Design optimization of multibody systems by sequential approximation


Multibody System Dynamics **2:** 393–415, 1998.

Design Optimization of Multibody Systems by Sequential Approximation

L.F.P. ETMAN, D.H. VAN CAMPEN and A.J.G. SCHOOFS
Department of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received: 29 December 1997; accepted in revised form: 31 July 1998)

Abstract. Design optimization of multibody systems is usually established by a direct coupling of multibody system analysis and mathematical programming algorithms. However, a direct coupling is hindered by the transient and computationally complex behavior of many multibody systems. In structural optimization often approximation concepts are used instead to interface numerical analysis and optimization. This paper shows that such an approach is valuable for the optimization of multibody systems as well. A design optimization tool has been developed for multibody systems that generates a sequence of approximate optimization problems. The approach is illustrated by three examples: an impact absorber, a slider-crank mechanism, and a stress-constrained four-bar mechanism. Furthermore, the consequences for an accurate and efficient accompanying design sensitivity analysis are discussed.

Key words: design optimization, approximation concept, design sensitivity analysis, transient dynamic behavior, mathematical programming.

1. Introduction

Multibody system analysis packages can automatically generate and solve the algebraic relations and differential equations of motion of user-defined mechanical systems. Many packages, however, do not include design optimization routines. To provide optimization facilities, the multibody analysis code has to be extended by a suitable optimization strategy. The question is then how numerical analysis and optimization can be effectively combined to a general multibody design tool. Important aspects are the mathematical formulation of the optimization problem, the type of optimization algorithm to solve this problem, and the actual implementation. Altogether, they should guarantee a reliable and efficient design optimization for a wide range of multibody systems.

The multibody system optimum design problem is defined by design variables, an objective function, and constraints. The design variables arise from the bodies, joints, and force elements present in the multibody system. Examples are the lengths of links, the sliding angles of translational joints, and the stiffness and damping coefficients in spring-damper elements. The objective function and constraints are usually determined by the transient responses following from the

numerical multibody analysis. Common responses are displacements, velocities, and accelerations, as well as induced forces and moments.

In literature on optimization of multibody systems usually a sensitivity-based optimization strategy is applied to a multibody system of fixed topology. Sensitivity-based optimization algorithms have proven to be very effective for smooth problems with large numbers of design variables and constraints. Several successful applications have been reported for planar linkage design [1]. Many have an *ad hoc* character and are concentrated on a specific type of mechanism. The multibody systems approach, however, requires the optimization to work in a more general framework, for both kinematics and dynamics.

Several authors have developed accurate and efficient design sensitivity methods that they consider as the missing link between multibody analysis and optimization. For kinematically driven systems Sohoni and Haug [2] were one of the first to computer-generate the governing equations for both analysis and sensitivity analysis. Dynamics were included by Haug et al. [3]. They studied the design sensitivity analysis of large-scale constrained dynamic systems. Ashrafiuon and Mani [4] proposed to generate the equations for both analysis and sensitivity analysis symbolically instead of numerically. Starting from symbolic formalisms, Bestle [5] gave an extensive description of analysis, sensitivity analysis, and optimization of multibody systems.

Nevertheless, the application of optimization methods to dynamic system design is lagging behind [6]. Most references start from a direct coupling of the analysis and design sensitivity analysis routines with the selected mathematical programming algorithm. However, this direct coupling is hindered by the computational complexity of many multibody systems. Optimization algorithms usually need a lot of objective function and constraint evaluations, whereas the computational cost of large-scale multibody system analysis is often high. Furthermore, the direct coupling yields a black-box optimization. The user is hardly able to control the optimization process, and cannot aid with engineering experience. Both Haug et al. [7] and Erdman [8] stress the importance of an interactive computer aided design tool for a successful design optimization.

This paper proposes to couple the multibody analysis and optimization algorithm by approximation concepts. The basic idea is to generate approximations of the objective function and constraints in a certain part of the design space, and to find the optimum point for this approximate optimization problem. The approximate problem can be easily solved using a mathematical programming algorithm, without any call to the analysis program. Such an interface is computationally more convenient than a direct coupling, and enables the design engineer to influence the optimization process. This approach is commonly employed in structural optimum design [9], but has hardly been used for the optimization of multibody systems. Therefore, we have developed a design optimization tool for multibody systems, starting from a sequence of single-point local approximations of the objective function and constraints. The effectiveness of the approximate optimization approach

is demonstrated for three examples: an impact absorber, a slider-crank mechanism, and a stress-constrained four-bar mechanism with flexible beams.

2. Optimization Problem Formulation

Within the time interval $t \in [t_0, t_f]$, the optimization problem is formulated as follows: find the set of design variable values $\mathbf{b} \in \Re^n$ that will minimize the objective function

$$F(\mathbf{b}) = F(\mathbf{r}(\mathbf{b}, t), \mathbf{b}, t), \quad t \in [t_0, t_f] \tag{1}$$

subject to the inequality constraints

$$g_j(\mathbf{b}, t) = g_j(\mathbf{r}(\mathbf{b}, t), \mathbf{b}, t) \leq c_j \quad \forall\, t \in [t_0, t_f], \quad j = 1, \ldots, m, \tag{2}$$

within the constrained design space

$$b_i^l \leq b_i \leq b_i^u, \quad i = 1, \ldots, n. \tag{3}$$

The functions stored in the column vector $\mathbf{r}(\mathbf{b}, t)$ represent the transient responses calculated from the multibody governing equations. So, the kinematic constraints of the joints and the equations of motion are not included as equality constraints, but are solved separately to yield the responses $\mathbf{r}$.

A great variety of design variables, objective functions and constraints can be defined depending on the specific design application. The above-mentioned formulation covers many different types of optimization problems that may occur for multibody systems. It includes functional relations that can be response dependent, response independent, time dependent or time independent. Usually, the time-dependent multibody responses play a central role. Objective function (1) often contains a max-value or integral operation on the time domain of, for example, a displacement or an acceleration. The constraints (2) usually are parametric of nature, i.e., they have to be satisfied for a complete interval of time points $[t_0, t_f]$. They include constraints such as displacements that may never surpass predefined bounds, or bending stresses that are constrained to a maximum value.

Mathematical programming algorithms generally cannot deal directly with parametric constraints like:

$$g(\mathbf{b}, t) \leq c \quad \forall\, t \in [t_0, t_f]. \tag{4}$$

Such a constraint has to be reformulated in order to remove the time dependence. The most straightforward way is to simply discretize the time interval into n_t time points. Then, the original constraint (4) is replaced by n_t constraints:

$$g_p(\mathbf{b}) = g(\mathbf{b}, t)\Big|_{t=t_p} \leq c, \quad t_p \in [t_0, t_f], \quad p = 1, \ldots, n_t. \tag{5}$$

The time point distribution has to be sufficiently dense to avoid large constraint violations between two adjacent time points. As a consequence of the discretization, time-dependent constraints can greatly increase the number of constraints, and thereby the costs of optimization [10].

Several equivalent constraint formulations have been proposed to remove the time dependence without increasing the number of constraints. In some references the time-dependent constraint (4) is replaced by an integral constraint function similar to

$$g(\mathbf{b}) = \frac{1}{t_f - t_0} \int_{t_0}^{t_f} \max\{g(\mathbf{b}, t), c\} \, \mathrm{d}t \leq c. \tag{6}$$

Constraint (6) will be satisfied as long as $g(\mathbf{b}, t)$ is smaller than or equal to the constraint bound c in the entire interval $[t_0, t_f]$. Whenever a violation occurs in between t_0 and t_f, the integral constraint will be violated as well, which means that the final optimum solution is not affected by the reformulation. Hsieh and Arora [11] stated, however, that from an optimization theory point of view, constraints (4) and (6) are different. This can be understood by noticing that an equivalent integrated constraint

$$g^e(\mathbf{b}) = \int_{t_0}^{t_f} f(g(\mathbf{b}, t)) \, \mathrm{d}t \tag{7}$$

represents the behavior of the time-dependent constraint $g(\mathbf{b}, t)$ on the complete time domain $[t_0, t_f]$ by a single constraint value $g^e(\mathbf{b})$. Information is lost and as a consequence, equivalent constraints tend to blur design trends [10]. Due to the max-value operator, the gradients of constraint (6) vanish at the point where it becomes active [12]. Numerical difficulties during convergence may therefore occur.

