Tolerancing kinematic couplings

Mathieu Barraja a, b, R. Ryan Vallance a, * a Mechanical and Aerospace Engineering, The George Washington University, 738 Phillips Hall, 801 22nd St., N.W. Washington, DC 20052, USA
b Department of Mechanical Engineering, University of Kentucky, 015 RGAN Building, Lexington, KY 40506, USA

Received 12 August 2002; received in revised form 9 February 2004; accepted 12 May 2004
Available online 20 July 2004

Abstract

This paper presents a method for allocating tolerances to dimensions in kinematic couplings, which are exact constraint devices used to locate one object with respect to another. The objective is to reduce the manufacturing cost without exceeding limits on the variation of the coupled position and orientation. The allocation procedure uses parametric models of the contacting surfaces and a solution for the resting position of the coupled body. A multivariate error analysis provides a relation between variation in manufactured dimensions and variation in the position and orientation of the coupled body. Optimal tolerances are then determined using a non-linear constrained optimization algorithm that minimizes the manufacturing cost while satisfying constraints on the variation of the coupled position and orientation. The method provides a useful tool when designing mass-produced kinematic couplings intended for applications where coupled bodies are exchanged.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Kinematic coupling, Tolerance allocation, Manufacturing variation, Manufacturing cost

1. Introduction

Kinematic couplings, illustrated in Fig. 1, are widely used for positioning one rigid body with respect to another. Contact between a ball body and a groove body occurs at six points, which is the minimum necessary for static equilibrium. Hence, kinematic couplings exactly constrain [1] all six degrees of freedom without over-constraint and are therefore extremely repeatable techniques for positioning two bodies [2,3]. However, the relative position and orientation of the two bodies are not necessarily accurate. Accuracy must be attained with either mechanical adjustments or tight production tolerances, both of which increase the manufacturing cost of the kinematic coupling.

As kinematic couplings increasingly find applications in manufacturing, fixturing, and material handling, it is necessary to consider the effect of inaccurate kinematic couplings. For instance, Vallance et al. [4] described the use of kinematic couplings for positioning pallets in flexible assembly systems. In this application, kinematically coupled pallets are routinely exchanged at multiple machine stations, and hence manufacturing errors in each pallet and station contribute to system-wide manufacturing variation.

This paper presents a method for allocating tolerances to the dimensions of kinematic couplings so that variation in the position and orientation of kinematically coupled bodies is less than a set of design constraints. The geometry of the contacting surfaces is modeled using parametric functions of dimensions that include manufacturing errors. The variation in the kinematic couplings' position and orientation errors are expressed as a function of the tolerances using a multivariate error analysis [5]. The tolerances of the coupled bodies are related to manufacturing costs via cost–tolerance relations for common processes (milling, drilling, grinding, etc.) published by Chase [6]. Finally, a constrained nonlinear optimization problem returns dimensional tolerances for the kinematic coupling that minimize manufacturing costs but satisfy constraints on variation in position and orientation.

2. Background and prior work

Kinematic couplings have been used in precision instruments for many years [7,8], and their utility in precision machines is widely recognized [9]. In traditional applications, often a single ball body and a single groove body are coupled together, and so the principal functions of the kinematic coupling are

1. to minimize variation in the position and orientation of the ball body after removing and replacing the ball body, and
2. to minimize elastic deformation induced in the ball body due to excessive constraints.

The success of similar anti-distortion mountings and kinematic couplings with regard to these two functions was studied and demonstrated by designers of precision instruments and machines [3,10,11]. As a result, they found increasing application within precision manufacturing equipment and processes [4,12,13]. For some of these applications, multiple ball bodies are coupled with a single or several groove bodies. This introduces system-level variation due to inaccurate production of the mating surfaces within the kinematic coupling, which was described and analyzed by Vallance [5]. To increase the accuracy of each coupling and thereby reduce the system-wide variation, the dimensional variation within the set of ball and groove bodies must be specified and controlled.

Limits on the dimensional variation within the ball and groove bodies can be specified on drawings using standard techniques for dimensional and geometric tolerances [14]. Early tolerancing research resulted in approaches for tolerance analysis that predict the effect of multiple tolerances on the dimensions and geometry of mechanical components [15,16]. The most common approaches use worst-case analyses [17], root-sum-square (RSS) analyses [18], statistical techniques [19], or Monte Carlo simulation [20]. More recent research extended tolerance analysis techniques to assemblies of components [21,22], and some of these techniques are available in tolerance analysis software and may even be integrated with CAD software [23].

Software for tolerance allocation [24], which is the problem of assigning values to the tolerances, is less available. Therefore, tolerance allocation is less common, but it has been demonstrated for particular mechanical systems [25]. Tolerance allocation often uses optimization techniques [26] that minimize cost [27,28] subject to constraints on variation using cost-tolerance relations [24,27,29].

