Res Eng Des (1991) 3:75–85

Research in Engineering Design

Dimensional Variable Expansion—A Formal Approach to Innovative Design

Jonathan Cagan[1,*] and Alice M. Agogino[2]
[1]Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania; and
[2]Department of Mechanical Engineering, University of California at Berkeley, Berkeley, California, USA

Abstract. A methodology called Dimensional Variable Expansion (DVE) is presented to formalize the process of design space expansion and design innovation. Unlike parametric design, where the structure of the design is specified and only the parameters are allowed to vary, DVE creates new structures, including new variables, constraints, and a reformulation of the objective function. Via DVE, a body is expanded into multiple regions by dividing along a dimensional variable, and then each region is permitted independent properties. Optimality conditions are used to determine which variables to expand and which regions should be subject to property modifications. With DVE the degrees-of-freedom of a design can be expanded and designs with unique features can be derived.

1 Introduction

Computational algorithms for design innovation are an important area of research. Nevins and Whitney [1989] report that 75% of the cost of a product for its producer is committed during the conceptual design phase. Unfortunately, time constraints during the conceptual design phase limit the length of time allotted for concept investigation. In addition, there are often too many variables to explore to determine the best solution to a design problem. Qualitative computational tools to aid in the rapid generation of innovative designs could greatly increase the quantity of quality concepts considered in the preliminary stages of design.

This paper presents a technique called *Dimensional Variable Expansion* (DVE) which innovates new designs by expanding the design space in an optimally directed manner. To define an innovative design we first classify a design as *routine* or *nonroutine*. Routine designs are derived from common prototypes with the same set of variables or features; the structure does not change. A nonroutine design expands the set of variables or features as compared to a previous prototype. Nonroutine designs are then further classified as *innovative* or *creative*. Although new features or variables are introduced in innovative designs, they bear a resemblance to existing variables or features. Creative designs introduce new design variables or features demonstrating no obvious similarity to variables or features in a previous prototype. We have found that most of the designs obtained by DVE transformations are innovative rather than creative nonroutine designs as they create new variables but bear some resemblance to the original prototype design from which they are derived. Other discussions of design classifications for creative, innovative, or routine design can be found in Coyne et al. (1987), Gero et al. (1988), and Brown and Chandrasekaran (1983, 1984).

In an earlier paper (Cagan and Agogino, 1987), we presented the 1stPRINCE (FIRST Principle Computational Evaluator) methodology which originally utilized the concept of integral division for design space expansion. Integral division is an interesting concept but provokes many important questions: How general is the approach? Can it be formalized? Is it domain independent? Is it automatable? A design concept alone can be an important contribution, but its formalization is imperative to define the general applicability of the technique. In addition, a formalized design approach contributes to an underlying theory of design, stimulating other formal design approaches.

This paper follows up on the 1987 paper by presenting a formal design technique which incorporates the intuitive notion of integral division. DVE is a rigorous approach which classifies expansive behavior of variables and constraints, formalizing a description such that automation of constraint expansion is made possible. DVE begins with an ini-

* *Offprint requests:* Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

tial problem formulation as an optimization problem with an objective function and a set of equality and inequality constraints, called a *primitive-prototype*. All design information of the initial design concept is modeled in the primitive-prototype. DVE expands the design space by introducing new variables and potentially new behavioral characteristics to the primitive-prototype by creating discontinuities along a spatial dimension. Because typical application of DVE demonstrates an obvious mapping to previous designs, we classify DVE as an innovative design method. Note that DVE is an independent technique which can be incorporated into the 1stPRINCE methodology or can be employed independently.

DVE finds its roots in optimization theory. Optimization analysis dictates what dimensional variables should be expanded by DVE. A hierarchy of optimization techniques can then be used to perform an analysis on an expanded primitive-prototype. A monotonicity analysis generates cases of active constraints based on a qualitative form of the Karush-Kuhn-Tucker conditions of optimality (Papalambros and Wilde, 1988). An active constraint is one which affects the location of the optimal solution. Active inequality constraints behave as tight equality constraints at optimality. Those cases which can be instantiated to at least one feasible artifact are defined as *prototypes*. A prototype represents a class of designs of equivalent features; it is the prototype which must be analyzed for its innovative potential and application of DVE if appropriate. Numerical optimization is also utilized when required to analyze a specific design artifact. Although symbolic solutions are preferred for their generality and insight, most industrially relevant problems require numerical solutions due to their complexity. Although only a simple example is presented in this paper to clarify the theory of DVE, it is the numerical analysis of the expanded prototypes which could make DVE useful in practice.

The prototypes generated with DVE are *optimally directed. Optimally directed* design is an approach to design which attempts to determine optimal regions of the design space by directing the search toward improving the objectives and eliminating suboptimal or dominated regions of the design space. The result is to reduce the size of the search space and gain insight as to the desirable directions for improving the design variables. The prototypes are not necessarily the globally optimal solutions but rather designs which converge toward that optimum based on available knowledge (Cagan and Agogino, 1991).

DVE is presented as a domain-independent technique with a formal theoretical framework that enables its general automation. In addition, as demonstrated in the next section, DVE presents a way of thinking about a design problem to refine its performance and to understand how its variables are coupled.

2 Conceptual Illustration of DVE

DVE expands the design space of a known primitive-prototype to develop superior designs; composite structural designs are one potential outcome of this process. Design of structural components using composite materials is an important area of design application in industry. Composite structures can exhibit superior stiffness and strength for a reduced weight over conventional designs. For example, automobile manufacturers are considering the use of composite designs for drive shafts. Instead of utilizing conventional steel rods of circular cross section, an aluminum hollow tube is used as a base with layers of graphite-epoxy on the exterior. For the same performance criteria, the weight can be reduced on the order of 60–70%. In order to develop the composite solution, the variables describing the initial design must be expanded to create new variables which lead to the superior composite solution.

