TOLERANCES IN GEOMETRIC CONSTRAINT PROBLEMS

JOHANNES WALLNER, HANS-PETER SCHRÖCKER, AND SHIMIN HU

ABSTRACT. We study error propagation through implicit geometric problems by linearizing and estimating the linearization error. The method is particularly useful for quadratic constraints, which turns out to be no big restriction for many geometric problems in applications.

1. INTRODUCTION

Geometric constraint solving is a term of Computer-Aided Design and means the problems which arise when the location of geometric objects (points, lines,...) is described via geometric relations (distances, angles,...) between them. The issues crucial for engineering applications are *solvability* of constraint problems and their *sensitivity to errors* [15]. Many methods have been proposed for geometric constraint solving: based on dependency graphs [4, 21, 9, 20, 23], rule-based [5, 10, 11, 29] and numerical ones [19, 22], and methods based on symbolic computing [10, 11, 17]. See also the survey article [14].

From the many works on geometric constraints with a view towards applications we mention the book [8], which extensively deals with distance constraints, and the paper [27].

In this paper we are concerned with the second of the two problems mentioned above, i.e., with the propagation of errors through implicit constraints. Based on the concept of *tolerance zone* [14, 16, 24, 26, 30], we show how to handle error propagation through implicit constraints in a way independent of solvability. The scope of this paper in so far exceeds CAD applications as systems of geometric constraints whose mere solution presents a challenge do not occur very often in practice.

We assume that a certain number of geometric objects is given imprecisely – each of them is only known to be contained in a certain tolerance zone. Other geometric objects are located via constraints, and we want to give tolerance zones for them. This is done by linearizing the system of constraints and estimating the linearization error. For each configuration, this works only up to a certain maximum size of tolerance zones, dependent on the particular instance of the constraint problem we wish to analyze, on the number of objects and constraints involved, and on the behavior of the constraints'

derivatives. In engineering applications, the maximum size of indecomposable constraint problems (see Section 2.2) usually is small, and we shall see that the maximum sizes of tolerance zones in our examples are well above the tolerances used in, say, mechanical engineering. The phenomenon that a tolerance analysis of a particular instance of a constraint problem yields unusually small maximal tolerance zones is then due to the fact that this instance is ill-conditioned (see Examples 10 and 11).

Estimating the linearization error in the way presented below is most efficient if the constraints are quadratic polynomials. The reason for this is that these constraints are reproduced exactly by their second order Taylor expansion. As it is hard to think of geometric relations which are not expressible via quadratic polynomials, this means that for many applications estimating the norms of second derivatives in a certain region as described in Section 5.3 can be replaced by computing those norms once.

The concept of tolerance zone in a certain way generalizes *interval arithmetic*. An interval may be seen as a tolerance zone of a real number, whose location on the real axis is known to lie in some interval. If the coordinates of a geometric object, like a point in $\mathbb{R}^3$, are known to lie in intervals, then the coordinate vector of that geometric object is contained in a certain box which is the product body of those intervals. So using boxes as tolerance zones for geometric objects is the same as using interval arithmetic for these objects' coordinates. One reason why we consider more general tolerance zones is that they can be made invariant with respect to the action of transformation groups acting on geometric objects. Another reason is that tolerance zones, being more general by definition, automatically yield bounds which are at least as tight as those obtained for intervals. A short discussion of these topics can be also found in the introduction to [30].

2. GEOMETRIC CONSTRAINTS

2.1. **Definitions and Coordinatization.** We consider two kinds of geometric entities: the *fixed* variables $p_1, p_2, \ldots$, and the *moving* variables $q_1, q_2, \ldots$ They can be real numbers, points or lines, for instance. We assume that certain equations $c_i(p_1, p_2, \ldots, q_1, q_2, \ldots) = 0$ (the *constraints*) have to hold true, and we assume that knowledge of the fixed variables together with the constraints determines the moving variables not necessarily uniquely, but locally so.

We assume that all geometric entities under consideration are coordinatized in some way:

Example 1. In the Euclidean plane we introduce a Cartesian coordinate system such that a point is represented by two real numbers (ξ, η). The line with equation $n_1\xi + n_2\eta + d = 0$ is represented by coordinates (n_1, n_2, d) which obey the side condition $n_1^2 + n_2^2 = 1$. $\diamond$

The coordinates of the fixed variables p_i are put together in a list "x", and likewise the list "y" comprises the list of coordinates of the variables q_i:

(1)
$$x = (\underbrace{x_1,\ldots,x_{r_1}}_{p_1},\underbrace{x_{r_1+1},\ldots,x_{r_1+r_2}}_{p_2},\ldots,x_n)$$
$$y = (\underbrace{y_1,\ldots,y_{s_1}}_{q_1},\underbrace{y_{s_1+1},\ldots,y_{s_1+s_2}}_{q_2},\ldots,y_m).$$

The system of constraints then is a twice continuously differentiable function

(2) $\quad F\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m : F(x_1,\ldots,x_n,y_1,\ldots,y_m) = (c_1(x,y),\ldots,c_m(x,y)),$

where each component $c_i(x,y)$ represents a constraint. Solving the constraint problem means finding x,y such that $F(x,y) = 0$.

Remark: In (2) the number of constraints equals the number of variable coordinates. In general, if the Jacobian of F with respect to $y_1,\ldots,y_m$ has maximal rank, there are exactly m constraints needed to determined m variables locally. If there are more than m constraints, the solvability of the system is due to a nontrivial relation between the constraints, and some constraints may be deleted in order to have the same set of solutions locally.

