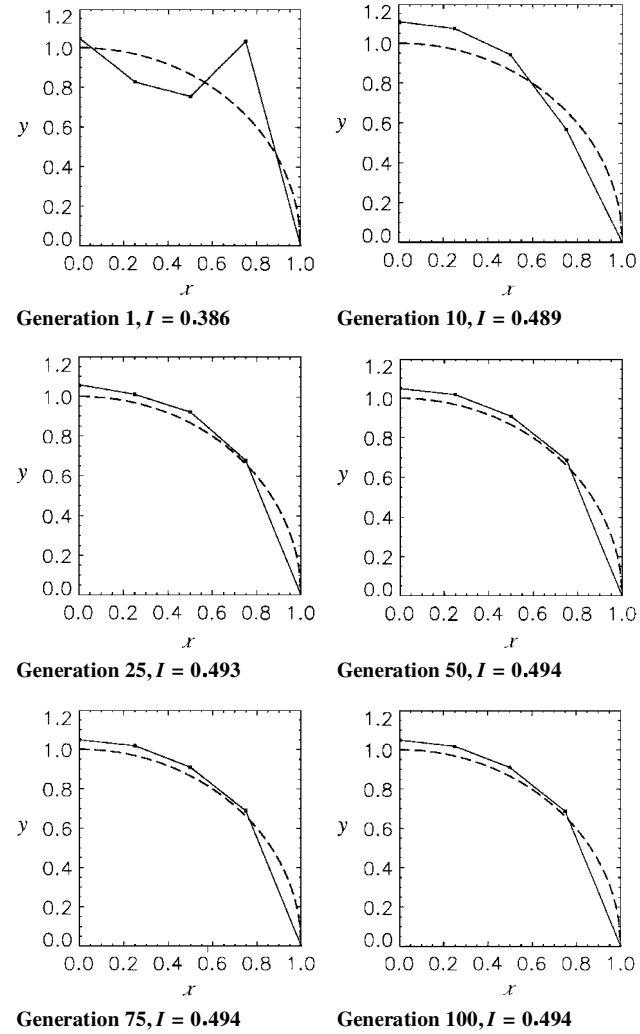



**Fig. 4 Evolution of the fittest shape for $N_{df}$ = 5, using the PLA: - - - -, exact shape (circular arc) and ——, fittest shape at corresponding generation (■).**

profiles were likewise in very good agreement with the exact solutions. Of course, their analysis did not have the present problem of infinite slope at the endpoint to reckon with.

*2. FSA*

Choosing odd cosine multiples of $(\pi x/2)$ as basis functions would be tantamount to cheating because this would presume a zero slope at $x = 0$, which we happen to know a priori. Therefore, our approach here is to use a Fourier sine series. The end-constraint function is

$$\mu(x) = a_{N_{df}+1}(1 - x) \tag{42}$$

where $a_{N_{df}+1}$ is the value of $f(0)$ and is solved for in tandem with the rest of the $a_k$.

Figure 5 shows the evolution of the optimal shape for $N_{df} = 5$. It is noteworthy that $f(0)$ has more or less converged by the 10th iteration, and the rest of the coefficients have readjusted themselves to ensure that $f'(0) = 0$. The infinite slope at $x = 1$ is still not adequately captured, and this cannot be expected since the derivatives of the basis functions have a finite slope there.

*3. Effect of Increasing $N_{df}$*

One would expect that, as more modes are considered, a closer approximation to a circular shape would emerge. Figure 6 reveals, however, that this may not be true in the case of the PLA. The plotted shapes are the fittest shapes at the end of the 100th generation. A consistent deterioration in performance of the GA is seen as $N_{df}$ increases. Although it is conceivable that the shapes may eventually converge to that of a circular arc, this does not appear to justify the computational expense involved. In light of the results of Fig. 6, we submit that the supersonic nozzle design obtained by Klockgether and Schwefel[9] using a similar approach is somewhat questionable in terms of its claimed optimality.

Figure 6 also shows that unlike in the case of the PLA, the performance of the GA does not deteriorate with increasing $N_{df}$ when the FSA is used. As in the preceding case, the results of Fig. 6 represent the average of 10 runs of the GA. Clearly the search for the coefficients in Fourier space is a far more effective approach. It is also noteworthy that far better agreement near $x = 1$ is obtained in almost all cases. For this problem, the values of $\alpha_0$ and $\alpha_1$ in the decay criterion in Eq. (16) were both set to unity. Setting $\alpha_0 = 1$ is justified on the grounds of Lipschitz continuity. However the condition $\alpha_1 = 1$ requires that $f'(x)$ be of bounded variation everywhere in $[0, 1]$. Given that the circular arc clearly violates this latter requirement at $x = 1$, it is interesting that the shapes predicted by the GA are nevertheless in remarkably good agreement with the exact shape.

