Chapter 9
Geometry of General Serial Robots

9.1 Introduction

Current serial robots, encountered not only in research laboratories but also in production or construction environments, include features that deserve a chapter apart. We will call here *general serial robots* all *non-redundant* serial robots that do not fall in the category of those studied in Chap. 4. Thus, the chapter is devoted to manipulators of the serial type that do not allow a decoupling of the positioning and the orientation problems. The focus of the chapter is, thus, the inverse displacement problem (IDP) of general six-revolute robots. While redundant manipulators of the serial type fall within this category as well, we will leave these aside, for their redundancy resolution calls for a more specialized background than what we have either assumed or given here.

A special feature of serial manipulators of the kind studied here is that they can admit up to sixteen inverse displacement solutions. Such manipulators are now in operation in industry, an example of which is the TELBOT System, shown in Fig. 9.1, which features all its six motors on its base, the motion and force transmission taking place via concentric tubes and bevel gears. This special feature allows TELBOT to have unlimited angular displacements at its joints, no cables traveling through its structure and no deadload on its links by virtue of the motors (Wälischmiller and Li 1996).

9.2 The IDP of General Six-Revolute Manipulators

As shown in Chap. 4, the IDP of six-revolute manipulators of the most general type leads to a system of six independent equations in six unknowns. This is a highly

---

Electronic supplementary material The online version of this article (doi: 10.1007/978-3-319-01851-5_9) contains supplementary material, which is available to authorized users.

J. Angeles, *Fundamentals of Robotic Mechanical Systems: Theory, Methods, and Algorithms*, Mechanical Engineering Series 124, DOI 10.1007/978-3-319-01851-5_9,

Fig. 9.1 The TELBOT System (courtesy of Wälischmiller GmbH, Meersburg, Germany)

nonlinear algebraic system whose solution posed a challenge to kinematicians for about two decades and that was not considered essentially solved until the late eighties. Below we give a short historical account of this problem.

Pieper (1968) reported what is probably the earliest attempt to formulate the inverse displacement problem of six-axis serial manipulators in a univariate polynomial form. He showed that decoupled manipulators, studied in Sect. 4.4, and a few others, allow a closed-form solution of the inverse displacement problem associated with them. However, apart from the simple architectures identified by Pieper, and others that have been identified more recently (Mavroidis and Roth 1992), a six-axis manipulator does not admit a closed-form solution. Attempts to derive the *minimal characteristic polynomial* for this manipulator were reported by Duffy and Derby (1979), Duffy and Crane (1980), Albala (1982), and Alizade et al. (1983), who derived a 32nd-degree polynomial, but suspected that this polynomial was not minimal, in the sense that the manipulator at hand might not be able to admit up to 32 postures for a given end-effector (EE) pose. Tsai and Morgan (1985) used a technique known as *polynomial continuation* (Morgan 1987) to solve *numerically* the nonlinear displacement equations, cast in the form of a system of quadratic equations. These researchers found that no more than 16 solutions were to be expected. Briefly stated, polynomial continuation consists basically of two stages, namely, reducing first the given problem to a system of polynomial equations; in the second stage, a continuous path, also known as a *homotopy* in mathematics, is defined with a real parameter t that can be regarded as time. The continuous path takes the system of equations from a given initial *state* to a final one. The initial state is so chosen that all solutions to the nonlinear system in this state are either apparent or much easier to find numerically than those of the originally proposed system. The final state of the system is the actual system to be solved. The initial system is thus deformed continuously into the final state upon varying its set of parameters, as t varies from 0 to 1. At each continuation step, a set of initial guesses for each of the solutions already exists, for it is simply the solution to the previous continuation step. Moreover, finding the solutions at the current continuation step is done using a standard Newton method (Dahlquist and Björck 1974).