Figure 1 is used as a conceptual illustration of application of DVE to a solid cylindrical rod under torsion load where weight is to be minimized. Analysis of the expanded primitive-prototype generates

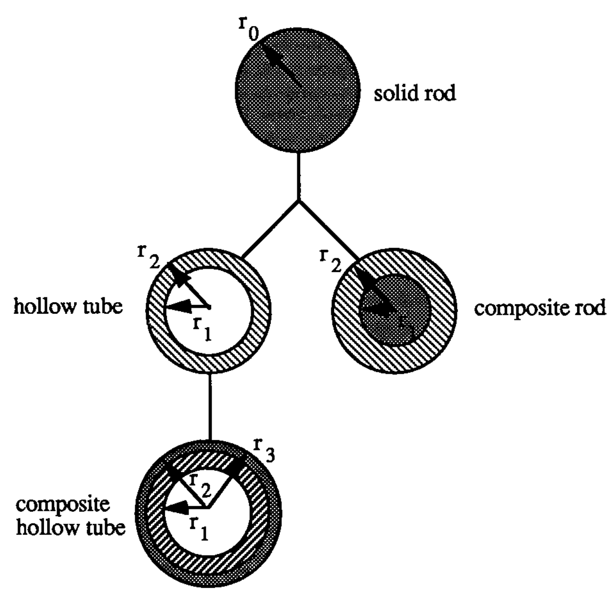

Fig. 1. Concept of DVE.

the solutions of a hollow tube and a composite rod. A further application of DVE on the hollow tube leads to a composite hollow tube. Optimization analysis after DVE leads to a basic set of design constraints and optimally directed solutions. With an appropriate choice of materials, the composite hollow tube demonstrates superiority to the solid rod, justifying the use of composite drive shafts and providing optimally directed design geometries for design application.

Note that DVE expands the design by introducing new radii and material properties in each region. It is this discontinuity of material properties which can lead to a superior design solution. The next section presents the theory of DVE, classifying the expansive properties of design variables and constraints. The way discontinuities along the spatial dimensions are asserted and the resulting introduction of new design variables and constraints is discussed.

3 DVE

3.1 Introduction

In this framework, a *design space* consists of a set of design variables bounded by the set of equality and inequality constraints which model a primitive-prototype. In the design space there are design variables, called *dimensional variables*, that are associated with the coordinate space. A *body* is the geometrical object associated with the dimensional variables of a primitive-prototype. Associated with the body are properties represented by other design variables. A body can be subdivided into a number of *regions*, where a region is a section of a body which may be independently modeled and which may have independent properties and features. Note that a body could be a subregion of a larger body.

In DVE, a body is subdivided along a dimensional coordinate, associated with dimensional variables, into subregions; certain properties and relations are then permitted to be independent of each other. A body is modeled by physical properties which define the body and constitutive relations which describe how the body behaves. When the properties and relations are equivalent between regions, the body behaves as if it were homogeneous; however, when the properties and relations are not equivalent between regions, as is possible after DVE, then the subdivided body may behave differently from the homogeneous body. By subdividing a body and permitting the properties in the regions to be independent from the properties in other regions, new designs can be innovated.

3.2 Primitive-Prototype Definition

More formally, a design space which models a primitive-prototype consists of five different types of variables which characterize the way the variables are used to model the design and their properties for design space expansion.[1] *Coordinate variables*, $\boldsymbol{w}$, represent the coordinate axes of the physical space. *Dimensional variables* are design variables associated with coordinate variables, $\boldsymbol{w}$, and define the boundary of the body, B, or of subregions of B. In the example of section 2, the radius is a dimensional variable. A *system variable*, X_i^{sys}, is defined within a subspace of B as an integral quantity:

$$X_i^{sys}(\boldsymbol{\chi}, z_0, z_1) = \int_{z_0}^{z_1} f_i(\boldsymbol{\chi}, w)\, dw, \qquad (1)$$

where $\boldsymbol{\chi}$ is defined as a vector of variables formed from the elements of a subset of the design variables, f_i is a function of those design variables and the coordinate axis of integration w, and z_0 and z_1 are the limits of integration over w. System variables are functions of $\boldsymbol{\chi}$ and z_0 and z_1; the coordinate variables are integrated out of the problem, a restriction of DVE. Equation (1) is also valid for multiple integrals over a vector of coordinates $\boldsymbol{w}$:

$$X_i^{sys}(\boldsymbol{\chi}, z) = \int_{\boldsymbol{Z}} f_i(\boldsymbol{\chi}, \boldsymbol{w})\, d\boldsymbol{w}, \qquad (2)$$

where $\boldsymbol{Z}$ is the volume under consideration of the body. The vector of dimensional design variables, $\boldsymbol{z}$, is defined from the set of variables z_i derived from the boundary of $\boldsymbol{Z}$. In the example in section 2, the weight and torque are system variables, both being integrated over radius.

Note that dimensional variables are actually system variables because they are determined by integrating over the differential of $\boldsymbol{w}$. We define dimensional variables separately because they have additional characteristics in that they provide geometric boundaries to body regions and provide the mechanism to perform DVE. All characteristics of system variables are applicable to dimensional variables. In summary, $\{X_\alpha^{sys}\} \supseteq \{z_\xi\}$, where $\{X_\alpha^{sys}\}$ is the set of system variables and $\{z_\xi\}$ is the set of dimensional variables.

The fourth type of variables are *region variables*, X_i^{reg}, defined over the space of B as characteristics

[1] This discussion will assume physical domains.

of a region which are not represented as integral quantities:

$$X_i^{reg} = f_i(\chi, z). \qquad (3)$$

In the same example of section 2, the material properties and stresses are region variables.