Both Grandhi et al. [13] and Hsieh and Arora [11] preferred to replace the original constraint (4) by critical time point constraints:

$$g_p(\mathbf{b}) = g(\mathbf{b}, t)\Big|_{t=t_{mp}} \leq c, \quad t_{mp} \in [t_0, t_f], \quad p = 1, \ldots, n_{mt}. \tag{8}$$

Instead of a complete time discretization (5), they monitored the local maxima and the boundary maxima at t_0 and t_f of the time-dependent constraint functions. Part of these maxima will act as constraints in the optimization problem. Hsieh and Arora [11] only retained the violated critical time points in the active set, where Grandhi et al. [13] used a cutoff level to mark the important maxima. The time points for which the maxima occur depend on the design variable values

$$t_{mp} = t_{mp}(\mathbf{b}), \quad p = 1, \ldots, n_{mt}(\mathbf{b}), \tag{9}$$

and may shift during the optimization process. This drift requires the critical time points to be frequently updated as the optimization progresses, for example after every iteration.

3. Sequential Approximate Optimization

Local approximations of the objective function and constraints are based on function values and derivative values with respect to the design variables in a single point of the design space. Examples are linear or reciprocal approximations. Usually, these approximations are only valid in the vicinity of this design point. Therefore, a search subregion is defined in which the approximate optimization problem is solved. A sequence of approximate optimization cycles has to be performed to reach the final optimum solution. Local approximation concepts are very popular, because large numbers of design variables and constraints can be handled without great difficulty. For multibody systems this is an important advantage, since many constraints may occur, especially if time discretization is used to deal with time-dependent constraints. Moreover, efficient and accurate design sensitivity analysis methods for multibody systems have shown major progress during the last decade (see Section 5). Both aspects indicate that local approximation concepts can be effectively used for the optimization of multibody systems.

The approximate optimization problem of the q-th cycle is defined as: minimize the approximate objective function

$$\tilde{F}^{(q)}(\mathbf{b}) = F(\tilde{\mathbf{r}}^{(q)}(\mathbf{b}, t_p), \mathbf{b}, t_p), \quad t_p \in [t_0, t_f] \tag{10}$$

subject to the approximate constraints

$$\begin{aligned} \tilde{g}_j^{(q)}(\mathbf{b}, t_p) \;&=\; g_j(\tilde{\mathbf{r}}^{(q)}(\mathbf{b}, t_p), \mathbf{b}, t_p) \le c_j \\ &\forall\, t_p \in [t_0, t_f], \quad j = 1, \ldots, m, \quad p = 1, \ldots, n_t^{(q)} \end{aligned} \tag{11}$$

within the search subregion

$$s_i^{l\,(q)} \le b_i \le s_i^{u\,(q)}, \quad i = 1, \ldots, n. \tag{12}$$

Time discretization is used to deal with the time-dependent constraints. Thus, the approximate optimization problem has been reduced to the standard form that can be iteratively solved by a mathematical programming algorithm. So, the term cycle is used for the sequence of approximations, whereas the term iteration applies to the mathematical programming algorithm.

The framework of the approximate optimization problem is depicted in Figure 1. Herein, two types of objective function and constraints are distinguished: objective function and constraints are either completely explicitly known resulting in a direct computation from $\mathbf{b}$, or are related to the approximate responses $\tilde{\mathbf{r}}$. In the latter case, the relationship between the objective function and constraints on

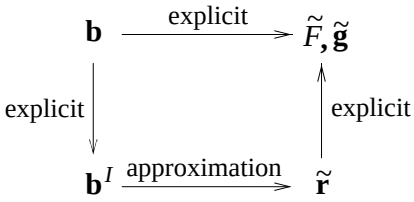


Figure 1. Framework of the approximate optimization problem.

the one hand, and the multibody responses on the other hand, is supposed to be explicitly known or easily calculable. The approximate responses $\tilde{\mathbf{r}}$ are treated as intermediate response quantities that are linearly approximated from the multibody responses $\mathbf{r}$ and their derivatives with respect to intermediate design variables $\mathbf{b}^I$:

$$\tilde{r}_h(\mathbf{b}_h^I, t_p) = r_h(\mathbf{b}_{0h}^I, t_p) + \sum_{k=1}^{n} (b_{hk}^I - b_{0hk}^I) \left(\frac{\partial r_h}{\partial b_{hk}^I} \right)_{\mathbf{b}_{0h}^I, t_p} . \tag{13}$$

The relationship between intermediate design variables $\mathbf{b}^I$ and design variables $\mathbf{b}$ is supposed to be explicitly known. Each response function r_h may have its own intermediate design variables $\mathbf{b}_h^I(\mathbf{b})$.

The introduction of intermediate design variables and intermediate response quantities aims at creating a high quality approximation that yields an efficient and reliable optimization process [14]. The key idea is to improve an approximation of objective function or constraint by incorporating nonlinear behavior that is explicitly known or physically present. Intermediate design variables and intermediate response quantities are commonly employed in structural optimization (see, e.g., [9, 14]). The basic principle is not restricted to any specific area of application whatsoever. For multibody systems at least the same potential is present.

4. Design Optimization Tool

In the current research a design optimization tool has been developed based on local approximations. It has been especially designed for multibody systems and covers time-dependent constraints. The main program structure originates from [14] and is described in further detail in [15]. The optimization process starts with the analysis of an initially proposed design, followed by an evaluation of all constraint functions (constraint screening). Approximation models are generated only for the critical and potentially critical constraints. For these constraints, design sensitivities are calculated. The approximate optimization problem is built, and the region of validity is bounded by so-called move limits. Within this search subregion, the approximate problem is iteratively solved by an optimizer. At the calculated optimum a new cycle can be started.

The approximate optimization problem has to be solved by an appropriate mathematical programming algorithm. When intermediate design variables and

intermediate response quantities are used, usually a smooth but nonlinear programming problem is obtained. To solve it, we use the SQP algorithm of the NAG library [16]. Since only a limited part of the design space is considered, the approximate optimization problem will often have only one optimum solution. However, the approximate optimization problem is not guaranteed to be convex, and therefore more than one local optimum solution may occur. In [17] a global optimizer is proposed for the solution of the approximate optimization problem to yield a globally convergent optimization strategy. For specific types of intermediate design variables convexity can be proved, and in that case an optimizer can be selected that utilizes this quality of the approximation. Well-known examples are convex linearization [18] and the method of moving asymptotes [19, 20].

The constraint screening tries to identify the important constraints in the optimization problem. Constraints that are not critical or potentially critical at the cycle start design $\mathbf{b}_0^{(q)}$ are removed. This can greatly reduce the number of constraints in the approximate optimization problem, and thereby the cost of the numerical optimization. Additionally, design sensitivity information is only required for the retained critical and potentially critical constraints. This gives the opportunity to reduce the cost of design sensitivity analysis (see Section 5).

Great potential exists for constraint deletion if time-dependent constraints are replaced by time point constraints. All constraints can be removed whose values at $\mathbf{b}_0^{(q)}$ are smaller than, e.g., 70% of the constraint bound. Furthermore, for each local maximum of the constraint $g(\mathbf{b}, t)$, only a few time point constraints $g_p(\mathbf{b}, t_p)$ near the maximum have to be retained. The developed optimization tool allows the user to define whether constraints below a prescribed bound are removed, and how many time points are considered near local maxima, and initial and final time. In the following examples, usually two before and two after the maxima above 70% of the constraint bound are selected. Multiple time points at a local maximum instead of a single point are selected to avoid oscillations due to the shift of constraint maxima. This shift may yield curved constraint functions which often cannot be predicted by a single time point constraint, unless a conservative one. Neighboring time point approximations can solve this problem (see Section 6.1).