Primrose (1986) proved conclusively that the problem under discussion admits at most 16 solutions, while Lee and Liang (1988) showed that the same problem leads to a 16th-degree univariate polynomial. Using different elimination procedures, as described in Sect. 9.3, Li[1] (1990) and Raghavan and Roth (1990, 1993) devised different procedures for the computation of the coefficients of the univariate polynomial. While the inverse displacement problem can be considered basically solved, research on finding all its solutions *safely* and *quickly* still continued into the nineties (Angeles et al. 1993). Below we describe two approaches to solving this problem: (a) the methods of Raghavan and Roth (1990, 1993) and of Li (1990), aimed at reducing the displacement relations to a single univariate polynomial; and (b) the bivariate-equation approach, introduced in (Angeles and Etemadi Zanganeh 1992).

It will become apparent, however, that a streamlined algorithm guaranteeing the reduction of the system of 14 *fundamental equations*, as derived in Sect. 9.2.2, to a lower number of equations in only one or two unknowns, is still lacking. A step in this direction is a method based on the concept of *kinematic mapping*, as reported by Husty et al. (2007). Within their method, the authors split the six-revolute kinematic chain into two three-revolute subchains, which allows the computation of the 16 inverse-displacement solutions using advanced geometric concepts. Once these solutions are available, the 16 possible values of a joint angle are known, the balance five joint angles are then computed by linear-equation solving, as in the case of the algorithms described here.

9.2.1 Preliminaries

We start by recalling a few definitions that were introduced in Chap. 4. In Sect. 4.2 we defined the matrices $\mathbf{Q}_i$ and the vectors $\mathbf{a}_i$ associated with the coordinate transformations from frame $\mathcal{F}_{i+1}$ to frame $\mathcal{F}_i$ or, equivalently, the displacement of the latter to the former. The 4×4 homogeneous matrix—see Sect. 2.5—transforming coordinates in $\mathcal{F}_{i+1}$ to coordinates in $\mathcal{F}_i$ is given by

$$\mathbf{A}_i = \begin{bmatrix} \mathbf{Q}_i & \mathbf{a}_i \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{9.1}$$

[1]N.B. Lee and Li of the references in this chapter are one and the same person, namely, Dr.-Ing. Hongyou Lee (a.k.a. Dr.-Ing. Hongyou Li).

where $\mathbf{0}$ is the three-dimensional zero vector, while the 3×3 rotation matrix $\mathbf{Q}_i$ and the three-dimensional vector $\mathbf{a}_i$ were defined in Chap. 4 as

$$\mathbf{Q}_i \equiv \begin{bmatrix} c_i & -\lambda_i s_i & \mu_i s_i \\ s_i & \lambda_i c_i & -\mu_i c_i \\ 0 & \mu_i & \lambda_i \end{bmatrix}, \qquad \mathbf{a}_i \equiv \begin{bmatrix} a_i c_i \\ a_i s_i \\ b_i \end{bmatrix} \tag{9.2}$$

In the above definitions we used the Denavit–Hartenberg notation, whereby a_i is the distance—and hence, $a_i \geq 0$—between the Z_i- and the Z_{i+1}-axes, while b_i is the offset—$-\infty < b_i < +\infty$—between the X_i- and X_{i+1}-axes, as measured along the positive direction of the Z_i-axis. Moreover,

$$c_i \equiv \cos\theta_i\,, \quad s_i \equiv \sin\theta_i\,, \quad \lambda_i \equiv \cos\alpha_i\,, \quad \mu_i \equiv \sin\alpha_i$$

where θ_i is the ith joint angle, measured from X_i to X_{i+1} in the positive direction of Z_i, and α_i denotes the twist angle from Z_i to Z_{i+1} in the positive direction of X_{i+1}, for $i = 1, \ldots, 6$. Furthermore, the factoring of matrix $\mathbf{Q}_i$, introduced in Eq. (4.2a), is reproduced below for quick reference:

$$\mathbf{Q}_i = \mathbf{Z}_i \mathbf{X}_i \tag{9.3}$$

with $\mathbf{X}_i$ and $\mathbf{Z}_i$ denoting two *pure reflections*, namely,

$$\mathbf{X}_i \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\lambda_i & \mu_i \\ 0 & \mu_i & \lambda_i \end{bmatrix}, \qquad \mathbf{Z}_i \equiv \begin{bmatrix} c_i & s_i & 0 \\ s_i & -c_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9.4a}$$

$$\mathbf{X}_i^T = \mathbf{X}_i = \mathbf{X}_i^{-1} \qquad \mathbf{Z}_i^T = \mathbf{Z}_i = \mathbf{Z}_i^{-1} \tag{9.4b}$$

the foregoing reflections thus being both symmetric and self-inverse—see Sect. 2.4. As a consequence,

$$\mathbf{Q}_i^T = \mathbf{X}_i \mathbf{Z}_i$$

We will also use the partitionings of $\mathbf{Q}_i$ displayed in Eq. (4.12), namely,

$$\mathbf{Q}_i \equiv \begin{bmatrix} \mathbf{p}_i & \mathbf{q}_i & \mathbf{u}_i \end{bmatrix} = \begin{bmatrix} \mathbf{m}_i^T \\ \mathbf{n}_i^T \\ \mathbf{o}_i^T \end{bmatrix} \tag{9.5}$$

A quick comparison between Eqs. (9.2) and (9.5) leads to the relations below:

$$\mathbf{m}_i = \begin{bmatrix} c_i \\ -\lambda_i s_i \\ \mu_i s_i \end{bmatrix}, \quad \mathbf{n}_i = \begin{bmatrix} s_i \\ \lambda_i c_i \\ -\mu_i c_i \end{bmatrix}, \quad \mathbf{o}_i = \begin{bmatrix} 0 \\ \mu_i \\ \lambda_i \end{bmatrix} \tag{9.6}$$

Further, let us recall the definition introduced in Eq. (4.13), $\mathbf{e} \equiv \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$, along with that in Eq. (4.3d), $\mathbf{a}_i = \mathbf{Q}_i \mathbf{b}_i$, which readily leads to $\mathbf{b}_i = \mathbf{Q}_i^T \mathbf{a}_i$. Hence,

$$\mathbf{u}_i \equiv \mathbf{Q}_i \mathbf{e} = \begin{bmatrix} \mu_i s_i \\ -\mu_i c_i \\ \lambda_i \end{bmatrix} \quad \text{and} \quad \mathbf{b}_i \equiv \begin{bmatrix} a_i \\ b_i \mu_i \\ b_i \lambda_i \end{bmatrix} \tag{9.7}$$

where we have reproduced Eq. (4.3e) for quick reference. Moreover, since $\mathbf{e}_i = [\mathbf{e}_i]_i = [\mathbf{e}_{i+1}]_{i+1}$, the above expression for $\mathbf{u}_i$ leads to

$$\mathbf{u}_i = \mathbf{Q}_i [\mathbf{e}_{i+1}]_{i+1} = [\mathbf{e}_{i+1}]_i \tag{9.8a}$$

which means that $\mathbf{u}_i$ represents $\mathbf{e}_{i+1}$ in $\mathcal{F}_i$. Likewise,

$$\mathbf{o}_i = \mathbf{Q}_i^T [\mathbf{e}_i]_i = [\mathbf{e}_i]_{i+1} \tag{9.8b}$$

Now, using Eqs. (9.4a) and the second of Eq. (9.7), we introduce the definitions

$$\boldsymbol{\gamma}_i \equiv \mathbf{Z}_i \mathbf{a}_i = \mathbf{X}_i \mathbf{b}_i = \begin{bmatrix} a_i & 0 & b_i \end{bmatrix}^T \tag{9.9}$$

whence

$$\mathbf{b}_i = \mathbf{X}_i \boldsymbol{\gamma}_i \tag{9.10}$$

Furthermore, vector $\mathbf{x}_i$ of Eq. (4.11) is reproduced below for quick reference as well:

$$\mathbf{x}_i \equiv \begin{bmatrix} \cos\theta_i \\ \sin\theta_i \end{bmatrix} \tag{9.11}$$

A useful concept in this context is that of *bilinear form*: An algebraic expression of the form Auv, where u and v are two given scalar variables and A is independent of u and v, is said to be *bilinear* in u and v. Likewise, an expression of the form Au^2v^2 is said to be *biquadratic* in u and v, with similar definitions for bicubic, trilinear, and multilinear forms. Moreover, the same definitions apply to vector and matrix expressions, as pertaining to their components and, correspondingly, their scalar entries.