Figure 7 shows the convergence of the fitness function $I\{f(x)\}$ for $N_{df} = 25$, when the FSA is used. Also shown is the convergence of $f(0)$. The rapid convergence can be attributed in part to the use of the Boltzmann selection procedure in the analysis. The CRA was also used and yielded results graphically indistinguishable from those of the FSA. These results are not presented here for the sake of brevity.

**B. Heat Transfer in a Rectangular Fin**

In this example, we demonstrate the ability of the GA to solve the differential equation of heat diffusion. Our approach differs from that of Davalos and Rubinsky[12] in that our objective function is not based on a finite difference approximation of the energy equation but directly on a variational principle. Steady heat diffusion in a constant

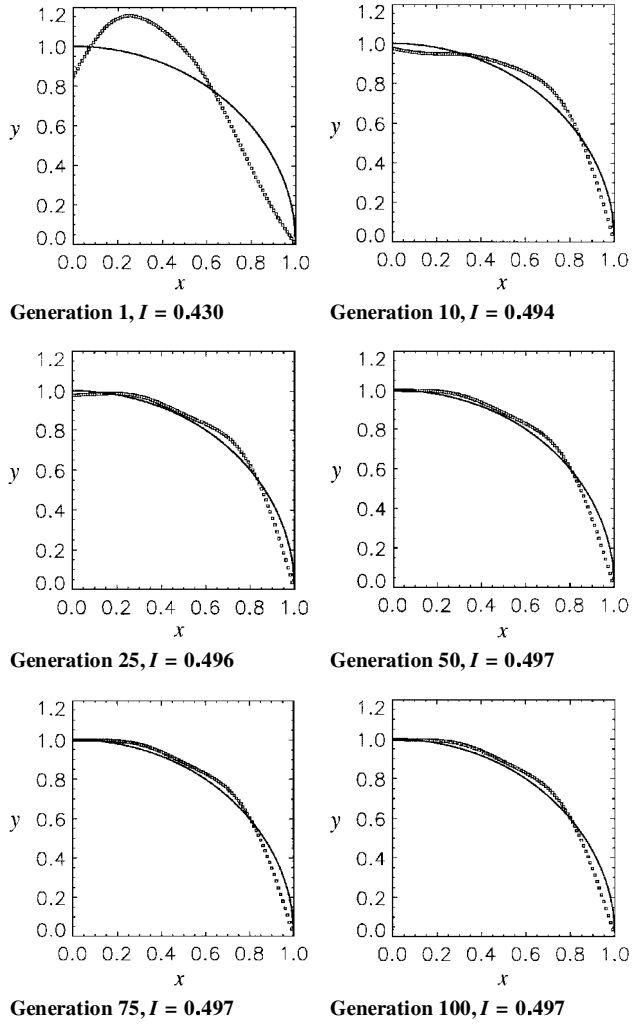



**Fig. 5 Evolution of the fittest shape for $N_{df} = 5$, using the FSA: ——, exact shape (circular arc) and ■, fittest shape at corresponding generation.**

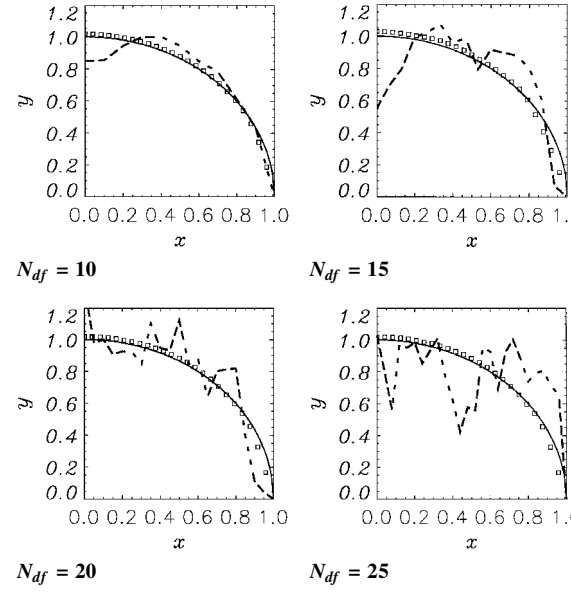



**Fig. 6 Fittest shapes after 100 generations for various values of $N_{df}$: ——, exact shape (circular arc); □, FSA; and - - - -, PLA.**