In case the Jacobian does not have full rank, it is possible that less than m equations determine m variables. For algebraic and analytic equations, such a singularity happens either for all x, or for almost no x. In fact, any system of constraints can be turned into one which is always singular, by simply considering the equation $\sum c_i(x,y)^2 = 0$ instead of $F(x,y) = 0$. Such problems may be considered ill-posed. In view of the 'almost no' statement above, we consider only systems of constraints whose number equals the number of variable coordinates. $\diamondsuit$

Example 2. If c_i represents a distance constraint regarding two points, one fixed, and one moving, then it has the form $c_i(x,y) = (x_j - y_k)^2 + (x_{j+1} - y_{k+1})^2 - d$. Here the fixed and the moving point under consideration are represented by the coordinates (x_j,x_{j+1}) and (y_k,y_{k+1}), respectively. An incidence constraint concerning a point and a line, both moving, has the form $c_i(x,y) = y_iy_k + y_{i+1}y_{k+1} + y_{k+2}$. The fact that the coordinates (y_k,y_{k+1},y_{k+2}) of a line must obey the side condition $y_k^2 + y_{k+1}^2 - 1 = 0$ is taken into account by adding just this equation as one further constraint. $\diamondsuit$

2.2. **Graphable Constraint Problems and Decomposability.** If each constraint involves exactly two of the given variables, a graph with vertices $p_1, p_2, \ldots, q_1, q_2, \ldots$ is defined in the obvious way, with one edge per constraint. Additional edges are those between any two fixed variables. A constraint which involves only one variable is shown as a loop in the graph.

Such a graph may contain information on the generic solvability of a constraint problem: Constraint graphs have been used in order to decompose the problem of solving a given system of constraints (i.e., to find the q_i's, if the p_j's are given) into smaller subproblems (see the introductory references which refer to graph-based methods). In this
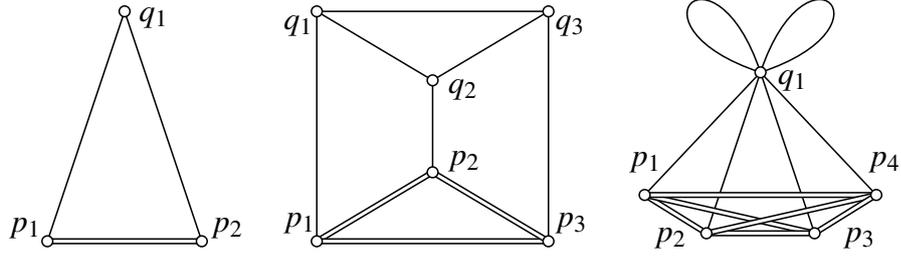
FIGURE 1. Constraint graphs for Examples 3, 4, and 5 (from left to right). The edges which connect fixed variables are indicated by double lines.

paper we are not concerned with this topic, which in Computer-Aided Design is a very important one, considering the sheer size of constraint problems which occur. This paper is relevant for applications in so far as it presents a method for the tolerance analysis of non-trivial indecomposable constraint problems. We give three simple examples.

Example 3. The points $p_1 = (x_1, x_2)$, $p_2 = (x_3, x_4)$, and $q_1 = (y_1, y_2)$ in the Euclidean plane $\mathbb{R}^2$ are connected by two distance constraints $c_i(x, y) = \|p_i - q_1\|^2 - d_i^2$ $(i = 1, 2)$. This problem is graphable and depicted in Figure 1. $\diamond$

Example 4. Here p_1, p_2, p_3 are lines in the Euclidean plane, and q_1, q_2, q_3 are points. The constraints are incidence of p_i and q_i $(i = 1, 2, 3)$ and known distances for q_1, q_2, and q_3. The corresponding graph is shown by Figure 1. We mention this particular set of geometric constraints here because q_1, q_2, q_3 can be found with ruler and compass from p_1, p_2, p_3, in contrast to the underlying graph being called 'not ruler-and-compass constructible' in [20]. The reader should be warned that in this area terminology can sometimes be misleading. $\diamond$

Example 5. Normalized Plücker coordinates for lines in $\mathbb{R}^3$ as defined e.g. in [25] consist of six numbers $u_1, \ldots, u_6$ which fulfill the relations $u_1^2 + u_2^2 + u_3^2 = 1$ and $u_1 u_4 + u_2 u_5 + u_3 u_6 = 0$. The condition that the two lines with coordinates $(u_1, \ldots, u_6)$ and $(v_1, \ldots, v_6)$ meet each other is expressed by $u_1 v_4 + u_2 v_5 + u_3 v_6 + u_4 v_1 + u_5 v_2 + u_6 v_3 = 0$. The right hand graph of Figure 1 corresponds to the constraint problem with four fixed lines $p_1, \ldots, p_4$ and one moving line q_1 which is to meet all of them. $\diamond$

Remark: Constraints which involve more than two variables can be split up by introducing new variables, thus making the constraint problem graphable. This is not relevant for tolerance analysis as presented in this paper, but may help with visualization and allow application of graph-based algorithms. $\diamond$

Remark: The special case of distance constraints between points of $\mathbb{R}^d$ is that of a *framework*. This problem is very well studied and has a long history, beginning with [6]. Here it is known that for almost all x, y the rank of the Jacobian of F is the same, and that for such x, y, the framework is *either* both infinitesimally rigid and rigid, *or* both infinitesimally flexible and flexible. The graph is consequently called rigid or flexible.

There might however be singular realizations of the underlying graph as an infinitesimally flexible, but rigid, framework. The possibility that a framework is flexible but infinitesimally rigid is known not to occur. For these notions and results, see [1, 2, 7]. For the Euclidean plane, a graph-theoretical characterization of rigid graphs is known [18]. The analogous problem for Euclidean three-space is still unsolved, however. $\diamondsuit$

3. TAYLOR EXPANSION

Our method of computing and estimating tolerance zones consists of linearization and estimating the second order remainder terms. We first define some notation concerning the Taylor expansion of the function F defined by (2).

3.1. **Linear and Bilinear Mappings: Notation.** We use the symbols U, V, W for linear spaces and $L(U, W)$ for the linear space of linear mappings of U to W. Further we use the symbol $B(U, V, W)$ for the linear space of bilinear mappings of $U \times V$ to W. We use the notation

$$(3) \qquad \alpha \cdot u, \quad \beta[u, v],$$

to indicate that we apply α to u and β to the pair (u, v). The reason for this is that we have linear and bilinear mappings which depend on some variables, like "$\alpha(u)$", and we want to avoid confusion.