For each optimization cycle, the quality of the approximations is checked by comparing the approximate objective and constraint values for the newly proposed design with the corresponding multibody analysis values. Differences between the approximated and calculated values are measures for the quality of the generated approximation. So, after the q-th cycle has been completed, the following approximation error is calculated for the objective function

$$E_f^{(q)} = \left| \frac{\tilde{F}^{(q)}(\mathbf{b}_*^{(q)}) - F(\mathbf{b}_*^{(q)})}{F(\mathbf{b}_*^{(q)})} \right| \times 100\%, \tag{14}$$

where $\mathbf{b}_*^{(q)}$ is the proposed optimum design computed from the approximate optimization problem of the q-th cycle. The maximum constraint approximation error is given by

$$E_g^{(q)} = \max_{j=1,\ldots,m} e_j^{(q)} \tag{15}$$

with

$$e_j^{(q)} = \begin{cases} \left| \dfrac{\tilde{g}_j^{(q)}(\mathbf{b}_*^{(q)}) - g_j(\mathbf{b}_*^{(q)})}{g_j(\mathbf{b}_*^{(q)})} \right| \times 100\% & \text{if } g_j \text{ is time independent,} \\ \max\limits_{p=1,\ldots,n_t^{(q)}} \left| \dfrac{\tilde{g}_j^{(q)}(\mathbf{b}_*^{(q)}, t_p) - g_j(\mathbf{b}_*^{(q)}, t_p)}{g_j(\mathbf{b}_*^{(q)}, t_p)} \right| \times 100\% & \text{if } g_j \text{ is time dependent.} \end{cases} \tag{16}$$

Many of the (time point) constraints will not contribute to this error $E_g^{(q)}$ since they have been removed from the approximate optimization problem after the constraint screening. In the same way the maximum constraint violation is calculated:

$$V^{(q)} = \max_{j=1,\ldots,m} v_j^{(q)} \tag{17}$$

with:

$$v_j^{(q)} = \begin{cases} \dfrac{g_j(\mathbf{b}_*^{(q)}) - c_j}{c_j} \times 100\% & \text{if } g_j \text{ is time independent,} \\ \max\limits_{p=1,\ldots,n_t^{(q)}} \dfrac{g_j(\mathbf{b}_*^{(q)}, t_p) - c_j}{c_j} \times 100\% & \text{if } g_j \text{ is time dependent.} \end{cases} \tag{18}$$

Both the approximation errors and the constraint violations are used to determine the size of the search subregion at the start of each new cycle. Poor approximations need the support of the move limit strategy during the optimization much more than high quality approximations. Therefore, a good choice of the search subregion is important for a good convergence of the optimization process. Details about the implemented move limit strategy can be found in [15].

5. Design Sensitivity Analysis

Local approximation concepts require gradient information. For each new approximate optimization cycle, a design sensitivity analysis has to provide the derivatives of the multibody responses with respect to the design variables. Finite differencing is probably the simplest method to obtain the design sensitivities, but it is computationally expensive and suffers from numerical inaccuracies [21]. A better approach is to derive the governing equations of the design sensitivities. A substantial amount of literature is available on this subject, see [15] for an overview. The direct and the adjoint method are briefly outlined below to see why direct differentiation is more suited for the approximate optimization process defined

in the previous sections than the adjoint method. For brevity, this is only shown for pure ordinary differential equations of motion, but for differential algebraic equations the same may be concluded.

Consider a multibody system that is made up of rigid bodies in tree structure (no closed loops). Then the equations of motion can be written as a set of ordinary differential equations:

$$\mathbf{M}(\mathbf{q}, \mathbf{b})\ddot{\mathbf{q}} = \mathbf{Q}^A(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{b}, t), \quad \mathbf{q}(t_0) = \mathbf{q}_0, \quad \dot{\mathbf{q}}(t_0) = \dot{\mathbf{q}}_0. \tag{19}$$

Herein, generalized coordinates $\mathbf{q}$, mass matrix $\mathbf{M}$, generalized forces $\mathbf{Q}^A$, and initial conditions for the state $\mathbf{q}(t_0)$ and $\dot{\mathbf{q}}(t_0)$ can be identified. The column of generalized forces includes externally applied forces and torques as well as forces due to gyroscopic and Coriolis effects. For simplicity, constant initial conditions are considered, although for some applications they may explicitly or implicitly depend on the design variables (see, e.g., [21] where the sensitivity equations of tree-type multibody systems are developed in a more general form). The set of second-order equations (19) can be rewritten in first-order form which yields for holonomic kinematic couplings:

$$\dot{\mathbf{y}} = \mathbf{z}, \quad \mathbf{y}(t_0) = \mathbf{y}_0, \tag{20}$$

$$\mathbf{M}(\mathbf{y}, \mathbf{b})\dot{\mathbf{z}} = \mathbf{Q}^A(\mathbf{y}, \mathbf{z}, \mathbf{b}, t), \quad \mathbf{z}(t_0) = \mathbf{z}_0. \tag{21}$$

The newly introduced generalized positions $\mathbf{y}$ and velocities $\mathbf{z}$ are equal to $\mathbf{q}$ and $\dot{\mathbf{q}}$, respectively.

Equations (20) and (21) are differentiated with respect to a design variable b to obtain

$$\frac{\mathrm{d}\dot{\mathbf{y}}}{\mathrm{d}b} = \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}b}, \tag{22}$$

$$\mathbf{M}\frac{\mathrm{d}\dot{\mathbf{z}}}{\mathrm{d}b} = -\left[(\mathbf{M}\dot{\mathbf{z}})_{\mathbf{y}} - \mathbf{Q}^A_{\mathbf{y}}\right]\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}b} + \mathbf{Q}^A_{\mathbf{z}}\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}b} - \left(\mathbf{M}_b\dot{\mathbf{z}} - \mathbf{Q}^A_b\right) \tag{23}$$

with initial conditions

$$\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}b}(t_0) = \mathbf{0}, \tag{24}$$

$$\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}b}(t_0) = \mathbf{0}. \tag{25}$$

The direct method computes $(\mathrm{d}\mathbf{y}/\mathrm{d}b)(t)$ and $(\mathrm{d}\mathbf{z}/\mathrm{d}b)(t)$ for each design variable $b \in \mathbf{b}$. Subscript $\mathbf{y}$ and b denote partial differentiation with respect to $\mathbf{y}$ and b, respectively.