This paper contributes a formulation and solution to tolerance allocation for kinematic couplings, which compliments other analytical tools that assist during design [30,31,32]. The technique for assigning tolerances is statistical, and it uses multivariate error analysis [33] and nonlinear constrained optimization [34] to minimize cost. The technique has been implemented and verified using a set of scripts that execute within Matlab. An additional set of Matlab scripts verifies the results of the allocation using random Monte Carlo simulations.

3. Tolerance allocation

Rigorously allocating tolerances to the dimensions of kinematic couplings, requires an algorithm that incorporates the four aspects described below.

3.1. Describe the geometry and dimensional variation in a mathematical form

Both bodies of the kinematic coupling should be represented parametrically, with respect to their nominal dimensions and their dimensional errors. The contact points between the two bodies are of primary interest for defining the assembly variations of the kinematic coupling, so the parametric representation should concentrate on contacting surfaces in terms of the dimension schemes for modeling the ball and groove bodies.

3.2. Combine dimensional variation in the ball body and groove body to estimate variation in the resting position and orientation of the ball body

The limits to dimensional variation in the ball and groove bodies are defined by tolerances. When a ball body and groove body with particular dimensional errors are assembled together, the ball body is positioned and oriented with errors in its resting position (x_r, y_r, z_r, γ_r, β_r, α_r). A relation between dimensional variation and variation in the resting position and orientation is provided by multivariate error analysis. This approach requires a robust method for determining the resting location of the ball body.

3.3. Relate assembly variation to the performance requirements of the kinematic coupling

The acceptable errors in the resting position and location are defined by the assembly tolerances specified by the designer. If the designer uses error budgeting techniques [35], then the limits on position and orientation errors associated with the kinematic coupling are known. However, these limits are usually specified at operating points, where manufacturing operations are performed, rather than at a reference coordinate system. The performance of the kinematic coupling should therefore be assessed using variation in the position and orientation of operating points.

---

Matlab for Windows is software from The MathWorks Inc., 24 Prime Park Way, Natick, MA.
3.4. Relate dimensional tolerances to manufacturing costs

The objective of the tolerance allocation is to minimize the manufacturing cost of the kinematic coupling, while satisfying tolerances on the assembly errors. It is then necessary to establish cost-tolerance functions relevant to the manufacturing operations used to produce the ball and groove bodies.

These four aspects are incorporated into the algorithm illustrated in Fig. 2. The multivariate error analysis is an iterative process in which one dimension is perturbed at a time: It returns the variation of the resting location and the manufacturing cost. This process is nested in an optimization loop.

4. Parametric representation of contacting surfaces

We require an analytical representation of the contacting surfaces within a kinematic coupling containing manufacturing errors. A common style of kinematic couplings uses three balls resting in three vee-grooves, as illustrated in Fig. 1 [30], and so we use the parametric equations for a sphere and flat surface. Our current limitation to spheres and flats can be generalized in future work with toroids having two radii; this improvement would accommodate other contacting surfaces like canoe balls and gothic-arch grooves. We distinguish six spherical surfaces since the effective diameter of the ball near the contact point may be slightly different due to out-of-roundness in the ball. The arrangement of the six spherical and flat surfaces is illustrated in Fig. 3. For computational convenience, coordinate systems are attached to each surface. The 12 surfaces are described in reference coordinate systems located at the coupling centroid of the ball body (BC) and groove body (GC).

Eq. (1) describes all points \( [x_B, y_B, z_B] \) that lie within a spherical surface with diameter, \( D_B \), and center located by the position vector, \( BC \vec{P}_B = [BC_{Px}, BC_{Py}, BC_{Pz}] \).

\[
(x_B - BC_{Px})^2 + (y_B - BC_{Py})^2 + (z_B - BC_{Pz})^2 - \frac{1}{4}D_B^2 = 0 \quad \text{for } i = 1, \ldots, 6
\] (1)

Eq. (2) describes all points \( [x_F, y_F, z_F] \) that lie within the flat plane in one of the coupling’s vee-grooves. The
vector that locates a point in the vee-groove, \( \text{GC} \), with the \( \text{GC} \rightarrow \text{P}\). The prescript, \( \text{GC} \), indicates that the vectors are measured in a coordinate system located at the grooves’ coupling centroid. The components of the position and normal vectors used in Eqs. (1) and (2) depend upon the manufactured dimensions of the kinematic coupling. Two sets of dimensions, \((d_{B1}, d_{B2}, \ldots, d_{Bn})\) and \((d_{F1}, d_{F2}, \ldots, d_{Fn})\), define the geometry of the ball body and groove body, respectively. The dimensions are measured with respect to two sets of metrology datum frames that define a coordinate system in the ball body denoted with a prescript, \( \text{BD} \), and a coordinate system in the groove body denoted with a prescript, \( \text{GD} \). The form of these relations depends upon the dimension scheme specified by the designer, but they are expressed generally as shown in Eqs. (3)–(5).