Finally, the fifth type of variables are *assignment variables*, X_i^{assign}, which are not expanded at all. Rather, an assignment variable is used to designate a global quantity or a temporary variable, associated with a system variable, used in modeling a design problem. For example, the weight of a body is the sum of the weights of its regions (which are system variables); an assignment variable for total weight could be utilized to designate this quantity. Because the assignment variables themselves are not expanded, they do not contribute to added degrees-of-freedom to the design problem. (The assignment variables are not expanded, but the system variables they designate are expanded.)

The set of design variables, V, can be defined as:

$$V = \{X_\alpha^{sys}\} \cup \{X_\beta^{reg}\} \cup \{z_\xi\} \cup \{X_\varphi^{assign}\}, \qquad (4)$$

from which a vector $\boldsymbol{V}$ is formed, where $\{X_\alpha^{sys}\}$ is the set of system variables, $\{X_\beta^{reg}\}$ is the set of region variables, $\{z_\xi\}$ is the set of dimensional variables, and $\{X_\varphi^{assign}\}$ is the set of assignment variables. $\boldsymbol{\chi}$ can now be defined as the vector formed from the set χ:

$$\chi = \{X_\alpha^{sys}\} \cup \{X_\beta^{reg}\} \cup \{X_\varphi^{assign}\}.$$

3.3. Serial and Parallel Constraints

For a primitive-prototype, the design space is bounded by a set of equality and inequality constraints. An inequality constraint, g_i, bounds a function f_i of variables $\boldsymbol{V}$, while an equality constraint, h_i, expresses an equality relation between variables of $\boldsymbol{V}$ represented by a function f_i. Constraints may also contain parameter and constant quantities. Inequality constraints, g_i, take the form:

$$g_i\colon \quad f_i(\boldsymbol{V}) \le 0,$$

while equality constraints, h_i, take the form:

$$h_i\colon \quad f_i(\boldsymbol{V}) = 0.$$

We define valid constraints in DVE as *serial* or *parallel constraints*. Conceptually, we differentiate between serial and parallel constraints to indicate which design criteria apply locally versus those that apply globally. A *serial constraint*, c_i^s, is defined to be a constraint which *defines* global bounds on relations between system variables across a set of regions, initially presented with a single variable in a single region:

$$c_i^s\colon \quad f_i(\boldsymbol{V}) \leqq 0. \qquad (5)$$

The constraint can be either an equality or inequality constraint as shown by Eq. (5). A serial constraint is expanded as a single constraint over a set of regions and often takes the form of a summation of system variables across that set; the summation is the default expansion of serial constraints in the DVE algorithm. In the example of section 2, the torque and weight are serial constraints summed across regions to evaluate their total value.

A *parallel constraint*, c_i^p, is a constraint which represents local constraints for expansion by DVE in each individual region:

$$c_i^p\colon \quad f_i(\boldsymbol{V}) \leqq 0. \qquad (6)$$

Again, the constraint can be either an equality or inequality constraint. A parallel constraint is uniquely expanded into each new region created by DVE. In the same example, bounds on the material properties of each region are parallel constraints.

The objective function, $obj(\boldsymbol{V})$, can be either a parallel or serial relation in that during expansion it can be either maintained as a relation for a specific region or expanded over a set of regions. The design problem is now formulated as:

$$\begin{aligned} &min\colon \quad obj(\boldsymbol{V}) \\ &s.t.\colon \quad \{c_i^s\}\ over\ (\boldsymbol{V}),\ \{c_i^p\}\ over\ (\boldsymbol{V}), \end{aligned}$$

and the primitive-prototype, previously described as an optimization problem with an objective function and a set of equality and inequality constraints, is now formally defined as:

$$\begin{aligned} &primitive\text{-}prototype = \\ &\quad \{obj\}\ bounded\ by\ \{c_i^s\} \cup \{c_i^p\} \\ &\quad over\ \{X_\alpha^{sys}\} \cup \{X_\beta^{reg}\} \cup \{z_\xi\} \cup \{X_\varphi^{assign}\}. \end{aligned} \qquad (7)$$

3.4 Expansion of Critical Variables

DVE expands the design space along a variable which is *critical*, where a *critical variable* is defined as one which influences the objective function and which, when expanded, will create new variables which will also influence the objective function. Criticality is determined both by optimization and heuristic information. Monotonicity analysis is used to identify irrelevant variables that are removed from the list of potentially critical variables. From the remaining variables, those which are dimensional variables are selected for expansion. We assume the dimensional variable is critical and then verify its criticality by observing the effect of its expansion on the objective function. If more than one critical dimensional variable is assumed, each associated coordinate is expanded independently by DVE.

During expansion over a critical dimensional variable by DVE, the body is expanded into subregions and discontinuities are permitted across the coordinate axis, thus introducing new variables. It is these discontinuities which cause design innovation.

3.5 Expansion over a Single Coordinate

When the body is expanded over a single coordinate, the body is divided into n regions and new dimensional variables are introduced at the boundary between those regions. Also during DVE, all region and system variables are expanded by being duplicated in each region. Utilizing Eq. (1), since expansion is performed across one dimensional variable at a time, system variable X_i^{sys} expands to n variables over n regions as:

$$(X_i^{sys}(\chi^k, z_0^j, z_1^j)^j = \int_{z_0^j}^{z_1^j} f_i(\chi^k, w)\, dw, \tag{8}$$

where $j = 1 \ldots n$ and $k = 1 \ldots n$. Thus, the system variable can be a function of all other design variables. Also in Eq. (8), $z_1^j = z_0^{(j+1)}$, z_0^0 replaces z_0 in the initial space, z_1^n replaces z_1 in the initial space, and $\{z_0^j$ *and* $z_1^j\}$ represents the new dimensional variables introduced by DVE, forming vector $\mathbf{z}$. Notation X_i^k designates variable X_i in region k. Considering Eq. (3), region variable X_i^{reg} expands to:

$$(X_i^{reg})^j = f_i(\chi^k, z^k). \tag{9}$$

The new vector of variables χ^j can now be defined from the set of variables as:

$$\{\chi\}^j = \{X_\alpha^{sys}\}^j \cup \{X_\beta^{reg}\}^j \cup \{X_\varphi^{assign}\},$$

where assignment variables are not expanded. The vector of design variables, $\mathbf{V}$, over n regions, is now defined from the set

$$\begin{aligned} V = \{X_\alpha^{sys}\}^1 \cup \cdots \cup \{X_\alpha^{sys}\}^n \cup \{X_\beta^{reg}\}^1 \\ \cup \cdots \cup \{X_\beta^{reg}\}^n \cup \{z_0, z_1\}^1 \\ \cup \cdots \cup \{z_0, z_1\}^n \cup \{X_\varphi^{assign}\}. \end{aligned} \tag{10}$$

Note that there could be other dimensional variables associated with other coordinate axes. However, those other dimensional variables are independent of the $\{z_0, z_1\}$ being expanded and, as far as DVE is concerned, they are expanded as system variables. Thus, for presentation and analysis, all dimensional variables not associated with DVE will be presented as system variables.

If there were ν non-assignment design variables and ν^{assign} assignment variables in the initial space, the number of variables is not expanded to $n\nu + \nu^{assign}$ given a single region expanded to n regions; rather, there are now $n(\nu - 1) + 1 + \nu^{assign}$ variables. There were two dimensional design variables in the original space, so the expansion of the dimensional variables creates $(n + 1)$ additional variables instead of $2n$ additional variables, since $z_1^i = z_0^{(i+1)}$. However, the remaining $(\nu - 2) + \nu^{assign}$ variables are expanded to $n(\nu - 2) + \nu^{assign}$ variables giving a total of $n(\nu - 2) + \nu^{assign} + (n + 1) = n(\nu - 1) + 1 + \nu^{assign}$ variables. Similarly, if r regions are expanded to n regions each, the number of variables is expanded from $\nu_r + \nu_r^{assign}$ total variables in the r regions to $n(\nu_r - r) + r + \nu_r^{assign}$ variables, where ν_r is the number of non-assignment variables and ν_r^{assign} is the number of assignment variables in the r regions.

The dimension of new variable vectors $(\mathbf{X}^{reg})^j$ and $(\mathbf{X}^{sys})^j$, $j = 1 \ldots n$, defined from sets $\{X_\beta^{reg}\}^j$ and $\{X_\alpha^{sys}\}^j$, respectively, is the same as variable vectors $\mathbf{X}^{reg}$ and $\mathbf{X}^{sys}$ in the initial space, and each individual element in each of the vectors $(\mathbf{X}^{reg})^j$ and $(\mathbf{X}^{sys})^j$ is derived directly from the analogous element of vectors $\mathbf{X}^{reg}$ and $\mathbf{X}^{sys}$.

Note that for the special case where the properties of the body remain homogeneous after expansion:

$$\int_{z_0}^{z_1} f_i(\chi, w)\, dw = \int_{z_0^1}^{z_1^1} f_i(\chi^1, w)\, dw + \cdots + \int_{z_0^n}^{z_1^n} f_i(\chi^n, w)\, dw, \tag{11}$$

given $z_1^i = z_0^{(i+1)}$. Thus, for that special case:

$$X_i^{sys} = (X_i^{sys})^1 + (X_i^{sys})^2 + \cdots + (X_i^{sys})^n. \tag{12}$$

So the expansion aspect of DVE is equivalent to division of the integral over system variables, and thus integral division is incorporated into DVE.

In general, during DVE, discontinuities *are* permitted across the coordinate axis at the region boundaries, defined by the new dimensional variables, and the body may not remain homogeneous. These discontinuities introduce independent design variables within each region. Thus:

$$X_i^{sys} \neq (X_i^{sys})^1 + (X_i^{sys})^2 + \cdots + (X_i^{sys})^n. \quad (13)$$

In addition to the variables, parallel and serial constraints are also expanded during DVE. A new parallel constraint is generated within each of the n regions as:

$$(c_i^p)^j: \quad f_i((\boldsymbol{X}^{sys})^k, (\boldsymbol{X}^{reg})^k, z_0^k, z_1^k, \boldsymbol{X}^{assign}) \leqq 0, \quad (14)$$

generating n constraints from initial constraint c_i^p. Again Eq. (14) represents both the equality and inequality constraints. Originally, constraint c_i^p related variables $(\{X_\alpha^{sys}\} \cup \{X_\beta^{reg}\} \cup \{z_0, z_1\} \cup \{X_\varphi^{assign}\})$. Now constraint $(c_i^p)^j$ relates variables $(\{X_\alpha^{sys}\}^k \cup \{X_\beta^{reg}\}^k \cup \{z_0, z_1\}^k \cup \{X_\varphi^{assign}\})$. Note that if the original formulation considered multiple regions, then each constraint could refer to variables from other regions; however, the number of system and region variables and parallel constraints generated remains the same.

Serial constraints are restricted to be inequality or equality constraints which define bounds on relations of variables over a set of regions. During DVE, the serial constraint is expanded in a predefined manner into a single constraint describing a global behavior. Assuming summation of the same function across regions, that summation is introduced as:

$$c_i^s: \quad f_i((\boldsymbol{V})^1) + f_i((\boldsymbol{V})^2) + \cdots + f_i((\boldsymbol{V})^n) \leqq 0, \quad (15)$$

Originally, constraint c_i^s related only variables $\{V_\zeta\}$. Now constraint c_i^s relates variables $\{V_\zeta\}^j$, summed over all j. Serial constraints can model the division of integrals over a set of dimensional variables as demonstrated in Eqs. (11–13), thus restricting the form of a serial constraint to be the summation of one or more system variables over a set of regions when used for integral division.