Subscripts indicate coefficients of vectors in U, V, W with respect to some previously defined standard bases. Then we say that $\alpha \in L(U, W)$ and $\beta \in B(U, V, W)$ have coefficients α_{ri} and β_{rij}, if

$$(4) \qquad [\alpha \cdot u]_r = \sum_i \alpha_{ri} u_i, \quad \beta[u, v]_r = \sum_{i,j} \beta_{rij} u_i v_j.$$

There are natural isomorphisms $\phi : B(U, V, W) \to L(U, L(V, W))$ and $\psi : B(U, V, W) \to L(V, L(U, W))$ defined by

$$(5) \qquad \beta[u, v] = \beta^\phi(u) \cdot v, \quad \beta[u, v] = \beta^\psi(v) \cdot u.$$

Obviously, the coefficients of $\beta^\phi(u)$ and $\beta^\psi(v)$ are computed via

$$(6) \qquad [\beta^\phi(u)]_{rj} = \sum_i u_i \beta_{rij}, \quad [\beta^\psi(v)]_{ri} = \sum_j v_j \beta_{rij}.$$

3.2. **Taylor Expansion of the constraints.** The function F of (2) takes as argument a vector of $\mathbb{R}^n \times \mathbb{R}^m$, which we symbolically denote by (x, y), with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. For all (u, v) and $(h, k) \in \mathbb{R}^n \times \mathbb{R}^m$, there is $\theta \in [0, 1]$ such that

$$
\begin{aligned}
(7) \qquad F(u+h, v+k) &= F(u, v) + F_{,x}(u, v) \cdot h + F_{,y}(u, v) \cdot k \\
&+ \frac{1}{2} F_{,xx}(u+\theta h, v+\theta k)[h, h] + F_{,xy}(u+\theta h, v+\theta k)[h, k] \\
&+ \frac{1}{2} F_{,yy}(u+\theta h, v+\theta k)[k, k].
\end{aligned}
$$

The symbols $F_{,x}, \ldots$ denote linear and bilinear mappings as follows:

$$(8) \qquad F_{,x} \in L(\mathbb{R}^n, \mathbb{R}^m), \; F_{,y} \in L(\mathbb{R}^m, \mathbb{R}^m),$$
$$F_{,xx} \in B(\mathbb{R}^n, \mathbb{R}^n, \mathbb{R}^m), \; F_{,xy} \in B(\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^m), \text{ and } F_{,yy} \in B(\mathbb{R}^m, \mathbb{R}^m, \mathbb{R}^m),$$

such that the coefficients in the sense of (4) of $F_{,x}$ and $F_{,y}$ are given by the partial derivatives $\partial c_r / \partial x_i$ and $\partial c_r / \partial y_i$, respectively; and analogously the coefficients of $F_{,xx}$, of $F_{,xy}$ and of $F_{,yy}$ are given by $\partial^2 c_r / \partial x_i \partial x_j$, by $\partial^2 c_r / \partial x_i \partial y_j$, and by $\partial^2 c_r / \partial y_i \partial y_j$. For the general theory of derivatives we refer the reader e.g. to [3, Section X.4].

4. LOCAL SOLUTIONS AND TOLERANCE ZONES

Suppose that we are given a solution (u, v) of the constraint problem F, which means that $F(u, v) = 0$. A local solution of the tolerance problem which extends the solution (u, v) is a function G, defined in a connected neighborhood U of u such that

$$(9) \qquad G : U \to \mathbb{R}^m, \; G(u) = v \text{ and for all } x \text{ in } U : F(x, G(x)) = 0.$$

It follows from the inverse function theorem that such a local solution exists if

$$(10) \qquad\qquad\qquad F_{,y}(u, v) \text{ is nonsingular,}$$

and that it is essentially unique. Tolerance analysis means that we allow the fixed entities $p_1, p_2, \ldots$ to vary in respective tolerance zones $P_1, P_2, \ldots$, and ask for the possible locations of $y = (q_1, q_2, \ldots)$. The most general meaning of 'possible' is that we seek all solutions of the equation

$$(11) \qquad\qquad F(x, y) = 0, \text{ such that } x \in X := P_1 \times P_2 \times \ldots,$$

where the symbol $P_1 \times P_2 \times \ldots$ means the set of vectors x such that each single p_i is contained in the corresponding P_i. It is usually not useful to ask this general question, but to restrict oneself to *local* solutions as defined above, i.e., we would like to compute

$$(12) \qquad\qquad\qquad G(X), \quad (X = P_1 \times P_2 \times \ldots).$$

We define the functions $G^{(j)}$ as those coordinates of G, which belong to the geometric object q_j. Then a tolerance zone for the geometric entity q_j, if the fixed entities are allowed to vary in the domains P_i, is given by

$$(13) \qquad\qquad\qquad Q_j = G^{(j)}(X).$$

We have not yet specified which sets P to allow as tolerance zones of a geometric entity $p \in \mathbb{R}^r$. The most general definition is that of a subset of $\mathbb{R}^r$ which contains p, and in this paper we require that P is connected and bounded. It is useful to imagine tolerance zones as nicely shaped sets which are not too big. Of course, when actually computing with tolerance zones we restrict ourselves to sets which are computationally tractable.

Example 6. Consider points $p_1 = (x_1, x_2) = (0, 0)$, $p_2 = (x_3, x_4) = (60, 0)$, $q_1 = (y_1, y_2)$ in the Euclidean plane and the distance constraints $c_1(x, y) = (x_1 - y_1)^2 + (x_2 - y_2)^2 - 2900$, $c_2(x, y) = (x_3 - y_1)^2 + (x_4 - y_2)^2 - 4100$. A solution is given by $q_1 = (20, 50)$. We want to find the tolerance zone Q_1, if tolerance zones P_1, P_2 are given (Figure 2).
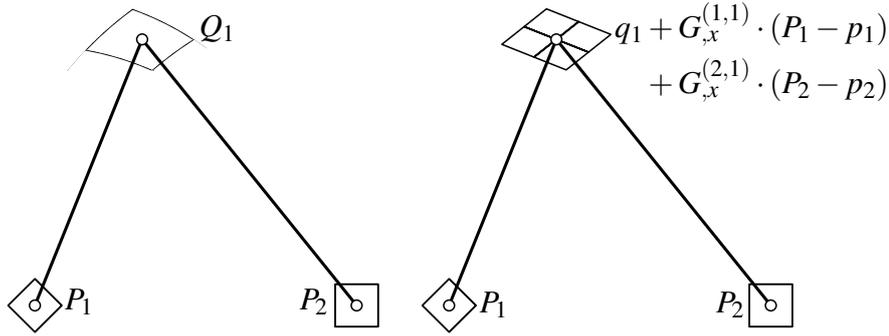
FIGURE 2. Exact and linearized tolerance zones for Examples 6 and 7.