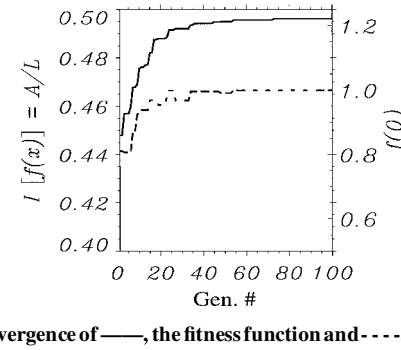



**Fig. 7 Convergence of ——, the fitness function and - - - -, $f(0)$ for $N_{df} = 25$, using the FSA.**

thermal conductivity fin exposed to invariant ambient conditions can be expressed in dimensionless form as

$$\frac{d^2\theta}{dx^2} + Bi\,\theta = 0, \qquad x \in [0, 1] \tag{43}$$

where $Bi$ is the scaled Biot number, $Bi = hL^2/kt$, and $L$ and $t$ are the fin length and thickness, respectively. Also, $h$ is the surface heat transfer coefficient, and $k$ is the constant fin thermal conductivity. The dimensionless temperature is defined as $\theta = (T - T_\infty)/(T_0 - T_\infty)$, where $T_0$ and $T_\infty$ are, respectively, the fin root and ambient temperatures. Typical boundary conditions involve the specification of the fin root temperature and the assumption that the fin tip is insulated. This translates to

$$\theta(0) = 1, \qquad \frac{d\theta}{dx}(1) = 0 \tag{44}$$

Straightforward integration yields the temperature distribution

$$\theta(x) = \exp(-\sqrt{Bi}\,x) + \exp(-\sqrt{Bi})\,\frac{\sinh\sqrt{Bi}\,x}{\cosh\sqrt{Bi}} \tag{45}$$

Equation (43) can also be posed as an optimization problem, wherein the optimal shape of the temperature distribution minimizes the variational integral[26] given by

$$I(\theta) = \frac{1}{2}\int_0^1 \left[\left(\frac{d\theta}{dx}\right)^2 + Bi\,\theta^2\right] dx \tag{46}$$

subject to the equality constraint

$$\frac{d\theta}{dx}(0) + Bi\int_0^1 \theta(x)\,dx = 0 \tag{47}$$

In terms of the basis functions and their derivatives, the global equality constraint can be written as

$$\sum_{k=1}^{N_{df}} a_k \left[\frac{d\psi_k}{dx}(0) + Bi\int_0^1 \psi_k(x)\,dx\right] = \tilde{C} \tag{48}$$

Equation (48) is consistent with the basic form of the equality constraint, Eq. (1), such that $N_d = 1$, $p_0 = p_1 = 0$, and $q_0 = q_1 = 1$. The corresponding weighting functions are $g_0(x) = Bi$ and $g_1(x) = \delta(x)$, where $\delta$ is the Dirac-delta function. The value of the constant $\tilde{C}$ depends on the discrete approximation used. Accordingly, when PLA, FSA, and CRA, respectively, are used,

$$\begin{aligned}\tilde{C} &= 0\\ &= 2 - (Bi/3)\\ &= -\{[1 + (Bi/12)]\,f_0' + (Bi/2)\,(1 + f_1)\}\end{aligned} \tag{49}$$

where $f_0'$ and $f_1$ are additional unknowns in the case of the CRA.

Equation (46) can be recast as the objective functional

$$I(\theta) = 1 - \int_0^1 \left[\left(\frac{d\theta}{dx}\right)^2 + Bi\,\theta^2\right] dx \tag{50}$$

so that the problem is now one of maximization, subject to Eq. (48). When the PLA is used, the first integral on the RHS of Eq. (50) is evaluated as

$$\int_0^1 \left(\frac{d\theta}{dx}\right)^2 dx = \sum_{j=1}^{N_{df}-1}\sum_{k=1}^{N_{df}} \{a_k\,\psi_k'(x_j)\}^2\,(x_{j+1} - x_j) \tag{51}$$

by exploiting the constancy of $\psi_k'(x)$ between breakpoints. When the FSA is used, the boundary conditions [Eq. (44)], suggest that the appropriate basis functions are sines of odd multiples of $\pi/2$

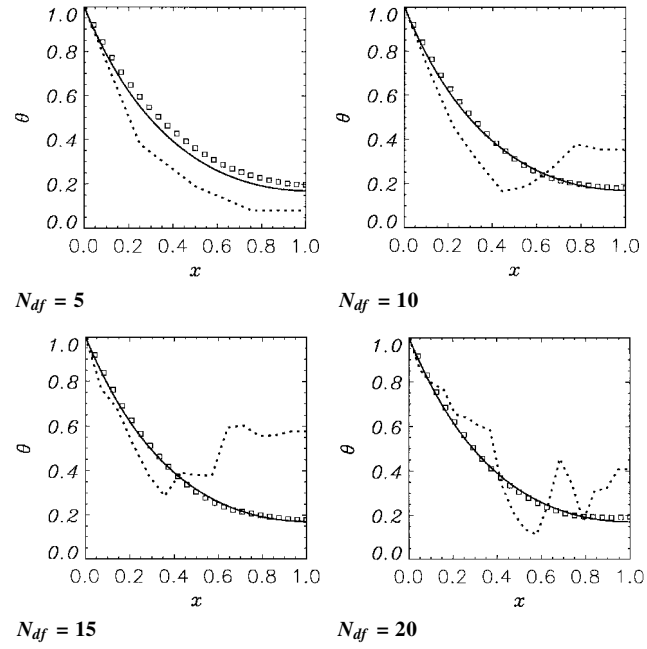



**Fig. 8 Temperature profiles after 100 generations for $Bi$ = 6 and various values of $N_{df}$: ——, exact shape (circular arc); □, CRA; and - - - -, PLA.**