\[
\begin{align*}
\text{BC} \text{F}_{i} &= f_{B}(d_{B1}, d_{B2}, \ldots, d_{Bn}) \\
\text{GD} \text{F}_{i} &= f_{F}(d_{F1}, d_{F2}, \ldots, d_{Fn}) \\
\text{GC} \text{F}_{i} &= f_{C}(d_{C1}, d_{C2}, \ldots, d_{Cn})
\end{align*}
\]

The position vectors that locate the spherical and flat surfaces are transformed from the coordinate systems determined by the manufacturing datums \( \text{BD} \) and \( \text{GD} \) to the centroidal coordinate systems \( \text{BC} \) and \( \text{GC} \) using homogeneous transformation matrices \( \text{HTM}_{\text{BD}}^{\text{BC}} \) and \( \text{HTM}_{\text{GD}}^{\text{GC}} \), as shown in Eqs. (6) and (7).

\[
\begin{align*}
\text{BC} \text{F}_{i} &= \text{BC} \text{F}_{i} \text{HTM}_{\text{BD}}^{\text{BC}} \text{BD}_{i} \\
\text{GC} \text{F}_{i} &= \text{GC} \text{F}_{i} \text{HTM}_{\text{GD}}^{\text{GC}} \text{GD}_{i}
\end{align*}
\]

The HTMs \( \text{HTM}_{\text{BD}}^{\text{BC}} \) and \( \text{HTM}_{\text{GD}}^{\text{GC}} \) are determined using a triangle defined by the centers of the balls. The origin of the centroidal coordinate system is located at the intersection of the triangle’s bisectors [36]. Its \( x \)-axis points towards the ball that contains contacting surfaces 5 and 6, and the three apices lie in the \( xy \)-plane. An algorithm for determining \( \text{HTM}_{\text{BD}}^{\text{BC}} \) and \( \text{HTM}_{\text{GD}}^{\text{GC}} \) is presented in Appendix A.

5. Resting position and orientation

When rigid ball and groove bodies are kinematically coupled, the ball body rests in a location that minimizes energy. If friction at the contact points and elastic deformation is neglected (these are only weakly dependent on variability of dimensions), the resting location is determined solely from the manufactured geometry of the bodies. The solution described here uses the geometric model presented in the previous section to calculate the relative position and orientation between kinematically coupled bodies that contain manufacturing errors. By avoiding assumptions such as a linear relation between manufacturing errors and resting position, the method remains valid for even large manufacturing errors.

Specification of the resting position and orientation requires that three translations \( x_{r}, y_{r}, \) and \( z_{r} \), and three rotations \( \alpha_{r}, \beta_{r}, \) and \( \gamma_{r} \), define the unknowns \( (x_{r}, y_{r}, z_{r}, \alpha_{r}, \beta_{r}, \gamma_{r}) \) and hence \( \text{HTM}_{\text{BC}}^{\text{BD}} \), but this cannot be done without also determining the position vectors that locate the six contact points, \( \text{GC} \rightarrow \text{P}_{i} \).

As illustrated in Fig. 4, the solution employs a system of 24 equations and unknowns that are solved iteratively using a nonlinear numerical technique. The inputs to the solver include the diameters of the spherical contacting surfaces, \( DB_{i} \), the position vectors that locate the balls in the \( \text{BC} \) coordinate, \( \text{BC} \rightarrow \text{P}_{i} \), the position vectors that locate the flat surfaces in the \( \text{GC} \) coordinate, \( \text{GC} \rightarrow \text{P}_{i} \), and the normal vectors at the flat surfaces, \( \text{GC} \rightarrow \text{F}_{i} \). The outputs of the algorithm include the translations and rotations of the resting position, \( x_{r}, \)
expressed mathematically in Eq. (9), and may be written six times for each contact point. This produces 18 independent components.

The system of 24 equations is obtained in two sets. The first set of six equations is obtained by requiring that the contact points lie in the plane defined by the flat surfaces. As shown in Eq. (8), this is accomplished by substituting the coordinates of the contact points into Eq. (2), which is done for each of the six contact points.

The second set of 18 equations is obtained from six equations for one ball and flat surface. One path in the loop originates at the GC coordinate system, but it proceeds to the unknown position of the contact points in the GC coordinate system, GC

The vector loop is illustrated in Fig. 5 and the positions vectors that locate the six contact points in the GC coordinate system, GC

The second path in the loop also originates at the GC coordinate system, and the rotations are small, then the matrix form could be derived approximations as described by Hale [37].