In light of the definition of $\mathbf{x}_i$, additionally, we shall refer to an expression of the form

$$E_1 \equiv A\cos\theta_i + B\sin\theta_i + C \tag{9.12}$$

in which coefficients A, B and C are independent of θ_i, as being linear in $\mathbf{x}_i$. Likewise, an expression of the form

$$E_2 \equiv A\cos\theta_i\cos\theta_j + B\cos\theta_i\sin\theta_j + C\sin\theta_i\cos\theta_j + D\sin\theta_i\sin\theta_j + F \tag{9.13}$$

with coefficients A, B, ..., F independent of both θ_i and θ_j, will be termed bilinear in $\mathbf{x}_i$ and $\mathbf{x}_j$. In fact, such an expression may also involve terms linear in $\mathbf{x}_i$ and $\mathbf{x}_j$ alone. More generally, an expression involving terms with products such as $\cos^2\theta_i \cos^2\theta_j$ and other terms with similar products of the same or lower degree will be termed biquadratic in $\mathbf{x}_i$ and $\mathbf{x}_j$. Now we have

Lemma 9.2.1. *Let matrix* $\mathbf{A}$ *be skew-symmetric and* $\mathbf{B}$ *be defined as the similarity transformation of* $\mathbf{A}$ *given below:*

$$\mathbf{B} \equiv \mathbf{Q}_i \mathbf{A} \mathbf{Q}_i^T \tag{9.14}$$

where $\mathbf{Q}_i$ *was recalled in Eq. (9.2) and* $\mathbf{A}$ *is assumed to be independent of* θ_i. *Then,* $\mathbf{B}$ *is linear in* $\mathbf{x}_i$.

Proof. This result follows from relation (2.139). Indeed, as the reader can readily verify, $\mathbf{B}$ is skew-symmetric, and the product $\mathbf{Bv}$, for any three-dimensional vector $\mathbf{v}$, can be expressed in terms of $\mathbf{b}$, defined as vect($\mathbf{B}$)—see Sect. 2.3.3. That is,

$$\mathbf{Bv} = \mathbf{b} \times \mathbf{v}$$

If $\mathbf{a}$ denotes vect($\mathbf{A}$), then $\mathbf{a}$ and $\mathbf{b}$, by virtue of Eq. (9.14) and the results of Sect. 2.6, obey the relation

$$\mathbf{b} = \mathbf{Q}_i \mathbf{a}$$

Hence,

$$\mathbf{Bv} = (\mathbf{Q}_i \mathbf{a}) \times \mathbf{v}$$

thereby showing that the resulting product is linear in $\mathbf{x}_i$, q.e.d.

Moreover, let

$$\tau_i \equiv \tan\left(\frac{\theta_i}{2}\right) \tag{9.15a}$$

which allows us to write the identities below, as suggested by Li (1990):

$$s_i - \tau_i c_i \equiv \tau_i, \quad \tau_i s_i + c_i \equiv 1 \tag{9.15b}$$

We now define $\mathbf{p}$ as the vector directed from the origin of $\mathcal{F}_1$ to the operation point (OP) P of Fig. 9.2. Moreover, we let $\mathbf{l} \equiv [l_x, l_y, l_z]^T$, $\mathbf{m} \equiv [m_x, m_y, m_z]^T$, and $\mathbf{n} \equiv [n_x, n_y, n_z]^T$ represent the three mutually perpendicular unit vectors parallel to the X_7, Y_7 and Z_7 axes, respectively, of $\mathcal{F}_7$, which has its origin at P—a layout of these axes is depicted in Fig. 4.3 for a decoupled manipulator. Hence, the pose of the EE is described in the base frame $\mathcal{F}_1$ by means of the homogeneous transformation $\mathbf{A}$ given as