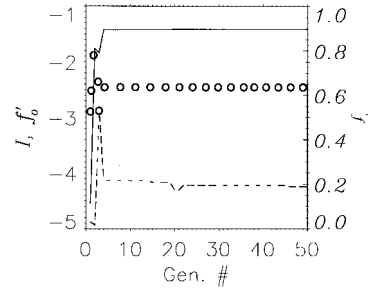



**Fig. 9 Convergence of ——, the objective functional $I$; ○, $f_0'$; and - - - -, $f_1$ (right axis) for $Bi$ = 6; CRA with $N_{df}$ = 10.**

and the equality constraint reduces to a homogeneous linear equation in $\psi_k$ when $Bi = 6$.

The fixed control parameters for the GA are chosen as follows: population size = 12, number of bits/substring = 10, crossover probability = 0.97, and mutation probability = 0.1.

Figure 8 shows the effect of increasing $N_{df}$ when the CRA and PLA are used. In the context of the FSA, the value of $Bi = 6$ was chosen to test the efficacy of the additive scaling procedure described in Sec. V, when a homogeneous equality constraint occurs. The FSA yielded results graphically indistinguishable from the CRA (and for that matter from the exact shape) and in the interest of clarity is not plotted. All of the results reflect the average of 10 runs of the GA with both binary and nonbinary representations. It is encouraging, however, that the additive correction procedure works very well. Confirming the observations of the preceding test problem, the PLA is seen to perform very poorly, especially when $N_{df}$ increases. The rapid convergence of the CRA is noteworthy, despite the presence of the two additional unknowns $f'(0)$ and $f(1)$ arising from Eq. (11).

The convergence of $f_1$ and $f_0'$, as well as the functional $I$, is shown in Fig. 9 when the CRA is used with $N_{df} = 10$. The temperature gradient at the fin root converges somewhat earlier than the tip root temperature. No explanation, however, can be volunteered for this behavior.

## IX. Application to the Micropumping Problem

Sen et al.[22] proposed the mechanism shown in Fig. 1 to pump fluids in microducts. The device essentially consists of an eccentrically placed rotating cylinder A in a parallel-plate channel. The fixed lower wall is B in Fig. 1, and the upper wall C is adjustable. At very low Reynolds numbers, viscous action is the dominant mechanism and the pump has been shown both experimentally[22] and

numerically[27,28] to generate bulk velocities as high as 10% of the rotor surface speed. In the presence of a load acting on the pump, however, smaller plate spacings were shown to yield higher flow rates. This led us to consider the existence of an optimal shape of the upper channel wall, such that the flow rate was a maximum. We use the GA described here for this purpose.

### A. Mathematical Model

The following (dimensionless) geometric and dynamic parameters may be defined, with regard to Fig. 1,

$$s = h/a \tag{52}$$

$$\epsilon = \frac{h_U + a - h}{2a} \tag{53}$$

$$\delta_L = h_L/2a \tag{54}$$

$$\Delta p^* = 4\left[\frac{(P_2 - P_1)a^2}{\rho \nu^2}\right] \tag{55}$$

where the subscripts 1 and 2 denote the inlet and exit stations of the channel. Further, $P$ is the static pressure, and $\rho$ and $\nu$ are, respectively, the fluid density and kinematic viscosity.

The objective functional for this problem is the bulk velocity

$$\bar{u} = \int_{y=0}^{2h} u(L, y)\, \mathrm{d}y \tag{56}$$

where $u(L, y)$ is the streamwise velocity profile at the exit station ($x = L$) of the duct. Here, all quantities are nondimensionalized with the rotor diameter ($2a$) and surface speed ($\omega a$) as length and velocity scales, respectively.