Since expressions for the six degrees of freedom are not actually known, the elements of [J] are estimated numerically. This is done by perturbing the value of each dimension from its nominal value, calculating the resting location that gives the six errors, and then evaluating a column in [J].

Assuming the dimensions are continuously distributed random variables and that a tolerance is equivalent to a 3σ range, the error analysis can be treated statistically. A covariance matrix of the form

\[
C_{T} = \begin{bmatrix}
\cos(u_i \cos(f_i)) & \cos(u_i \sin(f_i)) \sin(y_i) - \sin(u_i \cos(y_i)) \\
\sin(u_i \cos(f_i)) & \sin(u_i \sin(f_i)) \sin(y_i) - \cos(u_i \cos(y_i)) \\
-\sin(f_i) & \cos(f_i) \sin(y_i) \\
0 & 0
\end{bmatrix}
\]

6. Multivariate error analysis of variation in resting location

Tolerance allocation requires a relation between dimensional variation and system-wide variability in the resting position and orientation. This can be accomplished using a Monte Carlo simulation, but multivariate error analysis provides a more computationally efficient approach [5]. After allocating tolerances, a Monte Carlo simulation is an effective means for verifying the results.

Multivariate error analyses use linear approximations derived from Taylor series expansion. For instance, there exists a function, \(x = x(d_1, d_2, \ldots, d_{n+1})\), that gives the resting position in the x-direction in terms of the dimensions of the kinematic coupling. As shown in Eq. (11), an estimate of the deviation in the x-coordinate, \(x_{v}\), is expressed using a Taylor series expansion to X that includes only first-order terms consisting of partial derivatives and differential errors in the dimensions, \(\Delta d_{j}\).

\[
\delta x_{v} \approx \frac{\partial x}{\partial d_{1}} \Delta d_{1} + \frac{\partial x}{\partial d_{2}} \Delta d_{2} + \cdots + \frac{\partial x}{\partial d_{n+1}} \Delta d_{n+1}
\]

(11)

Similar expressions are written for the deviations in the remaining degrees of freedom, \(\delta y_{v}\), \(\delta z_{v}\), \(\delta x_{r}\), \(\delta y_{r}\), and \(\delta z_{r}\). All six approximations are expressed in matrix form by the transformation shown in Eq. (12). The \(6 \times (n + 1)\) matrix of partial derivatives is referred to as the Jacobian matrix, [J].

\[
\begin{bmatrix}
\frac{\partial x}{\partial d_{1}} & \frac{\partial y}{\partial d_{1}} & \frac{\partial z}{\partial d_{1}} \\
\frac{\partial x}{\partial d_{2}} & \frac{\partial y}{\partial d_{2}} & \frac{\partial z}{\partial d_{2}} \\
\vdots & \vdots & \vdots \\
\frac{\partial x}{\partial d_{n+1}} & \frac{\partial y}{\partial d_{n+1}} & \frac{\partial z}{\partial d_{n+1}}
\end{bmatrix}
\]

(12)

After the iterative solver returns values for the unknown variables, the HTM between the GC and BC coordinate systems is computed as shown in Eq. (10). If the tolerances are tight and the rotations are small, then the matrix form could be simplified using small angles approximations or second order approximations as described by Hale [37].

\[
C_{T} = \begin{bmatrix}
\cos(u_i \cos(f_i)) & \cos(u_i \sin(f_i)) \sin(y_i) - \sin(u_i \cos(y_i)) \\
\sin(u_i \cos(f_i)) & \sin(u_i \sin(f_i)) \sin(y_i) - \cos(u_i \cos(y_i)) \\
-\sin(f_i) & \cos(f_i) \sin(y_i) \\
0 & 0
\end{bmatrix}
\]

(10)
matrix organizes variances along its diagonal and covariances in the off-diagonal terms. The diagonal elements are therefore squares of the standard deviations of the corresponding random dimensions. The covariance matrix, \([C_D]\), of the kinematic coupling’s dimensions is given in Eq. (13). If the dimensions are independent and therefore uncorrelated, the off-diagonal covariance terms will equal zero. This is a common assumption during tolerance allocation.

\[
[C_D] = \begin{bmatrix}
\sigma^2_{d_1} & \text{cov}(d_1, d_2) & \cdots & \text{cov}(d_1, d_{n+1}) \\
\text{cov}(d_2, d_1) & \sigma^2_{d_2} & \cdots & \text{cov}(d_2, d_{n+1}) \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}(d_{n+1}, d_1) & \text{cov}(d_{n+1}, d_2) & \cdots & \sigma^2_{d_{n+1}}
\end{bmatrix}
\]  

(13)

A similar covariance matrix, \([C_E]\), for the variation in the resting location, is defined in Eq. (14).