If originally there were u^s serial constraints and u^p parallel constraints, then after DVE there are u^s serial constraints and nu^p parallel constraints. Thus, there are $(n-1)u^p$ additional constraints in the expanded primitive-prototype.

A new primitive-prototype is now defined by the updated objective function and the new set of variables and constraints:

$$\begin{aligned}&\textit{primitive-prototype} = \{obj\} \text{ bounded by } \{c_i^s\}\\ &\quad \cup \{c_i^p\}^1 \cup \{c_i^p\}^2 \cup \cdots \cup \{c_i^p\}^n \text{ over}\\ &\quad \{X_\alpha^{sys}\}^1 \cup \cdots \cup \{X_\alpha^{sys}\}^n \cup \{X_\beta^{reg}\}^1\\ &\quad \cup \cdots \cup \{X_\beta^{reg}\}^n \cup \{z_0, z_1\}^1\\ &\quad \cup \cdots \cup \{z_0, z_1\}^n \cup \{X_\varphi^{assign}\}. \end{aligned} \quad (16)$$

For expansion over multiple regions, Eq. (16) is defined after all expansions are performed. The problem formulation is now:

$$\begin{aligned}\textit{min}: \quad & obj((\boldsymbol{X}^{sys})^1, \ldots, (\boldsymbol{X}^{sys})^n, (\boldsymbol{X}^{reg})^1, \ldots,\\ & (\boldsymbol{X}^{reg})^n, z^1, \ldots, z^n, \boldsymbol{X}^{assign})\\ \textit{s.t.}: \quad & \{c_i^s\} \text{ over } ((\boldsymbol{X}^{sys})^j, (\boldsymbol{X}^{reg})^j, z^j, \boldsymbol{X}^{assign}),\\ & \{c_i^p\}^j \text{ over } ((\boldsymbol{X}^{sys})^j, (\boldsymbol{X}^{reg})^j, z^j, \boldsymbol{X}^{assign}).\end{aligned}$$

3.6 *Comments on Design Space Expansion*

Because 1stPRINCE expands a design space by expanding dimensional variables via DVE, it goes further by creating an independence in the new regions wherever possible. It is this independence in design variables which gives 1stPRINCE the power to innovate. DVE creates discontinuities in certain variables along the coordinate axis w at the location of the dimensional design variables z_0^j and z_1^j. The discontinuities permitted across the dimensional coordinates equivalently create new variables in the design space consisting of $\{X_\alpha^{sys}\}^1 \cup \cdots \cup \{X_\alpha^{sys}\}^n \cup \{X_\beta^{reg}\}^1 \cup \cdots \cup \{X_\beta^{reg}\}^n \cup \{z_0, z_1\}^1 \cup \cdots \cup \{z_0, z_1\}^n \cup \{X_\varphi^{assign}\}$. The original variables $\{X_\alpha^{sys}\}$, $\{X_\beta^{reg}\}$, and $\{z_0, z_1\}$ are no longer useful; rather the new set of variables, which replace $\{X_\alpha^{sys}\}$, $\{X_\beta^{reg}\}$, and $\{z_0, z_1\}$, are *created* to describe an expanded design space. Again, any dimensional variables not associated with the expansion are presented as system variables because they behave like (and are a subset of) system variables for the expansion.

Within the individual regions, properties may vary in values after expansion; however, the form of the constraints remains the same. Thus, discontinuities are permitted across the coordinate variables, but *the monotonicities of the constraint sets remain the same in the expanded regions as the initial regions*. For DVE, contraints need not be monotonic; however, more qualitative insight about constraint activity may be gained from a monotonicity analysis if the constraints are monotonic.

Boundary conditions, specified by the user, are dependent on the boundary of a region and not on the region itself. Boundary conditions (BCs), then, may or may not expand, depending on whether the region boundary expands. If a BC is dependent on a quantity in the direction being expanded then it will not expand. An example of such a BC will be given in section 5.

In practice, engineers often integrate the form of the integral out of a model and describe the body behavior without the use of the integral. This typically occurs with mechanical/structural design problems which are formulated with some system variables. For example, volume and load differential are all derived by integrating over a region of the body. The designer must formulate the problem by specifying which variables are system, region, dimensional, and assignment, and in addition, which constraints are parallel versus serial, and which constraints are boundary conditions. DVE is a domain-independent technique to expand the design space[2], but domain knowledge is required to properly formulate the problem for use by the technique.

3.7 DVE Algorithm

To summarize the discussion of design space expansion via DVE, we present the DVE algorithm:

```
BEGIN (*INPUT*)
  Formulate design problem as optimization problem;
  Specify dimensional variables, region variables,
      system variables, and assignment variables;
  Present constraints in serial and parallel form;
  Specify boundary conditions;
END;

BEGIN (*DVE*)
  Identify critical dimensional variables;
  FOR each critical dimensional variable DO
  BEGIN
      Expand set of design variables;
      Expand parallel constraints in each region
          expanded;
      Expand serial constraints over each region
          expanded;
      Expand boundary conditions as required;
  END;
END.
```

4 Expanding Degrees-of-Freedom

The degrees-of-freedom (DOF) of a problem indicate the number of variables which may be independently modified within a problem. In optimization problems, if the DOF of a design problem can be increased, then there exists the potential for determining a superior design because more variables can be independently modified. Greater potential for determining superior designs is also an indication of greater potential for determining innovative designs. Cagan (1990) presents a detailed analysis of the effect of DVE on the DOF of a prototype and of the design space under certain restrictions, as summarized below. These results show that DVE can increase the DOF of a prototype and thus increase the potential of creating an innovative design.