It is easy to see that in our example the boundary of Q_1 consists of circular arcs. The result is shown in Figure 2, left. ◇

5. LINEAR TOLERANCE ANALYSIS AND THE LINEARIZATION ERROR

5.1. Linearizing Local Solutions.
The Taylor expansion of a local solution reads

$$(14) \quad v = G(u), \ v+k = G(u+h) \ \Rightarrow \ v+k = v + G_{,x}(u) \cdot h + \frac{1}{2}G_{,xx}(u+\theta h)[h,h].$$

We combine (7) and (14) and get

$$0 = F_{,x}(u,v) \cdot h + F_{,y}(u,v) \cdot k + o(2), \quad k = G_{,x}(u) \cdot h + o(2) \quad \text{for all } h$$

$$(15) \quad \Longrightarrow \quad G_{,x}(u) = -F_{,y}(u,v)^{-1}F_{,x}(u,v).$$

Equation (15) is the basis of linear tolerance analysis. The first order approximation G_{lin} of a local solution is given by

$$(16) \quad G_{\text{lin}}(u+h) = G(u) + G_{,x}(u) \cdot h.$$

Instead of computing tolerance zones of the moving variables $q_1, q_2, \ldots$ via (12) or (13), we use G_{lin} and get *linearized tolerance zones* $\bar{Q}_j$ of the moving variables q_j, and a total linearized tolerance zone $\bar{Y}$ of y:

$$(17) \quad \bar{Y} = G_{\text{lin}}(X) = G(u) + G_{,x}(u) \cdot (X-u), \quad \bar{Q}_j = q_j + G_{\text{lin}}^{(j)}(u) \cdot (X-u).$$

The partition of vectors x and y into blocks which correspond to geometric objects p_i and q_j according to (1), defines a partition of the matrix $G_{,x}$ into block matrices $G_{,x}^{(i,j)}$. Then (17) becomes

$$(18) \quad \bar{Q}_j = G_{\text{lin}}^{(j)}(X) = q_j + \sum_i G_{,x}^{(i,j)} \cdot (P_i - p_i).$$

This Minkowski addition of the sets $G_{,x}^{(i,j)} \cdot (P_i - p_i)$ is particularly simple to compute if they are at most two-dimensional and convex [12].

Example 7. We continue Example 6 and compute

$$(19) \qquad G_{,x} = \left[G_{,x}^{(1,1)} \middle| G_{,x}^{(2,1)} \right] = \frac{1}{30} \begin{bmatrix} 10 & 25 & 20 & -25 \\ 8 & 20 & -8 & 10 \end{bmatrix}.$$

The resulting linearized tolerance zone is shown by Figure 2, right. Both $G_{,x}^{(1,1)}$ and $G_{,x}^{(2,1)}$ are singular, and $G_{,x}^{(i,1)} \cdot (P_i - p_i)$ is a straight line segment. It follows that $\bar{Q}_1$ is a parallelogram. $\diamondsuit$

Remark: The matrix $G_{,x}^{(i,j)}$ corresponds to the movement of the point q_j, if all points p_k except the point p_i are fixed. Therefore the straight line segment $G_{,x}^{(i,1)} \cdot (P_i - p_i)$ of Example 7 is orthogonal to the vectors $q_1 - p_i$.

A constructive method of finding those straight line segments and the matrices $G_{,x}^{(i,j)}$ for this and other examples is mentioned e.g. in [33] in connection with finding velocities of points in moving kinematic chains. It is based on orthogonality relations as the one mentioned above, but fails if the indecomposable parts of the underlying graph become too big. $\diamondsuit$

Remark: The infinitesimal rigidity of frameworks referred to in an earlier remark is characterized by the regularity condition (10). $\diamondsuit$

5.2. **Computing Norms of Linear and Bilinear Mappings.** For the convenience of the reader we repeat some facts concerning definition and computation of norms of linear and bilinear operators. They are needed for estimating the linearization error. We assume that $\alpha \in L(U,W)$, $\beta \in B(U,V,W)$, and that the linear spaces U,V,W are equipped with norms. Then

$$(20) \qquad \|\alpha\| := \sup_{\|u\| \leq 1} \|\alpha \cdot u\|, \quad \|\beta\| := \sup_{\|u\|,\|v\| \leq 1} \|\beta[u,v]\|.$$

Thus also $L(U,W)$, $L(V,W)$, and $B(U,V,W)$ become normed spaces. It is not difficult to show that with respect to these norms,

$$(21) \qquad \|\beta^{\phi}\| = \|\beta^{\psi}\| = \|\beta\|$$

(ϕ and ψ are the natural isomorphisms defined in Section 3.1). The examples in this paper are such that U, V and W are real vector spaces of finite dimension and the norms in U,V,W are L^p norms with $p = 1,2,\infty$. The reason for using the 2-norm is its geometric significance; the reason for using the ∞-norm is that computations become simple; and the reason for using the 1-norm is that also here the unit sphere is a convex polyhedron (like for the ∞-norm), so some computations are not difficult either.

The actual computation of $\|\alpha\|$ is well known for the cases of the 1-, 2-, and ∞-norms (cf. [13]). In general, if the unit sphere S_U in U is a convex polyhedron with vertices x_i, then for all $\alpha \in L(U,W)$,

$$(22) \qquad \|\alpha\| = \max_i \|\alpha \cdot x_i\|,$$

for any norm in W. It is therefore not difficult to compute the norm of a bilinear mapping β, if S_U is a convex polyhedron: We may use (21) and compute the norm of β^ψ with respect to the norms of U and $L(V,W)$. Analogously, if the unit sphere S_V of V is a convex polyhedron, we may compute $\|\beta\|$ as the norm of β^ϕ with respect to the norms in V and $L(U,W)$.