\[
[C_E] = \begin{bmatrix}
\sigma^2_{x_e} & \text{cov}(x_e, y_e) & \cdots & \text{cov}(x_e, z_e) \\
\text{cov}(y_e, x_e) & \sigma^2_{y_e} & \cdots & \text{cov}(y_e, z_e) \\
\vdots & \vdots & \ddots & \vdots \\
\text{cov}(z_e, x_e) & \text{cov}(z_e, y_e) & \cdots & \sigma^2_{z_e}
\end{bmatrix}
\]  

(14)

The covariance matrix of the resting location errors, \([C_E]\), is related to the covariance matrix of the dimensions \([C_D]\) by Eq. (15) (33).

\[
[C_E] = [J][C_D][J]^T
\]

(15)

By extracting the diagonal elements of the matrix, \([C_E]\), the multivariate error analysis returns the variation in the resting position and orientation in terms of the tolerances on dimensions.

7. Variation at operating points

The previous section presented a method for estimating the variation in the position and orientation of a coordinate system located at the coupling centroid in the ball body. Although this is useful, the utility of the tolerance allocation is greatly improved if it considers the variation at additional points in the ball body. For instance, in kinematic couplings intended for positioning pallets in flexible assembly operations, the assembly operations such as insertion and joining are performed to a product held within a fixture attached to the ball body. Hence, the designer’s specifications on variation, as determined with an error budget, are preferably specified at operating points.

Fig. 6 illustrates the definition of a single operating point. A coordinate system, denoted with the prescript, \(OP\), is defined at the 4th operation point. An HTM, \(BD_{OP}\), locates the operating point with respect to the manufacturing datum frame in the ball body, BD. A set of \(p\) operating points is similarly defined by a set of \(p\) HTMs.

After determining \(GC_T\) between the ball body’s datums and coupling centroid using the algorithm in the Appendix A, the position and orientation of the operating points can be calculated in the coordinate system at the groove body’s coupling centroid using the transformations shown in Eq. (16). Since \(GC_T\) contains the resting position errors resulting from coupling the ball and groove bodies, the transformation \(GC_{OP}\) reveals the effect of an inaccurate coupling on the position and orientation at the operating point. Larger position errors usually result from amplifying small rotations by the distance separating the operating point from the coupling centroid (Abbe offset).

\[
GC_{OP} = GC_T \times BD_{OP} \times OP
\]  

(16)

The multivariate error analysis is expanded to include the operating points. This is accomplished by expanding the vector of errors and the Jacobian matrix as shown in Eq. (17) so that they incorporate error terms associated with the set of \(p\) operating points. If all six degrees of freedom at each operating point are included, then the dimensions of the error vector become \((6 + 6 \times p) \times 1\), and the dimensions of the Jacobian matrix become \((6 + 6 \times p) \times (m + n)\). However, most manufacturing operations have sensitive and insensitive directions, so considering only the sensitive directions simplifies the problem and requires only a subset of the degrees of freedom at each operating point. Evaluation of the new terms in the Jacobian matrix is still determined by perturbing each dimension in the ball and groove body, calculating the resting position and orientation, and subsequently extracting changes in the values within \(GC_{OP}\).
With the changes shown in Eq. (17), the covariance matrix $[C_{E}]$ calculated with Eq. (15) takes the alternative form shown in Eq. (18). This form includes additional terms for the variances and covariances associated with the position and orientation at the operating points.

$$
[C_{E}] =
\begin{bmatrix}
\sigma_{x}^{2} & \text{cov}(x_{r}, y_{r}) & \cdots & \text{cov}(x_{r}, x_{\text{OP} k}) & \cdots & \text{cov}(x_{r}, \alpha_{\text{OP} p}) \\
\text{cov}(y_{r}, x_{r}) & \sigma_{y}^{2} & \cdots & \text{cov}(y_{r}, x_{\text{OP} k}) & \cdots & \text{cov}(y_{r}, \alpha_{\text{OP} p}) \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\text{cov}(x_{\text{OP} k}, x_{r}) & \text{cov}(x_{\text{OP} k}, y_{r}) & \cdots & \sigma_{x_{\text{OP} k}}^{2} & \cdots & \text{cov}(x_{\text{OP} k}, \alpha_{\text{OP} p}) \\
\text{cov}(\alpha_{\text{OP} p}, x_{r}) & \text{cov}(\alpha_{\text{OP} p}, y_{r}) & \cdots & \text{cov}(\alpha_{\text{OP} p}, x_{\text{OP} k}) & \cdots & \sigma_{\alpha_{\text{OP} p}}^{2}
\end{bmatrix}
$$

(18)