The adjoint variable approach starts from a response $\overline{r}$ that is an integral function of the generalized positions, velocities and accelerations:

$$\overline{r} = \int_{t_0}^{t_f} p(\mathbf{y}, \mathbf{z}, \dot{\mathbf{z}}, \mathbf{b}, t)\, \mathrm{d}t. \tag{26}$$

Differentiation of Equation (26) gives

$$\frac{\mathrm{d}\overline{r}}{\mathrm{d}b} = \int_{t_0}^{t_f} \left(p_{\mathbf{y}} \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}b} + p_{\mathbf{z}} \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}b} + p_{\dot{\mathbf{z}}} \frac{\mathrm{d}\dot{\mathbf{z}}}{\mathrm{d}b} + p_b \right) \mathrm{d}t. \tag{27}$$

The adjoint method avoids to explicitly calculate sensitivities $\mathrm{d}\mathbf{y}/\mathrm{d}b$, $\mathrm{d}\mathbf{z}/\mathrm{d}b$, and $\mathrm{d}\dot{\mathbf{z}}/\mathrm{d}b$. To accomplish this, one may apply the following procedure that originates from the Lagrange Multiplier Theorem. Firstly, multiply Equation (22) by adjoint variables $\boldsymbol{\mu}^T(t)$, and integrate by parts. Do the same for Equation (23) using adjoint variables $\boldsymbol{\zeta}(t)$. Furthermore, take Equation (23) again multiplied by another set of adjoint variables $\boldsymbol{\xi}(t)$ without integration by parts. Add these expressions to $\mathrm{d}\overline{r}/\mathrm{d}b$ in Equation (27), rearrange, and select the adjoint variables such that the terms with $\mathrm{d}\mathbf{y}/\mathrm{d}b$, $\mathrm{d}\mathbf{z}/\mathrm{d}b$ and $\mathrm{d}\dot{\mathbf{z}}/\mathrm{d}b$ vanish. In order to eliminate these state sensitivities, the adjoint variables should satisfy

$$\dot{\boldsymbol{\mu}} = \left[(\mathbf{M}\dot{\mathbf{z}})_{\mathbf{y}} - \mathbf{Q}_{\mathbf{y}}^A \right]^T (\boldsymbol{\zeta} + \boldsymbol{\xi}) - p_{\mathbf{y}}^T, \tag{28}$$

$$\mathbf{M}\dot{\boldsymbol{\zeta}} = -\boldsymbol{\mu} - \dot{\mathbf{M}}\boldsymbol{\zeta} - \mathbf{Q}_{\mathbf{z}}^{A\,T} (\boldsymbol{\zeta} + \boldsymbol{\xi}) - p_{\mathbf{z}}^T, \tag{29}$$

with boundary conditions

$$\boldsymbol{\mu}(t_f) = \mathbf{0}, \tag{30}$$

$$\boldsymbol{\zeta}(t_f) = \mathbf{0}, \tag{31}$$

where symmetry of the mass matrix is used. Herein, the 'intermediate' adjoint variables $\boldsymbol{\xi}$ follow from

$$\mathbf{M}\boldsymbol{\xi} = p_{\dot{\mathbf{z}}}^T. \tag{32}$$

The end conditions in Equations (30) and (31) require that the first-order differential equations (28) and (29) are integrated *backward* in time. Finally, Equation (27) becomes

$$\frac{\mathrm{d}\overline{r}}{\mathrm{d}b} = \int_{t_0}^{t_f} \left[p_b - \left(\boldsymbol{\zeta}^T + \boldsymbol{\xi}^T \right) \left(\mathbf{M}_b \dot{\mathbf{z}} - \mathbf{Q}_b^A \right) \right] \mathrm{d}t. \tag{33}$$

The sensitivity equations show that the computational cost of the direct method is proportional to the number of design variables, whereas the computational cost of the adjoint method is directly related to the number of response sensitivities that have to be calculated. The adjoint method is often more efficient if sensitivities are needed just for a small number of responses in combination with a large number of design variables and multiple load cases. This means that either equivalent integrated constraints such as (7) should be used, or critical time point constraints (8) using a constraint screening that includes only very few time points in the optimization. This is in contradiction to the constraint screening proposed in the previous section where multiple time points near local maxima are used to avoid oscillations for approximations that lack correct curvature. Possibly, an approximation concept with adjustable conservativeness may be able to overcome this difficulty.

In some cases the direct method is more suited in combination with the application of intermediate response variables than the adjoint method. Suppose we have an objective function that has integral form (26) with a nonlinear function p of the state variables, such as the comfort criterion (42) of the impact absorber example. If we would like to include the nonlinear behavior of p in the approximation of the objective function, then we have to approximate the states $\mathbf{y}$, $\mathbf{z}$, and $\dot{\mathbf{z}}$, and compute $\mathrm{d}\mathbf{y}/\mathrm{d}b$, $\mathrm{d}\mathbf{z}/\mathrm{d}b$, and $\mathrm{d}\dot{\mathbf{z}}/\mathrm{d}b$ on the time interval $[t_0, t_f]$. The direct method computes exactly these sensitivities. The adjoint method, however, uses integral (26) to directly compute the sensitivities of the objective function. But then we cannot benefit anymore from the intermediate response variables to improve the approximations.

6. Examples

6.1. Impact Absorber

An impact absorber is modeled as a single degree of freedom system with a constant mass m of 1 kg, a linear spring k and a linear damper c (Figure 2). It is a simplified version of the nonlinear impact absorber of Afimawala and Mayne [22]. At time $t = 0$ s, the initial mass position and velocity are $y(0) = 0$ m, and $\dot{y}(0) = 1$ ms^{-1}, respectively. Starting from the equation of motion and the initial conditions, the mass position as a function of time can be solved analytically:

$$y(k,c,t) = \begin{cases} \frac{e^{-t(c/2)}}{\sqrt{k-(c/2)^2}} \sin(t\sqrt{k-(c/2)^2}) & \text{if } 0 \le \frac{(c/2)}{\sqrt{k}} < 1, \\ te^{-t\sqrt{k}} & \text{if } \frac{(c/2)}{\sqrt{k}} = 1, \\ \frac{e^{-t(c/2)}}{2\sqrt{(c/2)^2-k}} \left[e^{t\sqrt{(c/2)^2-k}} - e^{-t\sqrt{(c/2)^2-k}} \right] & \text{if } \frac{(c/2)}{\sqrt{k}} > 1. \end{cases} \tag{34}$$

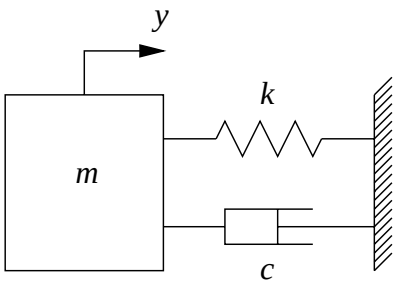


Figure 2. Impact absorber.

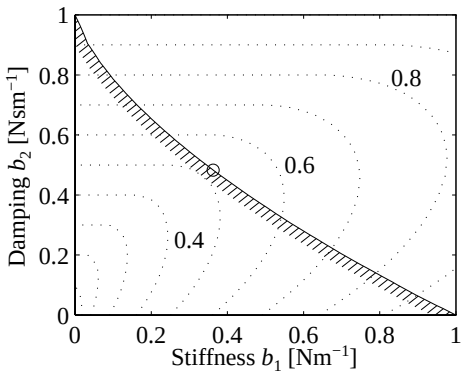


Figure 3. Optimization problem of the impact absorber for the maximum acceleration objective function. The optimum design is marked with ∘.

The optimum design problem is to find the stiffness coefficient $b_1 = k$ and the damping coefficient $b_2 = c$ that will minimize the maximum acceleration

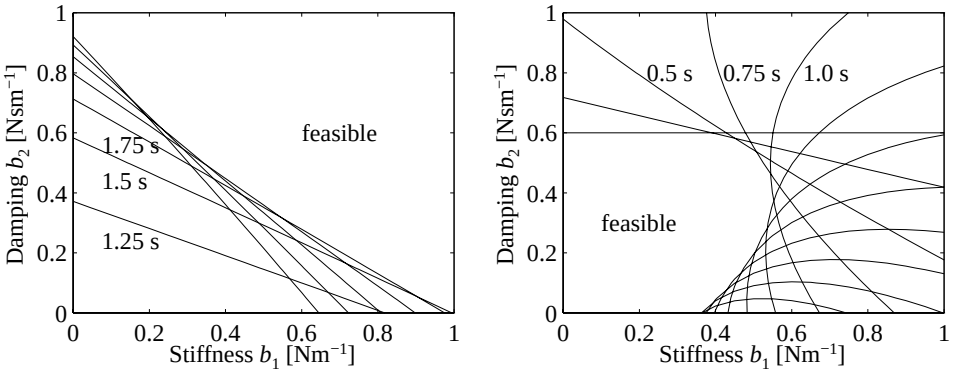
$$F(\mathbf{b}) = \max_{t\in[0,12]} |\ddot{y}(b_1, b_2, t)| \tag{35}$$

subject to the displacement constraint

$$g(\mathbf{b}, t) = |y(b_1, b_2, t)| \leq 1 \quad \forall\, t \in [0, 12], \tag{36}$$

within the design space $0 < b_1 < 1\,\mathrm{Nm}^{-1}$ and $0 < b_2 < 1\,\mathrm{Nsm}^{-1}$. A time period of 12 s includes all important response maxima. In Figure 3 the optimization problem is visualized. The hatched line represents the displacement constraint bound, and the dotted lines represent contour lines of the maximum acceleration. The feasible region is in the upper right part of the design space. A single optimum solution is present, determined by the curvature of the maximum acceleration. For computational convenience the problem is reformulated: minimize the artificial design variable

$$F(\mathbf{b}) = b_3 \tag{37}$$



(a) Displacement (b) Acceleration ($b_3 = 0.6$ ms^{-2}).