$$\mathbf{A} = \begin{bmatrix} \mathbf{Q} & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}, \qquad \mathbf{Q} \equiv \begin{bmatrix} \mathbf{l} & \mathbf{m} & \mathbf{n} \end{bmatrix} = \begin{bmatrix} l_x & m_x & n_x \\ l_y & m_y & n_y \\ l_z & m_z & n_z \end{bmatrix}$$

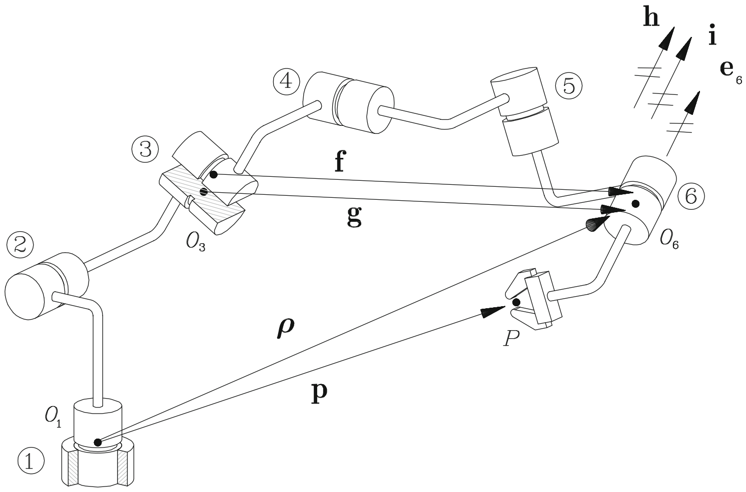

Fig. 9.2 Partitioning of the manipulator loop into two subloops

In the next step, we derive a set of scalar equations in five unknowns, upon eliminating one of these, that is *fundamental* in computing the solution of the problem at hand.

9.2.2 Derivation of the Fundamental Closure Equations

Given the geometric parameters of the manipulator and the pose of the EE with respect to the base frame, we derive the manipulator displacement equations, a.k.a. the *loop-closure equations*, from which all unknown angles are to be computed. We start by recalling the (matrix) rotation and (vector) translation equations of the general six-axis manipulator, as displayed in Eqs. (4.9a and b), and reproduced below for quick reference:

$$\mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_3\mathbf{Q}_4\mathbf{Q}_5\mathbf{Q}_6 = \mathbf{Q} \tag{9.16a}$$

$$\mathbf{a}_1 + \mathbf{Q}_1\mathbf{a}_2 + \mathbf{Q}_1\mathbf{Q}_2\mathbf{a}_3 + \ldots + \mathbf{Q}_1\mathbf{Q}_2\mathbf{Q}_3\mathbf{Q}_4\mathbf{Q}_5\mathbf{a}_6 = \mathbf{p} \tag{9.16b}$$

The use of 4×4 homogeneous transformations in the ensuing preparatory work will ease the suitable recasting of the foregoing equations. Thus, by using the matrices $\mathbf{A}_i$ of Eq. (9.1) in the above rotation and translation equations, we end up with a 4×4 matrix equation, namely,

$$\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\mathbf{A}_4\mathbf{A}_5\mathbf{A}_6 = \mathbf{A} \tag{9.17}$$

The unknown variables in the above equations are the joint angles $\{\theta_i\}_1^6$; the IDP thus consists in solving the closure equations (9.16a and b) or, equivalently, Eq. (9.17), for these unknowns. The said equations comprise 12 scalar equations and four identities; however, among these equations, only six are independent, for the columns (or the rows) of a rotation matrix must form an orthonormal—mutually orthogonal and of unit magnitude—set of vectors. The orthonormality property of the columns or rows of a rotation matrix, thus, brings about six scalar constraints.