### B. Solution Methodology

The velocity vector field $\boldsymbol{u}(\boldsymbol{x})$ is obtained along with the (scalar) pressure field $p(\boldsymbol{x})$, from the numerical solution of the constant-property, two-dimensional, steady Navier–Stokes equations in the arbitrarily shaped domain shown in Fig. 10. A generalized nonorthogonal curvilinear coordinate system was used in conjunction with a nonstaggered grid arrangement. Elliptically generated grids were deployed, as shown in Fig. 10. The power-law differencing scheme of Patankar[29] used here, reverts to central differencing on account of the low cell-Reynolds numbers involved. Pressure–velocity coupling was achieved with a hybrid PISO-SIMPLER algorithm, and the ensuing system of linearized equations was solved using a line-by-line tridiagonal-matrix algorithm. Additional information regarding the computational methodology is available in Ref. 30 or 31.

Parabolic velocity profiles were imposed on the inlet and outlet boundaries, and no-slip conditions at the upper and lower walls. The $32 \times 27$ grid used for all computations is admittedly coarse, but is considered sufficient for the purposes of demonstration of the GA methodology. The computations are performed for $s = 1.1$, $\epsilon/\epsilon_{\max} = 0.97$, $Re = 1$, and $\Delta p^\star = 10$. When the upper plate is flat [$a_k = 0, \forall k$, in Eq. (57)], the bulk velocity is 0.0485 times the rotor surface speed $\omega a$. It is envisaged that shapes other than the trivial (flat) shape could yield much higher bulk flow rates.

In the context of the GA each of the chromosomes in a particular generation corresponds to a particular wall shape. We shall restrict ourselves to functions that are Lipschitz of order 1 and differentiable everywhere in $[0, 1]$. Consequently, we set $\alpha_o = \alpha_1 = 1$. Our rationale for doing so is based on the intuitive notion that kinks in the upper wall could only result in adverse pressure gradients and, at Reynolds numbers beyond the Stokes regime, in flow separation. The control parameters for the GA were: population size

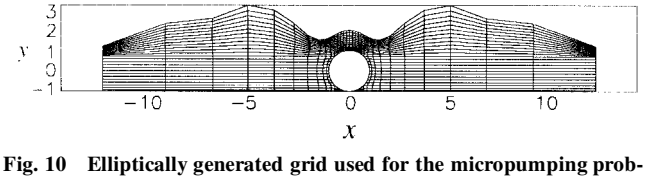



**Fig. 10 Elliptically generated grid used for the micropumping problem.**

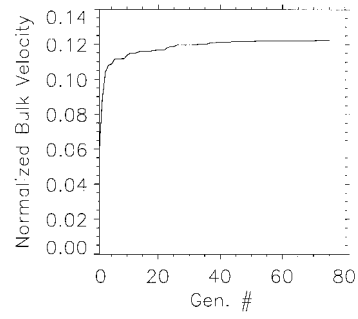
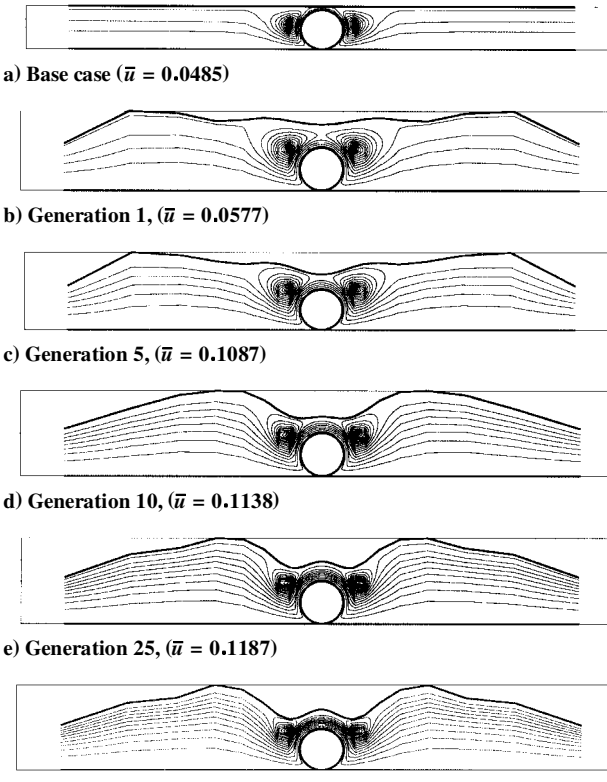



**Fig. 11 Convergence history of the GA for the micropumping problem for $Re = 1, s = 1.1$, and $\epsilon \sim \epsilon_{\max}$; FSA with 10 Fourier modes.**



**a) Base case ($\bar{u}$ = 0.0485)**

**b) Generation 1, ($\bar{u}$ = 0.0577)**

**c) Generation 5, ($\bar{u}$ = 0.1087)**

**d) Generation 10, ($\bar{u}$ = 0.1138)**

**e) Generation 25, ($\bar{u}$ = 0.1187)**

**f) Generation 50, ($\bar{u}$ = 0.1217)**

**Fig. 12 Evolution of optimal micropump shape using the FSA for $Re = 1, s = 1.1$, and $\epsilon \sim \epsilon_{\max}$.**

$= 12$, number of Fourier modes $= 10$, number of bits/substring $= 10$, crossover probability $= 0.97$, and mutation probability $= 0.1$.