Figure 4. Time point constraints of the impact absorber.

subject to the acceleration and displacement constraints

$$g_1(\mathbf{b}, t) = \ddot{y}(b_1, b_2, t) \leq b_3 \quad \forall\, t \in [0, 12], \tag{38}$$

$$g_2(\mathbf{b}, t) = -\ddot{y}(b_1, b_2, t) \leq b_3 \quad \forall\, t \in [0, 12], \tag{39}$$

$$g_3(\mathbf{b}, t) = y(b_1, b_2, t) \leq 1 \quad \forall\, t \in [0, 12], \tag{40}$$

$$g_4(\mathbf{b}, t) = -y(b_1, b_2, t) \leq 1 \quad \forall\, t \in [0, 12], \tag{41}$$

with $0 < b_1 < 1$, $0 < b_2 < 1$, and $b_3 \geq 0$.

Time discretization is used to deal with the time-dependent constraints. For a rather coarse discretization at every 0.25 s, the displacement and acceleration time point constraints of Figure 4 are obtained. The curvature that arises due to the shift of the response maxima is represented by intersecting time point constraints. As a result, the optimum is determined by two adjacent time point acceleration constraints and one time point displacement constraint, or one acceleration and two displacement constraints. Furthermore, notice the non-convex behavior of the time point acceleration constraints. This means that the discretized optimization problem may have multiple adjacent local optima in contrary to the time continuous problem.

The sequential approximate optimization approach is applied without intermediate design variables and responses. Displacements and accelerations are linearly approximated with respect to the design variables. As a result, the optimization is a sequential linear programming process. The discretization contains 101 (equally distributed) time points, and the accuracy is set to 10^{-5}. Starting from the initial design $b_1 = 1.0$ and $b_2 = 0.3$, and initial move limits of 40% of the design variable values, the optimum design is found after five cycles using six multibody analyses and five sensitivity analyses. The optimization history is summarized in Table I. For each design cycle, values of the design variables b_i, objective function F, maximum constraint violation V, and maximum approximation error E_g are tabulated.

Table I. Optimization history of the impact absorber with max-value objective function for two additional time points below and above local maxima.

Cycle	active mvlim	b_1 [Nm^{-1}]	b_2 [Nsm^{-1}]	$b_3 = F$ [ms^{-2}]	E_g [%]	V [%]
0 (y)	0	1.0000	0.3000	0.8500	0.000	$-7.296 \cdot 10^{-1}$
1 (y)	2	0.6000	0.4200	0.6232	$4.703 \cdot 10^{-1}$	$4.043 \cdot 10^{-1}$
2 (y)	0	0.3411	0.4934	0.5110	$1.625 \cdot 10^{-1}$	$8.996 \cdot 10^{-1}$
3 (y)	0	0.3628	0.4823	0.5203	$3.539 \cdot 10^{-3}$	$5.554 \cdot 10^{-4}$
4 (y)	0	0.3628	0.4823	0.5203	$9.055 \cdot 10^{-9}$	$6.510 \cdot 10^{-9}$
5 (y)	0	0.3628	0.4823	0.5203	0.000	$6.510 \cdot 10^{-9}$

All cycle optima are accepted as starting design of the next cycle, indicated by (y) in the first column. Move limits (two) are active only during the first design cycle. Constraints are included at a maximum of 14 time points. Both the maximum error and constraint violation remain small.

The optimization is restarted, but now only one time point constraint is included for each local maximum. As a result, 18 cycles are necessary to converge. Always one of the move limits remains active, and convergence can only be reached by a repeated reduction of the search subregion. An oscillatory behavior of the design variable values occurs.

Intermediate response variables can be employed when the maximum acceleration objective function is replaced by a comfort criterion such as:

$$F(\mathbf{b}) = \frac{1}{t_f - t_0} \int_{t_0}^{t_f} \ddot{y}^2(b_1, b_2, t)\, \mathrm{d}t. \tag{42}$$

Figure 5 shows that the new objective function has a completely different behavior. The optimum design moves towards the upper left corner of the design space. Again, displacements and accelerations are linearly approximated, however, the integral relation between objective function value and acceleration is included in the approximate optimization problem. Starting from $b_1 = 1.0$ and $b_2 = 0.3$, the final optimum design is found after eight cycles using nine multibody analyses and eight sensitivity analyses (see Table II). During the first four cycles always one move limit is active. Afterwards, automatic convergence occurs without any oscillations. When using a direct coupling of the SQP algorithm, 23 analyses and sensitivity analyses are required [15].

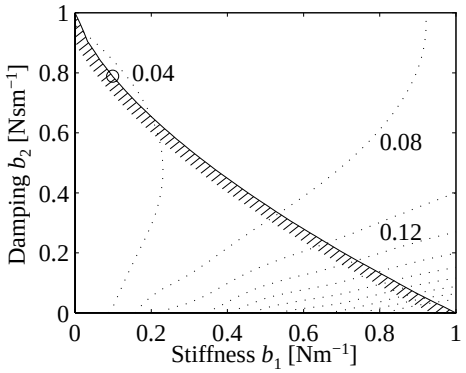


Figure 5. Optimization problem of the impact absorber for a comfort criterion objective function. The optimum design is marked with ○.

Table II. Optimization history of the impact absorber with comfort criterion objective function.

Cycle	b_1 [Nm^{-1}]	b_2 [Nsm^{-1}]	E_f [%]	E_g [%]	V [%]	F [m^{2}s^{-4}]
0 (y)	1.000	0.3000	0.000	0.000	$-1.938 \cdot 10^{1}$	0.1480
1 (y)	0.8657	0.4200	$5.901 \cdot 10^{-1}$	$1.095 \cdot 10^{-1}$	$-2.132 \cdot 10^{1}$	0.1029
2 (y)	0.6596	0.6640	$2.225 \cdot 10^{0}$	$4.781 \cdot 10^{-1}$	$-2.549 \cdot 10^{1}$	0.06949
3 (y)	0.3958	0.7933	$3.909 \cdot 10^{0}$	$2.882 \cdot 10^{-1}$	$-2.277 \cdot 10^{1}$	0.05384
4 (y)	0.1847	0.6117	$1.352 \cdot 10^{1}$	$3.801 \cdot 10^{0}$	$5.663 \cdot 10^{0}$	0.03805
5 (y)	0.1052	0.7750	$2.494 \cdot 10^{-2}$	$5.378 \cdot 10^{-1}$	$3.192 \cdot 10^{-1}$	0.03793
6 (y)	0.09856	0.7882	$7.365 \cdot 10^{-4}$	$2.286 \cdot 10^{-3}$	$1.472 \cdot 10^{-3}$	0.03804
7 (y)	0.09850	0.7883	$1.168 \cdot 10^{-7}$	$1.0 \cdot 10^{-7}$	$-1.000 \cdot 10^{-7}$	0.03804
8 (y)	0.09850	0.7883	0.000	0.000	$-1.000 \cdot 10^{-7}$	0.03804

6.2. SLIDER-CRANK MECHANISM

The mechanism design problem of [23] is used to illustrate that intermediate response variables can be employed for kinematic responses as well. A four-bar slider-crank mechanism has to be designed such that a desired coupler curve is generated. Figure 6 shows the four-bar mechanism with eight design variables b_i and the coupler point P that should generate the desired curve. The prescribed path and timing of the coupler curve is given in Table III. The desired (x_G, y_G) positions of P are tabulated as a function of the crank rotations $\Delta\gamma(k)$ relative to the starting angle b_8.