The basic approach to solving the IDP resorts to *disassembling* the kinematic chain of the manipulator at two joints, e.g., joints 3 and 6, to obtain two subchains or subloops (Li et al. 1991). The first subchain, as suggested in the foregoing reference, and depicted in Fig. 9.2, goes from joint 3 to joint 6 via joints 4 and 5, while the second subchain goes from joint 6 to joint 3 via the EE and joints 1 and 2. Algebraically, this is equivalent to rewriting Eq. (9.17) in the form

$$\mathbf{A}_3\mathbf{A}_4\mathbf{A}_5 = \mathbf{A}_2^{-1}\mathbf{A}_1^{-1}\mathbf{A}\mathbf{A}_6^{-1} \tag{9.18a}$$

Note that each side of Eq. (9.18a) bears a specific structure. Indeed, if we denote by $\mathbf{L}_s$ and $\mathbf{R}_s$ the left- and right-hand sides of Eq. (9.18a), we have

$$\mathbf{L}_s \equiv \begin{bmatrix} l_{11}(\theta_3,\theta_4,\theta_5) & l_{12}(\theta_3,\theta_4,\theta_5) & l_{13}(\theta_3,\theta_4,\theta_5) & l_{14}(\theta_3,\theta_4,\theta_5) \\ l_{21}(\theta_3,\theta_4,\theta_5) & l_{22}(\theta_3,\theta_4,\theta_5) & l_{23}(\theta_3,\theta_4,\theta_5) & l_{24}(\theta_3,\theta_4,\theta_5) \\ l_{31}(\theta_4,\theta_5) & l_{32}(\theta_4,\theta_5) & l_{33}(\theta_4,\theta_5) & l_{34}(\theta_4,\theta_5) \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9.18b}$$

$$\mathbf{R}_s \equiv \begin{bmatrix} r_{11}(\theta_1,\theta_2,\theta_6) & r_{12}(\theta_1,\theta_2,\theta_6) & r_{13}(\theta_1,\theta_2) & r_{14}(\theta_1,\theta_2) \\ r_{21}(\theta_1,\theta_2,\theta_6) & r_{22}(\theta_1,\theta_2,\theta_6) & r_{23}(\theta_1,\theta_2) & r_{24}(\theta_1,\theta_2) \\ r_{31}(\theta_1,\theta_2,\theta_6) & r_{32}(\theta_1,\theta_2,\theta_6) & r_{33}(\theta_1,\theta_2) & r_{34}(\theta_1,\theta_2) \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9.18c}$$

where l_{ij} and r_{ij} denote nontrivial components of the left- and the right-hand sides, respectively, of Eq. (9.18a). Note that, because of the forms of matrices $\mathbf{Q}_i$, whose third rows are independent of θ_i, the third row of $\mathbf{L}_s$, as made apparent in Eq. (9.18b), is free of θ_3. Likewise, the third and fourth columns of $\mathbf{R}_s$, as made apparent in Eq. (9.18c), are free of θ_6.

It should be apparent that other pairs of joints can be used to disassemble the kinematic chain of the manipulator into two subchains; what matters is that none of the two subchains contains more than three joints; else, the entries of the homogeneous matrices become unnecessarily complex on one side of the matrix equation, while the entries of the other side become unnecessarily simple.