The DOF of a design prototype ($DOF_{(prototype)}$) is defined as:

$$DOF_{(prototype)} = V - A,$$

where V is the number of *relevant variables* (those variables, and only those variables, in an active constraint) in the design solution, and A is the number of nonredundant *active constraints* (both active inequality constraints and relevant equality constraints) in the well-bounded candidate solution which forms a prototype.

In general, after DVE, constraint activity among the different prototypes can have varying DOF; however, if the activity of the constraints in the new prototype is maintained the same as the constraints from which they were derived in the original prototype (called prototypes of *dominant constraint activity*) then: *if the number of non-assignment design variables in the regions expanded over a critical dimensional variable by DVE exceeds the number of active parallel constraints in those regions plus the number of regions expanded, then it is* guaranteed *that DVE will expand a prototype to greater than the DOF of the previous prototype.*

The DOF of the feasible design space ($DOF_{(feasible\ design\ space)}$) is an upper bound on the DOF of any realizable prototype in that space, and is defined as:

$$DOF_{(feasible\ design\ space)} = V - E,$$

where V is the number of design variables and E is the number of nonredundant equality constraints. *If the number of non-assignment design variables in the regions expanded over a critical dimensional variable by DVE exceeds the number of nonredundant parallel equality constraints in those regions plus the number of regions expanded, then it is* guaranteed *that DVE will expand a feasible design space to greater than the DOF of the previous feasible design space.*

[2] The approach makes intuitive sense if the design space is a physical space.

5 Example of DVE

As an example of DVE, consider a primitive-prototype of a cylindrical beam under flexural load over region *1* (Fig. 2a).[3] The objective function is to minimize weight, designated as assignment variable W^{assign}; parallel equality constraint h_1 is a bending stress constraint bounded by parallel inequality constraint g_3; parallel equality constraint h_2 is the moment of inertia; serial constraint h_3 and parallel constraint h_4 define the weight of the body via assignment variable W^{assign} where h_3 is expanded by summation; parallel inequality constraints g_1 and g_2 are bounds on the dimensional variables, and serial equality boundary constraints $bc\text{-}h_1$ and $bc\text{-}h_2$ are bounds on length dimensions[4]:

- *objective*: W^{assign},

$$h_1{}^p: \quad \sigma_1{}^{reg} = P\, l1_1 r_1/(2I_1{}^{sys}),$$

$$h_2{}^p: \quad I_1{}^{sys} = \pi r_1{}^4/4,$$

$$h_3{}^s: \quad W^{assign} = W_1,$$

$$h_4{}^p: \quad W_1{}^{sys} = \pi r_1{}^2 \rho (l1_1 - l0_1),$$

$$g_1{}^p: \quad l1_1 - l0_1 \geq L_{min},$$

$$g_2{}^p: \quad r_1 \geq r_{min},$$

$$g_3{}^p: \quad \sigma_1{}^{reg} \leq \sigma_y,$$

$$bc\text{-}h_1{}^s: \quad l1_1 = L,$$

$$bc\text{-}h_2{}^s: \quad l0_1 = 0,$$

(*parameters*: $P, \rho, \sigma_y, L_{min}, r_{min}$),

(*design variables*: $r_1, l1_1, l0_1, \sigma_1{}^{reg}, I_1{}^{sys}, W_1{}^{sys}, W^{assign}$).

The weight is derived from an integral as:

$$W_1{}^{sys} = \int_{l0_1}^{l1_1} \rho \pi r_1{}^2 \, dx = \pi r_1{}^2 \rho (l1_1 - l0_1), \tag{17}$$

where r_1, $l1_1$ and $l0_1$ are dimensional variables. Thus, although there is no explicit integral in the objective, the influence of an integral over the dimensional coordinate x exists.

From a monotonicity analysis there are various prototypes generated; the prototype of interest for this discussion has inequality constraint g_3 active and equality constraints h_1, h_2, h_3, and h_4 relevant, as well as boundary condition equality constraints $bc\text{-}h_1$ and $bc\text{-}h_2$ which are serial constraints always referring to the same boundary region. This prototype has a $DOF_{(prototype)}$ of zero (constraint-bound). There are six non-assignment design variables and four active parallel constraints over one region in the solution, thus the DOF of this prototype will increase if the same constraints are maintained active in each region after DVE. The feasible design space has a $DOF_{(feasible\ design\ space)}$ of one. There are six non-assignment design variables and three parallel equality constraints over one region; thus the DOF of the feasible design space will also increase during DVE.

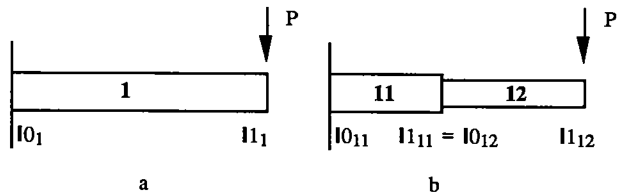

Fig. 2. Example of DVE.

In the prototype, $l1_1$ and $l0_1$ are critical variables in that they have a direct effect on the objective, so DVE is performed over l.[5] In this example the number of regions, n, generated from each initial region during DVE is two. After DVE, there are now two regions in the primitive-prototype, as shown in Fig. 2b, where region subscript 11 designates region 1 in the new space derived from region 1 in the previous space and subscript 12 designates region 2 derived from region 1 in the previous space. New dimensional variables, r_{11}, r_{12}, $l0_{11}$, $l1_{11}$, $l0_{12}$, and $l1_{12}$ are created, where $l0_{11}$ replaces $l0_1$ from the original space, $l1_{12}$ replaces $l1_1$ from the original space, and $l1_{11} = l0_{12}$ by definition of DVE. The number of unique dimensional length variables is now three.

Boundary constraints $bc\text{-}h_1$ and $bc\text{-}h_2$ are constraints on the length of the beam. Since DVE is applied over the length dimensions, the boundary of the end regions (i.e., the clamped condition at the wall and the free condition where the load is applied) will not expand. Thus, those boundary conditions will not expand and act as serial constraints dependent only on variables in the end regions. Cagan (1990) introduces a graph-based representation to implement DVE where decisions on boundary condition expansions are automated.