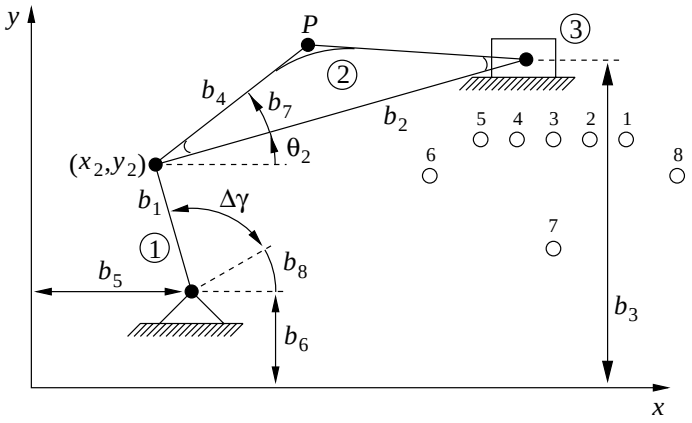


Figure 6. Slider-crank mechanism and coupler curve.

Table III. Path coordinates and prescribed timing.

Point k	1	2	3	4	5	6	7	8
$x_G(k)$	26	23	20	17	14	10	20	30
$y_G(k)$	16	16	16	16	16	13	7	13
$\Delta\gamma(k)$ in degrees	0	22	44	66	88	129	221	314

The optimization problem is to minimize the objective function

$$F(\mathbf{b}) = \sum_{k=1}^{8} \left[\{x_P(\Delta\gamma(k)) - x_G(k)\}^2 + \{y_P(\Delta\gamma(k)) - y_G(k)\}^2 \right] \tag{43}$$

subject to the movability constraints

$$g_1(\mathbf{b}) = -0.85b_2 + b_1 + b_3 - b_6 \leq 0, \tag{44}$$

$$g_2(\mathbf{b}) = -0.85b_2 + b_1 - b_3 + b_6 \leq 0, \tag{45}$$

and design space bounds $1 \leq b_1 \leq 30$, $1 \leq b_2 \leq 30$ and $1 \leq b_4 \leq 30$. The movability constraints ensure that the linkage can operate for the complete range of crank rotations. The position of coupler point P (x_P, y_P) follows from the response variables x_2, y_2 and θ_2 of body 2, and the design variables b_4 and b_7:

$$x_P = x_2 + b_4 \cos(b_7 + \theta_2), \tag{46}$$

$$y_P = y_2 + b_4 \sin(b_7 + \theta_2). \tag{47}$$

Usually, the state or response variable values have to be solved numerically from the governing equations of the multibody system. But for the slider-crank mech-

Table IV. Optimization history of the slider-crank mechanism.

Cycle	active mvlim	max Δb_i [%]	E_f [%]	F
0 (y)	0	$0.000 \cdot 10^{0}$	$0.000 \cdot 10^{0}$	$3.486 \cdot 10^{3}$
1 (n)	5	$4.000 \cdot 10^{1}$	$8.541 \cdot 10^{1}$	$1.303 \cdot 10^{2}$
1 (y)	8	$2.000 \cdot 10^{1}$	$2.870 \cdot 10^{1}$	$6.260 \cdot 10^{2}$
2 (y)	7	$1.500 \cdot 10^{1}$	$2.506 \cdot 10^{0}$	$8.500 \cdot 10^{1}$
3 (y)	4	$2.000 \cdot 10^{1}$	$1.691 \cdot 10^{-1}$	$3.225 \cdot 10^{1}$
4 (y)	4	$2.667 \cdot 10^{1}$	$6.177 \cdot 10^{0}$	$1.455 \cdot 10^{1}$
5 (y)	2	$2.000 \cdot 10^{1}$	$5.953 \cdot 10^{0}$	$5.658 \cdot 10^{0}$
6 (y)	1	$2.667 \cdot 10^{1}$	$2.263 \cdot 10^{0}$	$3.942 \cdot 10^{0}$
7 (y)	1	$3.556 \cdot 10^{1}$	$4.954 \cdot 10^{-1}$	$3.772 \cdot 10^{0}$
8 (y)	0	$1.288 \cdot 10^{1}$	$9.316 \cdot 10^{-3}$	$3.746 \cdot 10^{0}$
9 (y)	0	$1.085 \cdot 10^{0}$	$1.353 \cdot 10^{-4}$	$3.746 \cdot 10^{0}$
10 (y)	0	$3.248 \cdot 10^{-2}$	$4.511 \cdot 10^{-7}$	$3.746 \cdot 10^{0}$

anism, x_2, y_2 and θ_2 can be derived analytically as function of the rotation angle $\Delta\gamma$:

$$x_2 = b_1 \cos(b_8 + \Delta\gamma) + b_5, \tag{48}$$

$$y_2 = b_1 \sin(b_8 + \Delta\gamma) + b_6, \tag{49}$$

$$\theta_2 = \arcsin\left(\frac{b_3 - y_2}{b_2}\right). \tag{50}$$

We assume that responses x_2, y_2 and θ_2 are not explicitly known and treat them as intermediate response variables that are linearly approximated with respect to the design variables. This means that relations (43), (46) and (47) are included in the approximate optimization problem. The accuracy is again 10^{-5}. A sequence of ten approximate nonlinear programming problems yields the final optimum solution, and nearly exactly corresponds with the solution of Gabriele [1]. Gabriele, however, needed 44 iterations of the SQP algorithm, requiring 803 evaluations of the objective function (central difference gradients were used). So, the approximate optimization approach used much less 'expensive' multibody analyses and design sensitivity analyses. Instead, the objective function approximation is quite frequently evaluated. Due to the analytical example the actual savings in computer time are not so much present, but become much more apparent for numerically expensive multibody analyses.

The optimization history is given in Table IV. The first design cycle has to be repeated with smaller move limits, since the approximation error for the initial

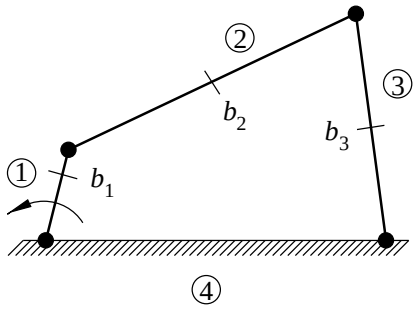


Figure 7. Stress-constrained four-bar mechanism.

move limit factors (40%) is far too large. Afterwards, a good convergence occurs with a steady reduction of the number of active move limits. During the last three cycles no move limits are active anymore and the maximum design variable change $\max \Delta b_i$ shows a superlinear decrease. Constraint violations do not occur, since the movability constraints are explicitly known and included in the approximate optimization problem. If the optimization is started with move limit factors of 30%, just eight cycles will be necessary to find the optimum solution. Then, a much smaller approximation error occurs at the first design cycle.