Note that there is no global equality constraint in this problem. It is readily envisaged from the results of the preceding section that the PLA is not suitable; it would be very difficult to capture the optimal shapes with only 10 degrees of freedom and rather large population sizes would be required, which detracts from the major objective of the present study. Therefore, we shall use only the FSA and the CRA.

### C. FSA

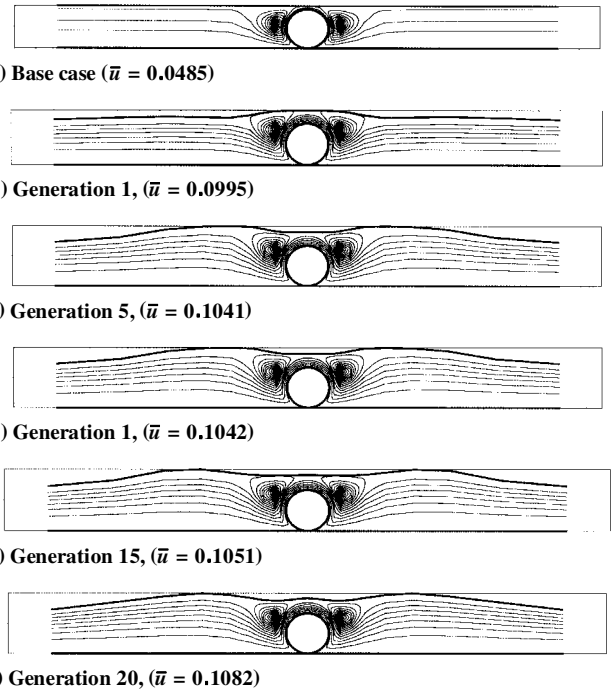
The upper wall shape is approximated as

$$y(x) = 2h + \sum_{k=1}^{N_{df}} a_k \cos(2k-1)\frac{\pi}{2}\frac{x}{L}, \qquad x \in -L, L \tag{57}$$

where we have assumed symmetry about $x = 0$. Trial runs of the GA with general shapes invariably yielded symmetric ones.

Figure 11 shows the convergence rate when the FSA is used. The bulk-velocity functional of Eq. (56) is quite expensive to compute, and this did not permit several runs with different values of the control parameters. In any case, rather rapid convergence to what we believe is the global optimum is reached in less than 50 generations. This corresponds to 551 objective functional evaluations ($N_{\text{eval}} = N_{\text{gen}}[N_{\text{pop}} - 1] + 1$).

Figure 12a shows the base case with a flat upper wall. The evolution of the optimal pump shape is shown in Figs. 12b–12f. Also



a) Base case ($\bar{u}$ = 0.0485)

b) Generation 1, ($\bar{u}$ = 0.0995)

c) Generation 5, ($\bar{u}$ = 0.1041)

d) Generation 1, ($\bar{u}$ = 0.1042)

e) Generation 15, ($\bar{u}$ = 0.1051)

f) Generation 20, ($\bar{u}$ = 0.1082)

**Fig. 13 Evolution of optimal micropump shape using the CRA for $Re = 1$, $s = 1.1$, and $\epsilon \sim \epsilon_{max}$.**

shown are the computed streamlines, as well as the bulk velocity for each case. The determined shapes are nonintuitive but physically plausible. Further computations probably on finer grids may be required, however, before an interpretation of the results can be made based on the physics of the problem. In the vicinity of $x = 0$, the effects of the changes from generation to generation are quite significant as evidenced in the values of the bulk velocities. The two corotating vortices surrounding the rotor are quite large for the tentative shape corresponding to the first generation. These vortices tend to act as blockages in the flow path. In subsequent generations, however, the evolution of the humps on either side of the rotor has the effect of reducing the size of these vortices, resulting in increased flow across the pump. The GA has resulted in wall shapes that yield 2.5 times the original bulk velocity.

### D. CRA

Despite recent impressive advances in the technology of microfabrication,[32–34] the rather complex shapes characteristic of Fig. 12 could pose difficulties in terms of actual fabrication. We, therefore, explore the use of shapes with gentler slopes, without a severe drop in performance. As a means of achieving this, we use the CRA with the additional conditions that $y'(-L) = y'(L) = 0$. Figure 13 shows the rather rapid evolution of the optimal shapes when this approach is employed. Insignificant changes were observed past the 20th generation. Note that the conditions $y'(-L) = y'(L) = 0$ do not penetrate into the channel and are not apparent even at the channel extremities owing to the relative coarseness of the grids therein. Comparing the labels of Figs. 12f and 13f, we find that a slightly more than 10% drop in the bulk velocity results when the additional constraints associated with the CRA are imposed. However, the flow rate in the pump corresponding to Fig. 13f is still more than twice that of the plane upper-wall shape corresponding to $s = 1.1$. This, coupled with the comparable ease in microfabrication of the shapes in Fig. 13, makes them the preferable optimum pump shapes.