### 8. Relative manufacturing cost and tolerances

The cost of manufactured ball and groove bodies depends upon the selected manufacturing process (assumed capable of producing necessary quantities) and dimensional tolerances. Once a manufacturing process is selected, the cost depends upon both the dimension’s nominal value and its tolerance. The manufacturing cost generally increases if the tolerance is tightened, and it is more expensive to hold a given tolerance on larger nominal dimensions. Several relationships were proposed to relate cost and manufacturing tolerance [24,25,27,29], we use the method recommended by Chase [6] that expresses tolerances for a given process with reciprocal power functions. Eq. (19) expresses the tolerance for the $j$th dimension, $t_{j}$, as a function of relative cost, $C_{j}$, range, $R_{j}$, and three constants $a_{j}$, $b_{j}$, and $c_{j}$. The values of the three constants depend upon the range and the manufacturing process; production quantity becomes relevant only when considering alternative processes. Eq. (19) would require adjustments to relate tolerance with absolute cost, and it is practically difficult to evaluate such constants for a general solution. The primary reason for using Chase’s relation was his extensive reported data [6]. Given sufficient data, alternative relations could also be used in the present work.

$$
t_{j} = c_{j} \times \left( \frac{R_{j}}{C_{j}} \right)^{b_{j}}
$$

(19)

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Process</th>
<th>$a_{j}$</th>
<th>$b_{j}$</th>
<th>$c_{j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thickness of the plate</td>
<td>Milling</td>
<td>0.4431</td>
<td>2.348</td>
<td>0.0355</td>
</tr>
<tr>
<td>Length of a leg</td>
<td>Grinding</td>
<td>0.4323</td>
<td>1.385</td>
<td>0.0217</td>
</tr>
<tr>
<td>Diameter of a ball</td>
<td>Lapping</td>
<td>0.3862</td>
<td>1.052</td>
<td>0.0130</td>
</tr>
<tr>
<td>Location of a hole</td>
<td>Milling</td>
<td>0.4431</td>
<td>2.257</td>
<td>0.0225</td>
</tr>
<tr>
<td>Height of a vee-groove</td>
<td>Grinding</td>
<td>0.4323</td>
<td>1.421</td>
<td>0.0228</td>
</tr>
</tbody>
</table>

Chase provides a set of cost–tolerance curves for some metal removal processes [6]. By extrapolating these curves, we determine values for the coefficients $a_{j}$, $b_{j}$ and $c_{j}$ for each dimension in the kinematic coupling. Table 1 presents the coefficients calculated for an exemplary kinematic coupling configuration.

Using the values in Table 1, Eq. (19) defines the manufacturing cost for every dimension in the exemplary kinematic coupling. Plots of the relations in Fig. 7 illustrate the effect of tightening tolerances. The portion of the total manufacturing cost that is attributable to tolerancing is then the sum of the costs for all $l$ dimensions in the ball and groove bodies, as shown in Eq. (20).

$$
C_{\text{total}} = \sum_{j=1}^{l} \left( \frac{t_{j} R_{j}}{C_{j}} \right)^{1/b_{j}}
$$

(20)

### 9. Tolerance allocation by optimization

Optimal tolerances for the dimensions are determined using nonlinear constrained optimization. The problem is formulated as shown in Eq. (21), where the total cost from Eq. (20) is used as the objective function that is minimized. Constraints are formulated by specifying that the standard deviation of the translation and rotation errors must be positive
yet below critical values. Additional bounds can be specified to prevent the optimization from driving the assigned tolerances to unreasonably high or low values.

\[
\min \left( \sum_{j=1}^{N} \left( \frac{C_j R_j}{T_j} \right) \right)^{1/\beta_j} \] such that

\[ 0 \leq \sigma_{\alpha_j} \leq \sigma_{\max_{\alpha_j}}, 0 \leq \sigma_{\beta_j} \leq \sigma_{\max_{\beta_j}}, 0 \leq \sigma_{\gamma_j} \leq \sigma_{\max_{\gamma_j}}, \quad 0 \leq \sigma_{\gamma_j} \leq \sigma_{\max_{\gamma_j}} \] (21)

The allocation method was used to allocate tolerances to an exemplary kinematic coupling. The parametric surface representation was based on 25 dimensions in the ball body \((m = 25)\) and 18 dimensions \((n = 18)\) in the groove body. The dimension schemes are completely described by Barraja [38]. The cost-tolerance coefficients listed in Table 1 and the constraints listed in Table 2 were used during the optimization. The tolerances that result from the optimization procedure are listed in Table 3. With these tolerances, the constraint on positioning variation in the \(y\) direction is most difficult to satisfy; this is indicated by the calculated \(\sigma_{yr}\) equaling the constraint value \(\sigma_{\max_{yr}}\). All other standard deviations will be less than the constraint values as summarized in Table 2.