6.3. Stress-Constrained Four-Bar Mechanism

Figure 7 shows a four-bar mechanism consisting of three solid but flexible links, connected to each other and the ground by revolute joints. The three mobile links have a constant circular cross section, a Young's modulus of $6.895 \cdot 10^{10}$ Pa, and a mass density ρ of 2757 kg·m^{-3}. The lengths of the bars are $l_1 = 0.3048$ m, $l_2 = 0.9144$ m, $l_3 = 0.762$ m, and $l_4 = 0.9144$ m, respectively. The input crank rotates with a constant angular velocity of 10π s^{-1}. Due to the motion, bending stresses occur in the mobile links.

Each link is modeled by six beam elements. The multibody analysis package MECANO [24] is used to compute the bending moments M_i^k in every node k of link i as a function of time. The bending stresses can then be calculated from

$$\sigma_i^k = \frac{4\sqrt{\pi}}{A_i^{3/2}} M_i^k, \tag{51}$$

where A_i is the cross sectional area of body i. A time interval of 0.3 to 0.5 s is considered which exactly covers one period of steady state motion after the transient has died away.

The optimization problem is to find the diameters of the mobile links $b_i = d_i$; $(i = 1, 2, 3)$ that will minimize the total mass

$$F(\mathbf{b}) = \rho(l_1 A_1 + l_2 A_2 + l_3 A_3) \tag{52}$$

subject to the stress constraints

$$\sigma_i^k(\mathbf{b}, t) \leq \sigma_a \quad \forall\, t \in [0.3, 0.5], \quad i = 1, 2, 3, \quad k = 1, \ldots, 7, \tag{53}$$

in the design space $b_i \geq 0$ ($i = 1, 2, 3$) with a maximum bending stress of $\sigma_a = 2.758 \cdot 10^7$ Pa. The time interval of 0.3 to 0.5 s is discretized into 201 time points. For the approximate optimization problem, the cross sectional areas are taken as intermediate design variables:

$$b_i^I = A_i = \frac{1}{4}\pi d_i^2, \quad i = 1, 2, 3. \tag{54}$$

The bending moments M_i^k are used as intermediate response quantities.

MECANO [24] has been coupled with the design optimization tool described in Section 4. However, MECANO cannot calculate the required design sensitivities. Therefore, the derivatives with respect to the design variables are computed by forward finite differencing. The optimization is started from initial design variable values of 356.8 mm [2], using initial move limits of 40% for the design variable values. This initial design has a total mass of 546 kg and is far from the maximum allowed stress. The critical constraint value for the constraint screening is set to 50% (see Section 4).

After 15 cycles, optimum design I of Table V is found. For other optimization runs with slightly different initial conditions we found optimum design II as well in accordance with [25]. Apparently multiple local optima are present. These designs differ from the final design of Sohoni and Haug [2], which is due to the different method of stress analysis – dynamic instead of kinematic – and the discretization of the mobile links. In [26] we also studied the stress-constrained four-bar mechanism, but did not use intermediate design variables and intermediate response quantities. About the same number of cycles was needed for convergence. At first glance, one may therefore conclude that the effect of the intermediate variables and responses is marginal, but a closer look at the convergence behavior reveals why still 15 cycles are necessary.

A quite typical optimization history is found (see Table VI). After five optimization cycles a great reduction of the move limits was necessary to keep the approximation errors at an acceptable level. Along with the decrease of the beam diameters, the links start to vibrate (see [15] for some plots). Additionally, the vibration frequency varies as the diameter changes. According to the beam theory, the frequency of the vibration will decrease if the diameter of a beam becomes smaller. This effect is indeed observed for the bending stresses of the mobile links. As a result, the position and the number of local stress maxima in the time interval [0.3, 0.5] vary. Approximations, however, are built at fixed time points and cannot predict this dynamic effect. For large steps in the design space, the approximation errors may rise, and if the errors become too large, a move limit reduction is necessary.

In order to reduce the contribution of the dynamic behavior, Young's modulus is increased to $6.895 \cdot 10^{11}$ Pa (imaginary experiment). With and without intermediate

Table V. Initial and optimum design of the stress-constrained four-bar mechanism. Optimum designs I and II have been calculated for the original Young's modulus of $6.895 \cdot 10^{10}$ Pa. Optimum III corresponds to an increased Young's modulus of $6.895 \cdot 10^{11}$ Pa.

b and $F(\mathbf{b})$	Initial design	Optimum design			Optimum reported by [2]
		I	II	III	
d_1 [mm]	356.8	38.5	37.3	35.5	36.2
d_2 [mm]	356.8	25.2	23.6	23.0	28.1
d_3 [mm]	356.8	20.0	19.8	18.0	12.2
F [kg]	546.0	2.89	2.67	2.42	2.69

Table VI. Optimization history of the stress-constrained four-bar mechanism for Young's modulus of $6.895 \cdot 10^{10}$ Pa, using cross sectional areas as intermediate design variables and bending moments as intermediate response quantities.

Cycle	active mvlim	b_1 [mm]	b_2 [mm]	b_3 [mm]	E_g [%]	V [%]	F [kg]
0 (y)	0	356.8	356.8	356.8	0.00	$-7.15 \cdot 10^{1}$	$5.46 \cdot 10^{2}$
1 (y)	3	214.1	214.1	214.1	0.00	$-5.28 \cdot 10^{1}$	$1.97 \cdot 10^{2}$
2 (y)	3	128.4	128.4	128.4	0.00	$-2.12 \cdot 10^{1}$	$7.08 \cdot 10^{1}$
3 (y)	2	84.7	77.1	77.1	$1.83 \cdot 10^{0}$	$-6.11 \cdot 10^{-1}$	$2.63 \cdot 10^{1}$
4 (y)	2	52.2	36.0	36.0	$2.57 \cdot 10^{1}$	$-6.65 \cdot 10^{0}$	$6.50 \cdot 10^{0}$
5 (n)	0	42.2	28.8	23.0	$1.21 \cdot 10^{2}$	$-2.77 \cdot 10^{0}$	$3.70 \cdot 10^{0}$
5 (n)	2	44.4	29.0	28.8	$6.07 \cdot 10^{1}$	$-1.39 \cdot 10^{0}$	$4.34 \cdot 10^{0}$
5 (y)	3	48.3	32.4	32.4	$2.63 \cdot 10^{1}$	$-4.19 \cdot 10^{0}$	$5.34 \cdot 10^{0}$
6 (y)	2	45.7	29.9	29.9	$7.51 \cdot 10^{0}$	$-2.68 \cdot 10^{0}$	$4.63 \cdot 10^{0}$
7 (y)	3	43.1	27.0	27.0	$1.69 \cdot 10^{1}$	$-6.35 \cdot 10^{0}$	$3.86 \cdot 10^{0}$
8 (n)	2	40.7	24.3	24.3	$4.16 \cdot 10^{1}$	$-3.26 \cdot 10^{0}$	$3.23 \cdot 10^{0}$
8 (y)	3	41.9	25.6	25.6	$1.86 \cdot 10^{1}$	$-4.30 \cdot 10^{0}$	$3.53 \cdot 10^{0}$
9 (y)	2	40.7	25.2	24.3	$9.99 \cdot 10^{0}$	$-1.06 \cdot 10^{0}$	$3.33 \cdot 10^{0}$
10 (y)	1	39.6	25.1	22.7	$1.20 \cdot 10^{1}$	$2.24 \cdot 10^{-1}$	$3.13 \cdot 10^{0}$
11 (y)	1	39.1	25.2	21.2	$1.40 \cdot 10^{1}$	$2.75 \cdot 10^{-2}$	$3.00 \cdot 10^{0}$
12 (y)	1	38.5	25.4	19.8	$1.70 \cdot 10^{1}$	$4.45 \cdot 10^{0}$	$2.90 \cdot 10^{0}$
13 (y)	0	38.5	25.3	20.1	$2.28 \cdot 10^{0}$	$-1.87 \cdot 10^{-1}$	$2.91 \cdot 10^{0}$
14 (y)	0	38.5	25.2	20.0	$2.35 \cdot 10^{-1}$	$5.05 \cdot 10^{-2}$	$2.89 \cdot 10^{0}$
15 (y)	0	38.5	25.2	20.0	$1.92 \cdot 10^{-3}$	$9.09 \cdot 10^{-4}$	$2.89 \cdot 10^{0}$

variables and responses, the optimization process converges towards optimum III of Table V. With intermediate variables and responses, only eight cycles are necessary with very small approximation errors [15]. Without, 14 cycles result with somewhat larger approximation errors, but still much better compared with the case of the original Young's modulus value. This shows that the intermediate design variables and intermediate response quantities do yield better approximations, but that this benefit can be spoiled by the dynamic behavior.