## X. Conclusions

A robust genetic algorithm for constrained functional optimization has been described. The desired objective of using much smaller population sizes than those typical of earlier studies was accomplished. The function being sought was represented both in a piecewise-linear fashion and in two different types of orthogonal series representations, satisfying in each case specified end conditions of both Dirichlet and Neumann types. In addition, the imposition of a linear global equality constraint was described in the context of all three types of representation. The search for the optimal function was translated to one of determining the coefficients of a series expansion and the developed GA was modified accordingly. The Boltzmann selection procedure was used to more effectively search for the global optimum. The method was validated in terms of two test problems for which the global optima are known a priori. With the orthogonal series representations, accurate results were obtained using population sizes an order of magnitude smaller than those used in previous studies. This is perhaps the most significant contribution of this paper because it offers an approach that can drastically reduce the number of often expensive functional evaluations. The results also indicate that if the population size of the chromosome pool is held constant, the performance of the piecewise-linear-representation approach deteriorates considerably, with an increase in the number of degrees of freedom. The Fourier series representation and the CRA do not suffer from this drawback and were found to be far more efficient approaches to functional optimization. A GA was employed to determine the optimal micropump shape. The GA uncovered shapes that were nonintuitive, but ones that yielded vastly superior pump performance compared to the trivial (flat) upper-wall shape. Finally, an alternative to the traditional approach of binary representation has been demonstrated. It is specifically shown that the offspring coefficients are merely complementary linear combinations of the parent ones. Future studies will explore the convergence characteristics of GAs that use convex linear combinations to mimic binary crossover, as some preliminary trials have indicated superior performance. In parting, the use of orthogonal series representations are by no means restricted to GAs. Thus, the most important results of the present study may be exploited in the context of optimization algorithms other than GAs.

## Acknowledgments

This study was performed under a grant from the National Science Foundation, under the Small Grants for Exploratory Research initiative (SGER) Grant CTS-95-21612. The Technical Monitor is Roger E. A. Arndt. We thank Michael de la Maza and Dietrich W. Bechert for their valuable help.

## References

[1]Darwin, C., *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, Oxford Univ. Press, London, 1859.

[2]Holland, J. H., "Genetic Algorithms and the Optimal Allocation of Trials," *SIAM Journal on Computing*, Vol. 2, No. 2, 1973, pp. 88–103.

[3]Holland, J. H., *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge, MA, 1992.

[4]Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989.

[5]Goldberg, D. E., Milman, K., and Tidd, C., "Genetic Algorithms: A Bibliography," IlliGAL Rept. 92008, Univ. of Illinois, Urbana, IL, 1992.

[6]Chambers, L. (ed.), *Practical Handbook of Genetic Algorithms: New Frontiers*, Vol. 2, CRC Press, Boca Raton, FL, 1995.

[7]Burgreen, G. W., and Baysal, O., "Three-Dimensional Aerodynamic Shape Optimization Using Discrete Sensitivity Analysis," *AIAA Journal*, Vol. 34, No. 9, 1996, pp. 1761–1770.

[8]Bechert, D. W., "Some Considerations on Measurements in Fluid Dynamics," *Near-Wall Turbulent Flows*, edited by R. M. C. So, C. G. Speziale, and B. E. Launder, Elsevier Science B.V., New York, 1993, pp. 1019–1027.

[9]Klockgether, J., and Schwefel, H. P., "Two-phase Nozzle and Hollow Core Jet Experiments," *Proceedings of the 11th Symposium of Engineering Aspects of Magnetohydrodynamics*, California Inst. of Technology, Pasadena, CA, March 1970.

[10]Rechenberg, I., *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frommann–Holzboog Verlag, Stuttgart, Germany, 1973.

[11]Hillebrand, N., Neise, W., and Bariskow, B., "Anwendung der Evolutionsstrategie auf die Formgebung eines Radialventilator-Gehäuses im Hinblick auf geringe Schallerzeugung," DAGA'80, VDE-Verlag, Berlin, 1980.

[12]Davalos, R., and Rubinsky, B., "Use of Concepts from Genetics and Evolution to Solve Problems in Heat Transfer," *Thermal Sciences Symposium in Honor of Chancellor C.L. Tien*, Berkeley, CA, 1992.

[13]Queipo, N., Devarakonda, R., and Humphrey, J. A. C., "Genetic Algorithms for Thermosciences Research: Application to the Optimized Cooling of Electronic Components," *International Journal of Heat and Mass Transfer*, Vol. 37, No. 6, 1994, pp. 893–908.