This method does not work if both U and V are equipped with the 2-norm. In that case we additionally use the ∞-norm in W. It is clear that

$$(23) \qquad \|\beta\|_{2,2,\infty} = \max_r \max_{\|u\|,\|v\|\le 1} |\beta_r[u,v]|,$$

where β_r is the r-th component function of β. It follows that $\|\beta\|_{2,2,\infty}$ is the maximum singular value of the matrices β_{rij}, r fixed. A case not covered so far is that also W is equipped with the 2-norm, i.e., the problem of computing $\|\beta\|_{2,2,2}$. In that case we use the estimate

$$(24) \qquad \|\beta\|_{2,2,2} \le \sqrt{\dim W}\,\|\beta\|_{2,2,\infty}.$$

5.3. **Estimating the Linearization Error.** The *linearization error* is the difference between an exact local solution G and the linearized one, G_{lin}. In the following computation we use (7), but drop the arguments "(u,v)" and "$(u+\theta h, v+\theta k)$".

$$F(u,v) = 0,\; v+k = G(u+h),\; v+k_{\mathrm{lin}} = G_{\mathrm{lin}}(u+h) \Rightarrow$$

$$0 = F(u+h,v+k) = F_{,x}\cdot h + F_{,y}\cdot k + \frac{1}{2}F_{,xx}[h,h] + F_{,xy}[h,k] + \frac{1}{2}F_{,yy}[k,k]$$

$$\Rightarrow\quad k - k_{\mathrm{lin}} = -\frac{1}{2}F_{,y}^{-1}\left(F_{,xx}[h,h] + 2F_{,xy}[h,k] + F_{,yy}[k,k]\right).$$

It follows that

$$(25) \qquad \|k - k_{\mathrm{lin}}\| \le \frac{1}{2}\|F_{,y}^{-1}\|\left(\|F_{,xx}\|\,\|h\|^2 + 2\,\|F_{,xy}\|\,\|h\|\,\|k\| + \|F_{,yy}\|\,\|k\|^2\right).$$

When computing norms, recall that (8) describes domain and range of each of the operators which occur in (25).

Definition 1. *Suppose that the three vector spaces $\mathbb{R}^n, \mathbb{R}^m$ and $\mathbb{R}^m$ used in the definition of F by (2) are equipped with norms, and suppose further that in looking for solutions of the given tolerance problem $F(u,v) = 0$ we restrict ourselves to (u,v) contained in such a subset of $\mathbb{R}^n \times \mathbb{R}^m$, where there are estimates*

$$(26) \qquad \|F_{,xx}(u,v)\| \le \alpha,\; \|F_{,xy}(u,v)\| \le \beta,\; \|F_{,yy}(u,v)\| \le \gamma,\quad (\alpha^2 + \beta^2 + \gamma^2 > 0),$$

with respect to the norms chosen previously. Then we define

$$(27) \qquad \Delta(s,t) := \frac{1}{2}(\alpha s^2 + 2\beta\, st + \gamma t^2).$$

Remark: Upper bounds as required by (26) are particularly simple to give if F is a *quadratic* function, because then $F_{,xx}$, $F_{,xy}$, and $F_{,yy}$ depend neither on x nor on y.

If F is linear, then the norms $\|F_{,xx}\|, \dots$ are zero and linearization is exact. For our purposes it is essential that $\Delta(s,t)$ is non-zero if $s,t > 0$. Therefore we require that $\alpha^2 + \beta^2 + \gamma^2 > 0$. $\diamond$

Example 8. We continue Examples 6 and 7. Table 1 shows $\Delta(s,t)$ for different choices of norms. The starred figures are upper bounds, cf. Equation (24). $\diamond$

A tolerancing concept for geometric constraint problems based on linearization must provide information under what circumstances the linearization error can be bounded. This is done by the following two results:

Lemma 1. *Consider a solution (u,v) of the constraint problem $F(x,y) = 0$, and assume that $\Delta(s,t)$ is defined according to Definition 1. We assume that there is a local solution G and its linearization G_{lin} which extend (u,v), such that $v + k = G(u+h)$ and $v + k_{\mathrm{lin}} = G_{\mathrm{lin}}(u+h)$. Then the linearization error is bounded by*

$$(28) \qquad \|k - k_{\mathrm{lin}}\| \leq \|F_{,y}(u,v)^{-1}\| \cdot \Delta(\|h\|, \|k\|).$$

Proof: This follows directly from (25) and (27). $\qquad \square$

Theorem 1. *Under the same assumptions as in Lemma 1, let*

$$(29) \qquad C_{\max} = \frac{\|G_{,x}(u)\|}{\|F_{,y}(u,v)^{-1}\| \cdot \Delta(1, 2\|G_{,x}(u)\|)}.$$

Choose $C < C_{\max}$ and let $C' = \|G_{,x}(u)\| C$. Then if $\|h\| \leq C$, the local solution $v + k = G(u+h)$ obeys the inequalities

$$(30) \qquad \|k\| < 2C', \quad \|k - k_{\mathrm{lin}}\| \leq \|F_{,y}(u,v)^{-1}\| \cdot \Delta(C, 2C') < C'.$$

Proof: First we give a condition on C such that

$$(31) \qquad C'' := \|F_{,y}(u,v)^{-1}\| \cdot \Delta(C, 2C') < C'.$$

This is easily seen to be equivalent to

$$(32) \qquad C^2 \|F_{,y}(u,v)^{-1}\| \, \Delta\big(1, 2\|G_{,x}(u)\|\big) < C \, \|G_{,x}(u)\|,$$

and so in turn is equivalent to the condition $C < C_{\max}$. It is our aim to show that the inequality $\|h\| < C$ implies that either $\|k\| \geq 2C'$ or $\|k\| \leq C' + C''$. Then by (31), there is a certain region, bounded by the spheres of radius $C' + C''$ and $2C'$, which contains no vector k. As G was supposed to be a local solution, and the sphere $\|h\| \leq C$ is connected, the local value of k must remain inside the sphere of radius $C' + C''$. This implies the statement of the theorem.