Now we extract one rotation and one translation equation from the 4×4 matrix equation (9.18a), namely,

$$\mathbf{Q}_3\mathbf{Q}_4\mathbf{Q}_5 = \mathbf{Q}_2^T\mathbf{Q}_1^T\mathbf{Q}\mathbf{Q}_6^T \tag{9.19a}$$

$$\mathbf{Q}_3(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5) = \mathbf{Q}_2^T\mathbf{Q}_1^T(\mathbf{p} - \mathbf{Q}\mathbf{b}_6) - (\mathbf{b}_2 + \mathbf{Q}_2^T\mathbf{b}_1) \tag{9.19b}$$

which are *kinematically equivalent* to Eqs. (9.16a and b), but *algebraically much simpler*. Note that, in Eq. (9.19b), we used the second Eq. (9.7) to substitute $\mathbf{a}_i$ by $\mathbf{Q}_i\mathbf{b}_i$. In the sequel, we will need the two products below:

$$\mathbf{Q}_5\mathbf{e} \equiv \mathbf{Q}_5[\mathbf{e}_6]_6 = [\mathbf{e}_6]_5 \tag{9.20a}$$

$$\mathbf{Q}_6^T\mathbf{e} \equiv \mathbf{Q}_6^T[\mathbf{e}_6]_6 = \mathbf{o}_6 = [\mathbf{e}_6]_7 \tag{9.20b}$$

where we have recalled relations (9.8a and b); whence,

$$\mathbf{Q}\mathbf{Q}_6^T\mathbf{e} \equiv \mathbf{Q}[\mathbf{e}_6]_7 = [\mathbf{e}_6]_1 \equiv \boldsymbol{\sigma}_6 \tag{9.20c}$$

Further, we equate the product of each of the two sides of Eq. (9.19a) by $\mathbf{e}$ from the right, to obtain, in light of Eqs. (9.20a and c),

$$\mathbf{Q}_3\mathbf{Q}_4[\mathbf{e}_6]_5 = \mathbf{Q}_2^T\mathbf{Q}_1^T[\mathbf{e}_6]_1 \tag{9.21a}$$

Both sides of Eq. (9.21a) thus represent the unit vector $\mathbf{e}_6$ in frame $\mathcal{F}_3$; the difference between the two sides should be apparent: while the left-hand side is obtained by transforming $[\mathbf{e}_6]_5$ into $[\mathbf{e}_6]_3$, the right-hand side by transforming $[\mathbf{e}_6]_1$ likewise. On the other hand, Eq. (9.19b) can be cast in the form

$$\mathbf{Q}_3(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5) = \mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\rho} - (\mathbf{b}_2 + \mathbf{Q}_2^T\mathbf{b}_1) \tag{9.21b}$$

where $\boldsymbol{\rho} \equiv \mathbf{p} - \mathbf{Q}\mathbf{b}_6 = [\mathbf{p} - \mathbf{a}_6]_1$. Hence, the left- and the right-hand sides of Eq. (9.21b) represent vector $\mathbf{a}_3 + \mathbf{a}_4 + \mathbf{a}_5$ in frame $\mathcal{F}_3$, the difference being that the left-hand side is obtained by carrying the $\mathcal{F}_4$-representation of the vector into $\mathcal{F}_3$, while the right-hand side does so from the $\mathcal{F}_1$-representation of the same vector.

Further, let the left- and the right-hand sides of Eq. (9.21a) be denoted by $\mathbf{h}$ and $\mathbf{i}$, respectively, while the counterparts of Eq. (9.21b) by $\mathbf{f}$ and $\mathbf{g}$, i.e.,

$$\mathbf{h} \equiv \mathbf{h}(\theta_3, \theta_4, \theta_5) = \mathbf{Q}_3\mathbf{Q}_4\mathbf{u}_5 \tag{9.22a}$$

$$\mathbf{i} \equiv \mathbf{i}(\theta_1, \theta_2) = \mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\sigma}_6 \tag{9.22b}$$

$$\mathbf{f} \equiv \mathbf{f}(\theta_3, \theta_4, \theta_5) = \mathbf{Q}_3(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5) \tag{9.22c}$$

$$\mathbf{g} \equiv \mathbf{g}(\theta_1, \theta_2) = \mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\rho} - (\mathbf{b}_2 + \mathbf{Q}_2^T\mathbf{b}_1) = \mathbf{Q}_2^T(\mathbf{Q}_1^T\boldsymbol{\rho} - \mathbf{b}_1) - \mathbf{b}_2 \tag{9.22d}$$

Further, notice that arrays $\mathbf{f}$ and $\mathbf{g}$ represent, in fact, the first three entries of the fourth columns of the matrices of Eqs. (9.18b) and (9.18c), respectively. Likewise, arrays $\mathbf{h}$ and $\mathbf{i}$ represent the third columns of the same matrices. Vectors $\mathbf{g}$ and $\mathbf{i}$ are thus free of θ_6.