7. Conclusions and Recommendations

Approximation concepts can effectively be applied to design optimization of multibody systems. They prove to be a valuable part in a multibody optimum design tool. Other important elements are an efficient multibody analysis and sensitivity analysis, a graphical user-interface, a database combined with for example an expert system, and a control program to manage the complete design process. An important future step is to put all parts together and build a completely integrated and interactive design tool. This will put a different complexion on each of the individual elements.

Intermediate design variables and intermediate response quantities can highly improve the approximations, and thus enhance the optimization process. For multibody systems several intermediate variables and responses can be identified. Their beneficial effect has been illustrated by the impact absorber, the slider-crank mechanism, and the four-bar constrained mechanism example. For the last example, however, vibrations induced by the motion of the system deteriorated the approximations. Possibly, new intermediate design variables and intermediate response quantities may be defined that include the change of the vibration frequencies.

The design sensitivity equations can be obtained either by direct differentiation of the equations of motion, or by the adjoint variable method. If approximate optimization with time discretization is used, usually the direct method is most suited. It works effectively in combination with the proposed constraint screening, and allows the utilization of intermediate design variables and intermediate response quantities. The adjoint variable method is preferable if the number of (time point) responses for which sensitivities are needed is small, the number of design variables is large, and multiple load cases occur.

Due to the large possible number of design variables, symbolic generation of sensitivity equations seems to be appropriate. For multibody codes that evaluate the equations of motion numerically from a library of joint elements, implementation of the direct and the adjoint method may become an immense programming task. In that case, an option is to use the semi-analytical method, i.e., to calculate the partial derivatives with respect to the design variables in the sensitivity equations by finite differencing which is commonly applied in structural optimization. Another option is to use automatic differentiation [27, 28], and produce new FORTRAN or C code at the joint element level for the extra partial derivatives in the sensitivity equations.

Acknowledgement

The authors thank the reviewers for their valuable suggestions and remarks.

References

1. Gabriele, G.A., 'Optimization in mechanisms', in *Modern Kinematics: Developments in the Last Forty Years*, A.G. Erdman (ed.), John Wiley & Sons, New York, 1993, Chapter 11, 471–519.
2. Sohoni, V.N. and Haug, E.J., 'A state space technique for optimal design of mechanisms', *ASME Journal of Mechanical Design* **104**, 1982, 792–798.
3. Haug, E.J., Wehage, R.A. and Mani, N.K., 'Design sensitivity analysis of large-scale constrained dynamic mechanical systems', *ASME Journal of Mechanisms, Transmissions, and Automation in Design* **106**, 1984, 156–162.
4. Ashrafiuon, H. and Mani, N.K., 'Analysis and optimal design of spatial mechanical systems', *ASME Journal of Mechanical Design* **112**, 1990, 200–207.
5. Bestle, D., *Analyse und Optimierung von Mehrkörpersystemen*, Springer-Verlag, Berlin, 1994.
6. Schiehlen, W., 'Multibody system dynamics: roots and perspectives', *Multibody System Dynamics* **1**, 1997, 149–188.
7. Haug, E.J., Choi, K.K. and Komkov, V., *Design Sensitivity Analysis of Structural Systems*, Academic Press, London, 1986.
8. Erdman, A.G., 'Computer-aided mechanism design: now and the future', *ASME Special 50th Anniversary Design Issue* **117(B)**, 1995, 93–100.
9. Barthelemy, J.-F.M. and Haftka, R.T., 'Approximation concepts for optimum structural design – A review', *Structural Optimization* **5**, 1993, 129–144.
10. Haftka, R.T. and Gürdal, Z., *Elements of Structural Optimization*, Kluwer Academic Publishers, Dordrecht, 1992.
11. Hsieh, C.C. and Arora, J.S., 'Design sensitivity analysis and optimization of dynamic response', *Computer Methods in Applied Mechanics and Engineering* **43**, 1984, 195–219.
12. Bestle, D., 'Zur Optimierung von Mehrkörpersystemen', Institutsbericht IB-14, Institut B für Mechanik, University of Stuttgart, 1989.
13. Grandhi, R.V., Haftka, R.T. and Watson, L.T., 'Design-oriented identification of critical times in transient response', *AIAA Journal* **24**, 1986, 649–656.
14. Vanderplaats, G.N., 'Thirty years of modern structural optimization', *Advances in Engineering Software* **16**, 1993, 81–88.
15. Etman, L.F.P., 'Optimization of multibody systems using approximation concepts', Ph.D. Thesis, Eindhoven University of Technology, 1997.
16. *NAG FORTRAN Library Manual*, Mark 15, Numerical Algorithms Group, Oxford, 1991.
17. Sepulveda, A.E. and Schmit, L.A., 'Approximation-based global optimization strategy for structural synthesis', *AIAA Journal* **31**, 1993, 180–188.
18. Fleury, C. and Braibant, V., 'Structural optimization: A new dual method using mixed variables', *International Journal for Numerical Methods in Engineering* **23**, 1986, 409–428.
19. Svanberg, K., 'The method of moving asymptotes – A new method for structural optimization', *International Journal for Numerical Methods in Engineering* **24**, 1987, 359–373.
20. Svanberg, K., 'A globally convergent version of MMA without linesearch', in *WCMSO-1 First World Congress of Structural and Multidisciplinary Optimization*, N. Olhoff and G.I.N. Rozvany (eds), Pergamon, Oxford, 1996, 9–16.
21. Bestle, D. and Eberhard, P., 'Analyzing and optimizing multibody systems', *Mechanics of Structures and Machines* **20**, 1992, 67–92.
22. Afimawala, K.A. and Mayne, R.W., 'Optimum design of an impact absorber', *ASME Journal of Engineering for Industry* **96**, 1974, 124–130.

23. Paradis, M.J. and Willmert, K.D., 'Optimal mechanism design using the Gauss constrained method', *ASME Journal of Mechanisms, Transmissions, and Automation in Design* **105**, 1983, 187–196.
24. *SAMCEF MECANO Manual, Version 5.1*, Samtech, Liège, 1994.
25. Etman, L.F.P., van Campen, D.H. and Schoofs. A.J.G., 'Optimization of multibody systems using approximation concepts', in *IUTAM Symposium on Optimization of Mechanical Systems*, D. Bestle and W. Schiehlen (eds), Kluwer Academic Publishers, Dordrecht, 1996, 81–88.
26. Etman, L.F.P., Thijssen, E.J.R.W., Schoofs, A.J.G. and van Campen, D.H., 'Optimization of multibody systems using sequential linear programming', in *ASME Advances in Design Automation*, Vol. DE 69-2, B.J. Gilmore, D.A. Hoeltzel, D. Dutta and H.A. Eschenauer (eds), American Society of Mechanical Engineers, New York, 1994, 525–530.
27. Bischof, C.H., 'On the automatic differentiation of computer programs and an application to multibody systems', in *IUTAM Symposium on Optimization of Mechanical Systems*, D. Bestle and W. Schiehlen (eds), Kluwer Academic Publishers, Dordrecht, 1996, 41–48.
28. Eberhard, P., *Zur Mehrkriterienoptimierung von Mehrkörpersystemen*, VDI Fortschritt-Berichte 11 (227), VDI Verlag, Düsseldorf, 1996.