The weight is expanded serially, summed over all of the regions of the body. Except for the weight

[3] System variables will be designated by superscript "sys," region variables by "reg," and assignment variables by "assign"; remaining variables are dimensional variables.

[4] When utilizing variable names which indicate region numbers in functional relations, subscripts will be used in this example to differentiate region notation from exponents.

[5] Expansion could also be performed over the radius, but this will not be pursued in this example.

constraint, all constraints are parallel constraints as they are relevant to the individual regions. Equality constraint h_2 defines the moment of inertia over an integral of area and is thus a system variable. The problem formulation in the expanded primitive-prototype is given as:

- *objective*: W^{assign},

$$h_{12}{}^p:\ \sigma_{12}{}^{reg} = Pl1_{12}r_{12}/(2I_{12}{}^{sys}),$$

$$h_{11}{}^p:\ \sigma_{11}{}^{reg} = Pl1_{11}r_{11}/(2I_{11}{}^{sys}),$$

$$h_{21}{}^p:\ I_{12}{}^{sys} = \pi r_{12}{}^4/4,$$

$$h_{22}{}^p:\ I_{11}{}^{sys} = \pi r_{11}{}^4/4,$$

$$h_3{}^s:\ W^{assign} = W_1 + W_2,$$

$$h_{42}{}^p:\ W_{12}{}^{sys} = \rho r_{12}{}^2\rho(l1_{12} - l0_{12}),$$

$$h_{41}{}^p:\ W_{11}{}^{sys} = \rho r_{11}{}^2\rho(l1_{11} - l0_{11}),$$

$$h_5{}^p:\ l1_{11} = l0_{12},$$

$$g_{12}{}^p:\ l1_{12} - l0_{12} \geq L_{min},$$

$$g_{11}{}^p:\ l1_{11} - l0_{11} \geq L_{min},$$

$$g_{22}{}^p:\ r_{12} \geq r_{min},$$

$$g_{21}{}^p:\ r_{11} \geq r_{min},$$

$$g_{32}{}^p:\ \sigma_{12}{}^{reg} \leq \sigma_y,$$

$$g_{31}{}^p:\ \sigma_{11}{}^{reg} \leq \sigma_y,$$

$$bc\text{-}h_{11}{}^s:\ l1_{12} = L,$$

$$bc\text{-}h_{21}{}^s:\ l1_{11} = 0,$$

(*parameters*: P, ρ, σ_y, L_{min}, r_{min}),

(*design variables*: r_{11}, r_{12}, $l1_{11}$, $l1_{12}$, $l0_{11}$, $l0_{12}$, $\sigma_{11}{}^{reg}$, $\sigma_{12}{}^{reg}$, $I_{11}{}^{sys}$, $I_{12}{}^{sys}$, $W_{11}{}^{sys}$, $W_{12}{}^{sys}$, W^{assign}),

where parallel constraint h_5 is introduced to define the relation between the dimensional variables. Maintaining constraints h_{12}, h_{11}, h_{22}, h_{21}, h_3, h_{42}, h_{41}, g_{32}, g_{31}, $bc\text{-}h_1$, and $bc\text{-}h_2$ active and relevant, adding constraint h_5 to the active constraints,[6] and considering the 13 new design variables, the prototype generated now has one DOF. The feasible design space has three DOF. Thus, DVE increases the DOF of the prototype which maintains dominant constraint activity, and increases the DOF of the feasible design space. Note that consideration of the material properties as design variables would provide a greater increase in DOF. Although one benefit of DVE is designer materials, a single specified material is considered in this example for simplicity of presentation. Geometrically, for a single material, the beams appear as shown in Fig. 2 and the solution of Fig. 2b is considered innovative as derived from the beam of Fig. 2a. The solution to the optimally directed beam of two regions with the same material in each region is presented in closed-form as:

$$W^{assign} = \pi\rho r_{11}{}^2 l_{11} + \pi\rho r_{12}{}^2(L - l_{11}),$$

$$r_{11} = \left(\frac{4Pl_{11}}{\pi\sigma_y}\right)^{1/3}, \quad r_{12} = \left(\frac{4PL}{\pi\sigma_y}\right)^{1/3},$$

and

$$l_{11} = \left(\frac{3}{5}\right)^{3/2} L = 0.465\,L. \tag{18}$$

The new design is 19% lighter than the original prototype and thus is superior, based on the objective function. In the next generation of four regions, the result is a 13% reduction in weight from this design and a 29% reduction from the first generation. Cagan and Agogino (1991) apply inductive techniques to this problem where the ideal limit representing a tapered beam has a reduction in weight of 40% from the initial cylindrical prototype.

6 Discussion

DVE has been incorporated into the 1stPRINCE design methodology. The methodology is designed to allow automation of as much of the process as possible in symbolic form, such as automation of monotonicity analysis as with the SYMON program [SYmbolic MONotonicity analyzer by Choy and Agogino (1986)], and automation of the Karush-Kuhn-Tucker optimality conditions, as with SYMFUNE [SYMbolic Functional Evaluator by Agogino and Almgren (1987)]. In other publications (Cagan and Agogino, 1987, 1991; Cagan, 1990), the authors have applied the 1stPRINCE methodology to various structures and dynamics problems. By minimizing weight, 1stPRINCE has innovated hollow tubes and composite rods from a solid cylindrical rod under torsion load where the solutions are presented in closed-form and are optimally directed. From a solid rectangular cross-section rod under flexural load, a hollow tube and an I-beam are innovated. Also, by minimizing resistance to spinning, a wheel is invented from a solid rectangular block. Aelion *et al.* (1991) have applied 1stPRINCE

[6] h_5 is introduced because of the *definition* of DVE as discussed in section 3.5; thus it will be active but does not modify the consistency of the dominant constraint activity of prototypes.

to the design of a mixed reactor for maximal conversion and derived the solution of a plug flow reactor.