[14]Gage, P., and Kroo, I., "A Role of Genetic Algorithms in a Preliminary Design Environment," AIAA Paper 93-3933, 1993.

[15]Anderson, M. B., "The Potential of Genetic Algorithms for Subsonic Wing Design," AIAA Paper 95-3925, 1995.

[16]Anderson, M. B., and Gebert, G. A., "Using Pareto Genetic Algorithms for Preliminary Subsonic Wing Design," *Proceedings of the AIAA/NASA/ISSMO 6th Symposium on Multidisciplinary Analysis and Optimization*, Pt. 1, AIAA, Reston, VA, 1996, pp. 363–371 (AIAA Paper 96-4023).

[17]Doorly, D. J., Peiró, J., and Oesterle, J.-P., "Optimisation of Aerodynamic and Coupled Aerodynamic-Structural Design Using Parallel Genetic Algorithms," *Proceedings of the AIAA/NASA/ISSMO 6th Symposium on Multidisciplinary Analysis and Optimization*, Pt. 1, AIAA, Reston, VA, 1996, pp. 401–409 (AIAA Paper 96-4027).

[18]Yamamoto, K., and Inoue, O., "Applications of Genetic Algorithms to Aerodynamic Shape Optimization," AIAA Paper 95-1650, June 1995.

[19]Obayashi, S., and Oyama, A., "Three Dimensional Aerodynamic Optimization with Genetic Algorithm," *Proceedings ECCOMAS*, Wiley, New York, 1996.

[20]Oesterle, J.-P., "Aeronautical Optimisation Using Parallel Genetic Algorithms," M.Sc. Thesis, Imperial College of Science and Technology, London, 1996.

[21]Chandrasekhar, S., and Reid, W. H., "On the Expansion Functions which Satisfy Four Boundary Conditions," *Proceedings of the National Academy of Sciences*, Vol. 43, 1957, pp. 521–527.

[22]Sen, M., Wajerski, D., and Gad-el-Hak, M., "A Novel Pump for Low-Reynolds-Number Flows," *Journal of Fluids Engineering*, Vol. 118, No. 3, 1996, pp. 624–627.

[23]Titchmarsh, E. C., *The Theory of Functions*, Oxford Univ. Press, London, 1932.

[24]Jeffreys, H., and Jeffreys, B. S., *Methods of Mathematical Physics*, Cambridge Univ. Press, Cambridge, England, UK, 1962.

[25]de la Maza, M., and Tidor, B., "Increased Flexibility in Genetic Algorithms: The Use of Variable Boltzmann Selective Pressure to Control Propagation," *New ORCA CSTS Conference: Computer Science and Operations Development in their Interfaces*, 1992, pp. 425–440.

[26]Finlayson, B. A., *The Method of Weighted Residuals and Variational Calculus*, Academic, New York, 1972.

[27]Sharatchandra, M. C., Sen, M., and Gad-el-Hak, M., "Navier–Stokes Simulations of a Novel Micropump," *Symposium on Application of Microfabrication to Fluid Mechanics*, DSC-Vol. 59, edited by K. S. Breuer, P. R. Bandyopadhyay, and M. Gad-el-Hak, American Society of Mechanical Engineers, New York, 1996, pp. 225–238.

[28]Sharatchandra, M. C., Sen, M., and Gad-el-Hak, M., "Navier–Stokes Simulations of a Novel Viscous Pump," *Journal of Fluids Engineering*, Vol. 119, No. 2, 1997, pp. 372–382.

[29]Patankar, S. V., *Numerical Heat Transfer and Fluid Flow*, Hemisphere, Washington, DC, 1979.

[30]Sharatchandra, M. C., and Rhode, D. L., "A New Strongly Conservative Finite-Volume Formulation for Fluid Flows in Irregular Geometries using Contravariant Velocity Components, Part 1—Theory," *Numerical Heat Transfer*, Pt. B, Vol. 26, 1994, pp. 39–52.

[31]Sharatchandra, M. C., "A Strongly Conservative Finite-Volume Formulation for Fluid Flows in Irregular Geometries Using Contravariant Velocity Components," Ph.D. Thesis, Texas A&M Univ., College Station, TX, 1995.

[32]Angell, J. B., Terry, S. C., and Barth, P. W., "Silicon Micromechanical Devices," *Scientific American*, Vol. 248, April 1983, pp. 44–55.

[33]Gabriel, K. J., "Engineering Microscopic Machines," *Scientific American*, Vol. 260, Sept. 1995, pp. 150–153.

[34]Hogan, H., "Invasion of the Micromachines," *New Scientist*, Vol. 156, June 29, 1996, pp. 28–33.

A. D. Belegundu
*Associate Editor*