### Table 2

<table>
<thead>
<tr>
<th>Variation</th>
<th>Constraint</th>
<th>Calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sigma_{\max_{\alpha_j}})</td>
<td>6.67 (\mu)m</td>
<td>5.23 (\mu)m</td>
</tr>
<tr>
<td>(\sigma_{\max_{\beta_j}})</td>
<td>6.67 (\mu)m</td>
<td>4.87 (\mu)m</td>
</tr>
<tr>
<td>(\sigma_{\max_{\gamma_j}})</td>
<td>(1.164 \times 10^{-3}) rad</td>
<td>(0.115 \times 10^{-3}) rad</td>
</tr>
<tr>
<td>(\sigma_{\max_{\alpha_j}})</td>
<td>(1.164 \times 10^{-3}) rad</td>
<td>(0.068 \times 10^{-3}) rad</td>
</tr>
<tr>
<td>(\sigma_{\max_{\alpha_j}})</td>
<td>(1.164 \times 10^{-3}) rad</td>
<td>(0.078 \times 10^{-3}) rad</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Nominal dimension</th>
<th>Assigned tolerance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ball Pallet</td>
<td>Thickness of the plate</td>
<td>6.35 mm</td>
</tr>
<tr>
<td></td>
<td>Length of a leg ((x \times 3))</td>
<td>19.05 mm</td>
</tr>
<tr>
<td></td>
<td>Roundness of the plate ((x \times 6))</td>
<td>0 mm</td>
</tr>
<tr>
<td></td>
<td>(x) coordinate of a leg axis at the top of the plate</td>
<td>Log 1</td>
</tr>
<tr>
<td></td>
<td>(y) coordinate of a leg axis at the top of the plate</td>
<td>Log 1</td>
</tr>
<tr>
<td></td>
<td>(x) coordinate of a leg axis at the bottom of the plate</td>
<td>Log 1</td>
</tr>
<tr>
<td></td>
<td>(y) coordinate of a leg axis at the bottom of the plate</td>
<td>Log 1</td>
</tr>
<tr>
<td>Groove Body</td>
<td>Height of the vertices for the groove body ((x \times 3))</td>
<td>2.54 mm</td>
</tr>
<tr>
<td></td>
<td>Orientation angle of a groove</td>
<td>Vee 1</td>
</tr>
<tr>
<td></td>
<td>Half-angle of aperture of a groove ((x \times 6))</td>
<td>(\pi/4) rad</td>
</tr>
<tr>
<td></td>
<td>(x) coordinate of a groove</td>
<td>Vee 1</td>
</tr>
<tr>
<td></td>
<td>(y) coordinate of a groove</td>
<td>Vee 1</td>
</tr>
</tbody>
</table>

Total number of dimensions: 43.
Ten thousand cases were generated for the simulation using the nominal dimensions and the assigned tolerances listed in Table 3. The positioning error for each case was computed, and the population’s mean and standard deviation were estimated after completing all 10,000 cases. The means were all very near zero (no mean shift), and the confidence intervals on the standard deviation (95% confidence) are listed in Table 4. The standard deviations from the optimization all fall within the minimum and maximum confidence intervals from the Monte Carlo simulations.

11. Conclusions

Kinematic couplings are known as an economical method for precisely locating one body with respect to another, but the absolute position and orientation between the coupled bodies depends upon manufacturing tolerances. In systems that exchange coupled bodies, system-wide variation results from the inaccuracy of dimensions in each body. Therefore, tolerances should be selected so that the system-wide variation is within a specified range.

This paper presents and demonstrates a method for allocating tolerances to the dimensions of the bodies. A parametric representation of the contacting surfaces is constructed and combined with a procedure that calculates the resting location based on the inaccurate dimensions. An analytical relation between dimensional variation and variation in the resting location is obtained from multivariate error analysis. Optimal tolerances are computed by minimizing the relative manufacturing cost while respecting constraints on variation in the resting position and orientation. Future work should consider expanding the techniques described in this paper to estimate relative costs as functions of requirements on positioning accuracy.

Appendix A. Determining the HTM from a metrology datum frame to a centroidal coordinate system

The coordinates of the triangles’ three apices $B_1 (D_xB_1, D_yB_1, D_zB_1), B_2 (D_xB_2, D_yB_2, D_zB_2)$ and $B_3 (D_xB_3, D_yB_3, D_zB_3)$ are measured in a metrology datum frame, denoted with the prescript $D$, as shown in Fig. 9.

The centroidal coordinate system, denoted with the prescript $C$, is defined by three criteria

(1) Its origin is located at the intersection of the triangle’s bisectors, which is the centroid $C$.

(2) Its $x$-axis points towards $B_3$.

(3) The three apices $B_1$, $B_2$ and $B_3$ lie in the $xy$-plane.