The primitive-prototype is a framework to represent a design problem. Once in the proper form, DVE can be automatically applied. The detailed definitions have been presented to help the designer in formulating a primitive-prototype, as well as in understanding the characteristics of the design problem itself. The assumption is that the designer has an understanding of an initial design configuration which might satisfy the design constraints; the design innovation occurs based on the initial design configuration.

7 Relation to Other Work

There has been other research into design space expansion for design innovation. Murthy and Addanki (1987) present PROMPT for innovative design of structures, a complementary process to 1stPRINCE. PROMPT takes a satisficing, numerical approach to design using precompiled heuristics and thus solutions are not in closed-form and are not optimally directed; however, when the PROMPT process is relevant the analysis is more efficient than the approach used here by DVE. Ulrich and Seering (1988) utilize the concept of function sharing to derive more elegant designs than previously considered by combining independent functional properties into the same form. Lenat (1983) implements a theory of heuristics in the EURISKO program which mutates a LISP representation to derive new design concepts; mutations are pursued by a quantification of "interestingness." Joskowicz and Addanki (1988) present an innovative design process to derive shape from kinematic description. They have taken a similar philosophy as we take in the need for *formal* design space representations and expansions.

8 Conclusions

DVE has been implemented in Allegro Common Lisp on a MAC II and applied to various structures problems and recently to the design of a mixed reactor in the chemical engineering domain. We foresee application of DVE to numerous high-tech optimization problems, such as disc drives, optical support structures, and aerospace structures, as well as mass-produced composite structures, where minimization of a design feature under tight constraints is imperative to the success of the design. Although the symbolic approach employed in this paper lends insight into design space expansion techniques, our investigation of industrially relevant problems indicates the need for additional numerical analysis.

DVE is a formal mechanism for expansion of a design space to promote innovative design based on a hierarchy of optimization conditions, ranging from qualitative to numerical. Within specific limitations, expanding the design space increases the DOF, an influence in deriving innovative designs, of the design prototypes which maintain dominant constraint activity and of the feasible space of designs.

Acknowledgments. The authors would like to thank Nestor Michelena, Ann Almgren, and Vital Aelion for their discussions on this work. This work was partially funded by the National Science Foundation under grant #DMC-8451622.

References

Aelion, V., J. Cagan, and G. Powers (1991), "Inducing Optimally Directed Innovative Designs from Chemical Engineering First Principles," *Computers and Chemical Engineering* (in press)

Agogino, A.M. and A. Almgren (1987), "Techniques for Integrating Qualitative Reasoning and Symbolic Computation in Engineering Optimization," *Engineering Optimization,* Vol. **12**(2):117–135

Brown, D.C., and B. Chandrasekaran (1983), "An Approach to Expert Systems for Mechanical Design," in: *Proceedings of IEEE Computer Society Trends and Applications, '83,* NBS, Gaithersburg, MD, May, pp. 173–180

Brown, D.C., and B. Chandrasekaran (1984), "Expert Systems for a Class of Mechanical Design Activity," in: *Proceedings of IFIP WG5.2 Working Conference on Knowledge Engineering in Computer Aided Design,* Budapest, Hungary, September

Cagan, J. (1990), *Innovative Design of Mechanical Structures from First Principles,* Ph.D. Dissertation, University of California, Berkeley, CA, April

Cagan, J., and A.M. Agogino (1987), "Innovative Design of Mechanical Structures from First Principles," *AI EDAM: Artificial Intelligence in Engineering Design, Analysis and Manufacturing,* **1**(3):169–189

Cagan, J., and A.M. Agogino (1991), "Inducing Constraint Activity in Innovative Design," *AI EDAM: Artificial Intelligence in Engineering, Design, Analysis and Manufacturing,* **5**(1):47–61

Choy, J.K. and A.M. Agogino (1986), "SYMON: Automated SYbolic MONotonicity Analysis System for Qualitative Design Optimization," *Proceedings of the 1986 International Computers in Engineering Conference,* ASME, Vol. 2, 1986, pp. 305–310

Coyne, R.D., M.A. Rosenman, A.D. Radford, and J.S. Gero (1987), "Innovation and Creativity in Knowledge-Based CAD," in: *Expert Systems in Computer-Aided Design* (Gero, J.S., ed.), North-Holland, Amsterdam, pp. 435–465

Gero, J.S., M.L. Maher, and W. Zhang (1988), "Chunking Structural Design Knowledge as Prototypes," in: *Artificial*

Intelligence in Engineering: Design (Gero, J.S., ed.), Elsevier, Amsterdam, pp. 3–21

Joskowicz, L., and S. Addanki (1988), "From Kinematics to Shape: An Approach to Innovative Design," in: *Proceedings of AAAI-88,* St. Paul, August 21–26, **1**:347–352

Lenat, D.B. (1983), "The Nature of Heuristics II: Theory Formulation by Heuristic Search," *Artificial Intelligence,* **21:** 31–59

Murthy, S.S., and S. Addanki (1987), "PROMPT: An Innovative Design Tool," in: *Proceedings of AAAI-87,* Seattle, WA, July 13–17, **2**:637–642

Nevins, J.L., and D.E. Whitney (1989), *Concurrent Design of Products and Processes,* McGraw Hill, New York

Papalambros, P., and D.J. Wilde (1988), *Principles of Optimal Design,* Cambridge University Press, Cambridge

Ulrich, K., and W.P. Seering (1988), "Function Sharing in Mechanical Design," in: *Proceedings of AAAI-88,* St. Paul, August 21–26, **1**:342–346