So we assume that $\|k\| < 2C'$. According to Lemma 1, the linearization error is bounded by C''. By definition of k_{lin}, we have

$$(33) \qquad \|k\| \leq \|k_{\mathrm{lin}}\| + \|k - k_{\mathrm{lin}}\| \leq \|G_{,x}(u)\| \, \|h\| + C'' \leq C' + C''.$$

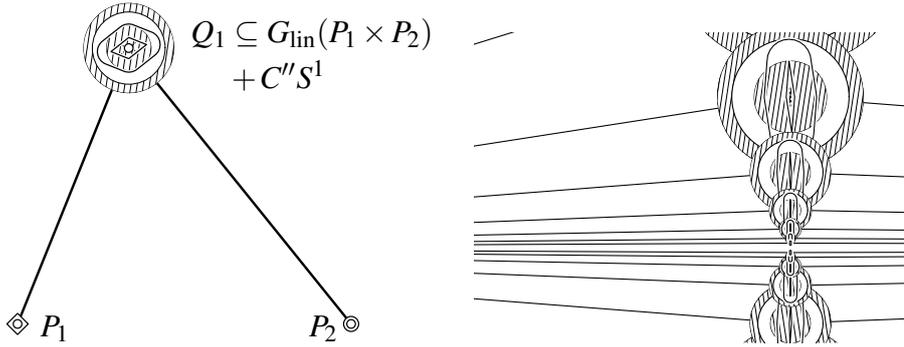This concludes the proof. $\qquad \square$

FIGURE 3. The framework of Example 9 and a detail of Example 10. The concentric circles have the radii $2C'$, $C' + C''$, and C'.

Remark: The estimates given here can be sharpened a little without much effort: Theorem 1 gives an upper bound of the following form: $\|h\| < C$ implies that $\|k_{\text{lin}}\| \leq C'$ and $\|k - k_{\text{lin}}\| \leq C''$. We get the relation

$$(34) \qquad \|k - k_{\text{lin}}\| \leq \|F_{,y}^{-1}\| \Delta(\|h\|, \|k\|) \leq \|F_{,y}^{-1}\| \Delta(C, C' + C'') =: C''' < C''.$$

by applying Lemma 1 a second time. $\Diamond$

6. EXAMPLES AND IMPLEMENTATION ISSUES

Example 9. Figure 3 illustrates the statement of Theorem 1, when applied to the framework of Example 6. We have chosen the 1-, 2-, and ∞-norms for u, v, and $F(u,v)$. Computing Δ and C_{max} presents no problems, as the constraints are quadratic.

The tolerance zones P_1, P_2 have been selected such that their product $P_1 \times P_2$ is contained in a ball of radius $C < C_{\text{max}}$. Then Theorem 1 applies. The concentric circles shown in the picture have the radii $2C'$, $C' + C''$, and C'. We can see clearly that the linearized tolerance zone of Q_1 (the parallelogram) is contained in the circle of radius C', and we know that the linearization error is bounded from above by C''. Hence, the tolerance zone Q_1 must be contained in an outer offset of radius C'' of the linearized tolerance zone and, at the same time, in the circle of radius $C' + C''$. $\Diamond$

Example 10. The constraint problem of Example 6 has two symmetric solutions. For each solution q_1, also q_1's reflection in the line $p_1 p_2$ is one. It follows from Theorem 1 that the ball of radius $\|G_{,x}\| C_{\text{max}}$ (which is the maximum size of the moving variables' tolerance zone such that Theorem 1 applies), must not intersect its own reflection. By changing the constraints as to allow solutions near the line $p_1 p_2$, this causes the phenomenon that C_{max} becomes smaller as the solution q_1 approaches its reflection, and the tolerance problem becomes ill-conditioned. $\Diamond$

Example 11. The choice of norms for u, v, and $F(u,v)$ influences the value of C_{max}. Table 1 shows this dependence for the hexagonal framework of Figure 4. We have chosen $C = 0.6 C_{\text{max}}$ and computed the tolerance zones with respect to the 2, ∞, and

TABLE 1. Numerical data for Examples 8 (left) and 11 (right).

$\|u\|$	$\|v\|$	$\|F\|$	$100\Delta(s,t)$	$\|u\|$	$\|v\|$	$\|F\|$	$C_{\max}$
∞	∞	∞	$3.3\ s^2 + 3.3\ st + 3.3\ t^2$	∞	∞	∞	$1.03 \cdot 10^{-2}$
∞	1	∞	$4.0\ s^2 + 2.0\ st + 2.0\ t^2$	∞	1	∞	$0.39 \cdot 10^{-2}$
∞	1	2	$4.8\ s^2 + 2.2\ st + 2.2\ t^2$	∞	1	2	$0.59 \cdot 10^{-2}$
1	2	∞	$1.7\ s^2 + 1.7\ st + 1.7\ t^2$	1	2	∞	$2.65 \cdot 10^{-2}$
2	2	2	$1.7^\star s^2 + 1.7^\star st + 1.7^\star t^2$	2	2	2	$1.79 \cdot 10^{-2}$
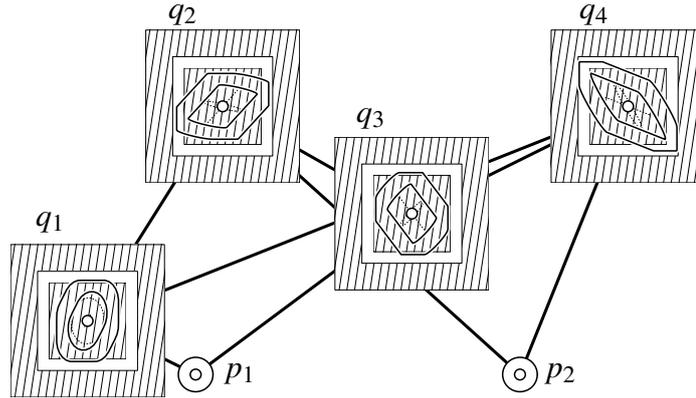
FIGURE 4. Linearized tolerance zones (scaled by a factor of 60) for a framework with respect to the 2-, ∞- and ∞-norms. Some values of $C_{\max}$ are given in Table 1.

∞-norms. The tolerance zones are scaled by a factor of 60 for better visualization. Note that the C''-offsets of the linearized tolerance zones have to be taken with respect to the ∞-norm.