Criterion (3) implies that the $z$-coordinates of the apices are equal to zero, as shown in Eq. (A.1)

$$C_{zB_1} = C_{zB_2} = C_{zB_3} = 0 \quad (A.1)$$
Since the three angles of a triangle are supplementary, triangle \( \triangle B \) gives Eqs. (A.8)–(A.12).

Applying the law of sines in triangle \( \triangle B \) gives Eq. (A.8)

\[
\frac{CB}{\sin \angle CB_1} = \frac{B_1B_2}{\sin \angle B_1CB} = \frac{B_2B_3}{\sin \angle B_2CB_3}
\]

Inserting Eqs. (A.10)–(A.12) into Eq. (A.9) provides Eq. (A.13):

\[
\angle B_1CB = \frac{\pi}{2} - \frac{\angle B_1}{2}
\]

Eq. (A.14) results from a trigonometry relation applied on Eq. (A.13)

\[
\sin(\angle B_1CB) = \cos \left( \frac{\angle B_1}{2} \right)
\]

Eq. (A.15) is obtained by injecting Eq. (A.14) into Eq. (A.8)

\[
CB_1 = B_1B_3 \times \frac{\sin(\angle B_1/2)}{\cos(\angle B_2/2)}
\]

Criteria (1) and (2) imply that the \( x \)-coordinate of apex \( B_1 \) is equal to the length \( CB_1 \). This condition gives Eqs. (A.16) and (A.17):

\[
\begin{align*}
\bar{c}_{x B_1} &= B_1B_3 \times \frac{\sin(\angle B_1/2)}{\cos(\angle B_2/2)} \\
\bar{c}_{y B_1} &= 0
\end{align*}
\]

The relative position of apex \( B_1 \) with respect to apex \( B_2 \) is known, so its coordinates in the centroidal coordinate system can be found as shown in Eqs. (A.18) and (A.19):

\[
\begin{align*}
\bar{c}_{x B_1} &= \bar{c}_{x B_1} - B_1B_3 \times \cos \left( \frac{\angle B_1}{2} \right) \\
\bar{c}_{y B_1} &= B_1B_3 \times \sin \left( \frac{\angle B_1}{2} \right)
\end{align*}
\]

Coordinates of apex \( B_2 \) are similarly defined in Eqs. (A.20) and (A.21):

\[
\begin{align*}
\bar{c}_{x B_2} &= \bar{c}_{x B_2} - B_2B_3 \times \cos \left( \frac{\angle B_1}{2} \right) \\
\bar{c}_{y B_2} &= B_2B_3 \times \sin \left( \frac{\angle B_1}{2} \right)
\end{align*}
\]

Finding the rotation angles about the three axes of the centroidal coordinate system requires the definition of three distinct unit vectors that start from apex \( B_1 \) and point respectively towards \( B_1, B_2 \) and \( C \). The coordinates of these unit vectors in the metrology datum frame are presented in Eqs. (A.22)–(A.24):

\[
\begin{align*}
\bar{D}_{x B_1B_3} &= \frac{1}{B_1B_3} \times \begin{bmatrix} D_{x B_1} - D_{x B_3} \\ D_{y B_1} - D_{y B_3} \\ D_{z B_1} - D_{z B_3} \end{bmatrix} \\
\bar{D}_{y B_1B_3} &= \frac{1}{B_1B_3} \times \begin{bmatrix} D_{x B_1} - D_{x B_3} \\ D_{y B_1} - D_{y B_3} \\ D_{z B_1} - D_{z B_3} \end{bmatrix} \\
\bar{D}_{z B_1B_3} &= \frac{1}{B_1B_3} \times \begin{bmatrix} D_{x B_1} - D_{x B_3} \\ D_{y B_1} - D_{y B_3} \\ D_{z B_1} - D_{z B_3} \end{bmatrix}
\end{align*}
\]
The following step is the definition of the rotation between the two coordinate systems is defined by Eq. (A.31) .

\[
C\theta_y = \arcsin \left( \frac{D_y B_C - D_z B_B \times D_x B_C}{\sqrt{D_x B_B^2 + D_y B_C^2 + D_z B_B^2}} \right)
\]

Finally, the rotation \( C\theta_y \) about the \( y \)-axis, between the two coordinate systems, is defined by Eq. (A.30)

\[
C\theta_y = \arcsin \left( \frac{D_y B_C - D_z B_B \times D_x B_C}{\sqrt{D_x B_B^2 + D_y B_C^2 + D_z B_B^2}} \right)
\]

To conclude, the homogeneous transformation matrix \( C\theta_T \) between the two coordinate systems is defined by Eq. (A.31).

\[
C\theta_T = \begin{bmatrix}
\cos(C\theta_y) & 0 & 0 & C\theta_y \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
\end{bmatrix}
\]