This example of an indecomposable constraint graph is unlikely to occur in CAD. The reason why maximum tolerance zones are much smaller than in Example 9 (Figure 2) is the size of the problem.

It is well known that this framework is singular (i.e., $F_{,y}(u,v)$ is singular), if and only if the six points p_1, p_2, q_1, q_2, q_3, q_4 are contained in an algebraic curve of degree two, which includes conics and pairs of lines. This is true in general for planar frameworks where the constraint graph is bipartite, cf. [32, 31, 28]. If the points q_j are close to a conic section, then the corresponding values for $C_{\max}$ become even smaller. $\diamond$

Obviously the local solutions do not change if we multiply some constraints by factors, but the computation of $C_{\max}$ is affected by it. In Examples 9 and 11, all constraints are of the same type. This is not true for Example 4, where we have both distance and incidence constraints. It is not clear a priori how equations which express different geometric relations are to be scaled so as not to badly influence the computations. As this scaling does not change G, $\|G_{,x}(u)\|$ remains the same, and therefore so does the ratio
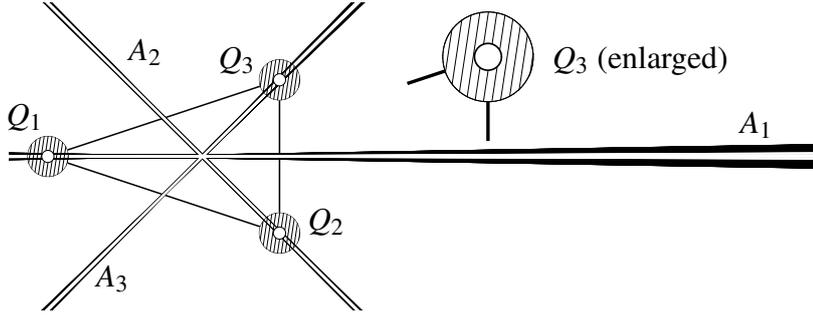
FIGURE 5. Visualization of tolerance zones of radius $C_{\max} = C$ and $\|G_{,x}\|C_{\max}$ before and after constraint balancing (Example 12).

C'/C in Theorem 1. This means that balancing might enlarge the region of validity of the theorem without deteriorating the ratio of the tolerance zones of fixed and moving entities. The dependence of $C_{\max}$ on the scaling of equations is not easy to analyse in general.

Remark: For a list of geometric constraint problems, the dependence of $C_{\max}$ on the normalization of constraints has been analysed by means of varying the unit length for measuring coordinates. It turned out that $C_{\max}$ often is optimal or nearly optimal either for a specific choice of unit length, or for all unit lengths smaller than a given value. The unit lengths in question are not difficult to compute. Unfortunately there is no general theory available at present. $\diamondsuit$

Example 12. We continue Example 4 and compute tolerance zones for the straight lines p_i and the points q_j. In Figure 5, the maximal regions of validity of Theorem 1 for a balanced and an unbalanced constraint system are visualized.

We use the ∞-norm for the fixed variables and choose tolerance zones P_1, P_2, P_3 in the coordinate space $\mathbb{R}^3$ for lines such that $P_1 \times P_2 \times P_3$ is contained in a ball of radius $C_{\max}$. Here P_i is chosen as an ∞-sphere of radius $0.99C_{\max}$.

The choice of constraints as in Example 2 is not optimal. For this specific problem, it seems reasonable that $C_{\max}$ will be maximal if the coefficients in the various constraint equations are of the same magnitude. This is indeed the case, as numerical experiments show. Figure 5 shows the area A_j traced out by the lines of P_j – it is filled in white before and in black after balancing. The corresponding tolerance zones Q_j are shown as small white and bigger hatched disks. Some data for both the balanced and the unbalanced case is presented in Table 2. $\diamondsuit$

Example 13. We present an example with nonquadratic constraints and consider the following problem (Figure 6): Given is a Cartesian coordinate frame, defined by its origin $p_1 = (x_1, x_2)$ and a point $p_2 = (x_3, x_4)$ on one axis. This frame defines the position of a certain curve (here, a sine curve). The point q_1 is found by intersecting this curve with a fixed line l. We assume that p_1 and p_2 are given imprecisely by their respective tolerance zones and we are interested in a tolerance zone of q_1 on l.

TABLE 2. Comparison between unbalanced and balanced constraint equations (Example 12).

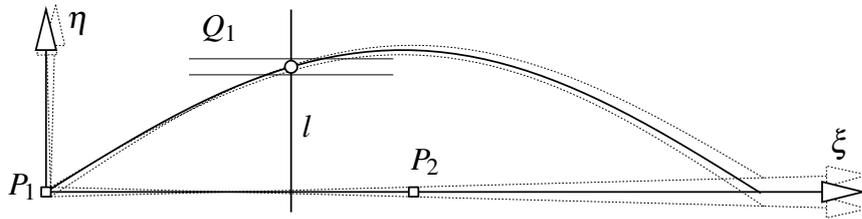| $\|u\|$	$\|v\|$	$\|F\|$	$C_{\max}$	$C^{\mathrm{b}}_{\max}$	$C^{\mathrm{b}}_{\max}/C_{\max}$
∞	∞	1	$0.302 \cdot 10^{-2}$	$1.365 \cdot 10^{-2}$	4.51
∞	2	2	$0.355 \cdot 10^{-2}$	$1.569 \cdot 10^{-2}$	4.41
1	∞	1	$0.769 \cdot 10^{-2}$	$4.054 \cdot 10^{-2}$	5.27
1	1	∞	$0.867 \cdot 10^{-2}$	$2.046 \cdot 10^{-2}$	2.35
1	2	∞	$1.669 \cdot 10^{-2}$	$4.116 \cdot 10^{-2}$	2.46
2	2	∞	$1.277 \cdot 10^{-2}$	$3.545 \cdot 10^{-2}$	2.77

FIGURE 6. Tolerance zones for Example 13.

In Figure 6, the tolerance zones P_i and Q_1 for the points p_i and q_1 are visualized (the latter by a pair of parallel lines as it is only one-dimensional). The bilinear mapping $F_{,yy}$ is not constant and its norm has to be estimated. $\diamondsuit$

7. CONCLUSION

We have studied the propagation of errors in the form of tolerance zones through implicit constraints. The usage of tolerance zones generalizes interval arithmetic in the sense that intervals are tolerance zones of real numbers. The method works by linearization and giving upper bounds for the linearization error. Depending on the problem, there is a maximum size of tolerance zone for which this method is applicable. Computing this radius of validity and the linearization error requires upper bounds on the constraints' second derivatives in the form of bilinear mappings. It turns out that such bounds are found easily if the constraints are quadratic, which for geometric constraints is very often the case.

ACKNOWLEDGEMENTS

The authors are grateful to the reviewers for their careful reading of the manuscript and their valuable suggestions.

References

[1] L. Asimov and B. Roth. The rigidity of graphs. *Trans. Amer. Math. Soc.*, 245:171–190, 1978.

[2] L. Asimov and B. Roth. The rigidity of graphs II. *J. Math. Anal. Appl.*, 68:171–190, 1979.

[3] R. Bhatia. *Matrix Analysis*. Springer, 1997.

[4] W. Bouma, I. Fudos, C. Hoffmann, J. Cai, and R. Paige. Geometric constraint solver. *Computer-Aided Design*, 27:487–501, 1995.

[5] B. Bruederlin. Using geometric rewriting rules for solving geometric problems symbolically. *Theoret. Comput. Sci.*, 116:291–303, 1993.

[6] A. L. Cauchy. Ii^e mémoire sur les polygones et les polyèdres. *Journal de l'Ecole Polytechnique*, 19:87–98, 1813.

[7] R. Conelly. On generic global rigidity. In P. Gritzmann and B. Sturmfels, editors, *Applied Geometry and Discrete Mathematics — The Victor Klee Festschrift*. American Mathematical Society, 1989.

[8] G. M. Crippen and T. F. Havel. *Distance geometry and molecular conformation*, volume 15 of *Chemometrics Series*. Research Studies Press Ltd., Chichester, 1988.

[9] I. Fudos and C. Hoffmann. A graph-constructive approach to solving systems of geometric constraints. *ACM Transactions on Graphics*, 16:179–216, 1997.

[10] X. S. Gao and S. C. Chou. Solving geometric constraint systems. I. a global propagation approach. *Computer-Aided Design*, 30:47–54, 1998.

[11] X. S. Gao and S. C. Chou. Solving geometric constraint systems. II. a symbolic approach and decision of rc-constructibility. *Computer-Aided Design*, 30:115–122, 1998.

[12] P. Ghosh. A unified computational framework for Minkowski operations. *Computers & Graphics*, 17:357–378, 1993.

[13] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*, chapter Matrix Norms. Soc. Industrial and Appl. Math, 1996.

[14] C. M. Hoffmann, A. Lomonosov, and M. Sitharam. *Computing in Euclidean Geometry*, pages 266–298. World Scientific, 1995.

[15] Christoph M. Hoffmann. Robustness in geometric computations. *Journal of Computing and Information Science in Engineering*, 1:143–156, 2001.

[16] S.-M. Hu and J. Wallner. Error propagation through geometric transformations. Technical Report 102, Institut für Geometrie, TU Wien, 2003.

[17] K. Kondo. Algebraic method for manipulation of dimensional relationships in geometric models. *Computer-Aided Design*, 24:141–147, 1992.

[18] G. Laman. On graphs and rigidity of plane skeletal structures. *J. Engrg. Math.*, 4:331–340, 1970.

[19] H. Lamure and D. Michelucci. Solving geometric constraints by homotopy. *IEEE Trans. Vis. Comp. Graph.*, 2:28–34, 1996.

[20] J. Y. Lee and K. Kim. A 2-D geometric constraint solver using DOF-based graph reduction. *Computer-Aided Design*, 30:883–896, 1998.

[21] K.-Y. Lee, O-H. Kwon, J.-Y. Lee, and T.W. Kim. A hybrid approach to geometric constraint solving with graph analysis and reduction. *Adv. Eng. Software*, 34:103–113, 2003.

[22] Y.-T. Li, S.-M. Hu, and J.-G. Sun. A constructive approach to solving 3-D geometric constraint systems using dependence analysis. *Computer-Aided Design*, 34:97–108, 2002.

[23] R. A. Light and D. C. Gossard. Modification of geometric models through variational geometry. *Computer-Aided Design*, 14:209–214, 1982.

[24] H. Pottmann, B. Odehnal, M. Peternell, J. Wallner, and R. Ait Haddou. On optimal tolerancing in computer-aided design. In R. Martin and W. Wang, editors, *Geometric Modeling and Processing 2000*, pages 347–363. IEEE Computer Society, 2000.

[25] H. Pottmann and J. Wallner. *Computational Line Geometry*. Springer, 2001.

[26] A. A. G. Requicha. Towards a theory of geometric tolerancing. *Internat. J. Robotics Res.*, 2:45–60, 1983.

[27] Brigitte Servatius and Walter Whiteley. Constraining plane configurations in computer-aided design: Combinatorics of directions and lengths. *SIAM J. Discret. Math.*, 12:136–153, 1999.

[28] H. Stachel. Higher-order flexibility for a bipartite planar framework. In A. Kecskeméthy, M. Schneider, and C. Woernle, editors, *Advances in Multi-body Systems and Mechatronics*, pages 345–357. Inst. f. Mechanik und Getriebelehre der TU Graz, Duisburg, 1999.

[29] A. Verroust, F. Schonek, and D. Roller. Rule-oriented method for parametrized computeraided design. *Computer-Aided Design*, 25:531–540, 1993.

[30] J. Wallner, R. Krasauskas, and H. Pottmann. Error propagation in geometric constructions. *Computer-Aided Design*, 32:631–641, 2000.

[31] Walter Whiteley. Infinitesimal motions of a bipartite framework. *Pacific J. Math.*, 110:233–255, 1984.

[32] W. Wunderlich. Über Ausnahmefachwerke, deren Knoten auf einem Kegelschnitt liegen. *Acta Math.*, 47:291–300, 1983.

[33] Walter Wunderlich. *Ebene Kinematik*, volume 447/447a of *BI-Hochschultaschenbücher*. Bibliograph. Inst., Mannheim, 1970.