Now, the six scalar equations (9.21b) and (9.21a) reduce, correspondingly, to

$$\mathbf{f} = \mathbf{g} \qquad \text{or} \qquad \begin{bmatrix} f_x(\theta_3, \theta_4, \theta_5) \\ f_y(\theta_3, \theta_4, \theta_5) \\ f_z(\theta_4, \theta_5) \end{bmatrix} = \begin{bmatrix} g_x(\theta_1, \theta_2) \\ g_y(\theta_1, \theta_2) \\ g_z(\theta_1, \theta_2) \end{bmatrix} \tag{9.23a}$$

$$\mathbf{h} = \mathbf{i} \qquad \text{or} \qquad \begin{bmatrix} h_x(\theta_3, \theta_4, \theta_5) \\ h_y(\theta_3, \theta_4, \theta_5) \\ h_z(\theta_4, \theta_5) \end{bmatrix} = \begin{bmatrix} i_x(\theta_1, \theta_2) \\ i_y(\theta_1, \theta_2) \\ i_z(\theta_1, \theta_2) \end{bmatrix} \tag{9.23b}$$

It should be noted that **h** and **i** are both unit vectors. Thus, each side of Eq. (9.23b) is subjected to a quadratic constraint, i.e.,

$$\mathbf{h} \cdot \mathbf{h} = 1, \quad \mathbf{i} \cdot \mathbf{i} = 1$$

and hence, out of the above six scalar equations, only five are independent. However, the number of unknowns in these six equations is also five. Therefore, Eqs. (9.23a) and (9.23b) suffice to determine the five unknown joint angles contained therein.

Although we already have one redundant equation to compute the six unknown angles, it will prove convenient to derive eight additional equations with the same power products[2] as **f**, **g**, **h** and **i**, namely,

$$\mathbf{f} \cdot \mathbf{f} = \mathbf{g} \cdot \mathbf{g} \tag{9.23c}$$

$$\mathbf{f} \cdot \mathbf{h} = \mathbf{g} \cdot \mathbf{i} \tag{9.23d}$$

$$\mathbf{f} \times \mathbf{h} = \mathbf{g} \times \mathbf{i} \tag{9.23e}$$

$$(\mathbf{f} \cdot \mathbf{f})\mathbf{h} - 2(\mathbf{f} \cdot \mathbf{h})\mathbf{f} = (\mathbf{g} \cdot \mathbf{g})\mathbf{i} - 2(\mathbf{g} \cdot \mathbf{i})\mathbf{g} \tag{9.23f}$$

It is noteworthy that Eq. (9.23f) is derived by first equating the *reflection*[3] of vector **h** onto a plane normal to **f** with its counterpart, the reflection of vector **i** onto a plane normal to **g**. The final form of Eq. (9.23f) is obtained upon clearing denominators in the foregoing reflection equation.

Equations (9.23a–9.23f) amount to 14 scalar equations in five unknown joint variables $\{\theta_i\}_1^5$. These are the *fundamental closure equations* sought. Some facts pertaining to the degree of the two sides of Eqs. (9.23c-f) are proven below:

Fact 9.2.1. *The inner products* $\mathbf{f} \cdot \mathbf{f}$ *and* $\mathbf{f} \cdot \mathbf{h}$ *are both free of* $\mathbf{x}_3$ *and bilinear in* $\{\, \mathbf{x}_i \,\}_4^5$*, while their counterparts* $\mathbf{g} \cdot \mathbf{g}$ *and* $\mathbf{g} \cdot \mathbf{i}$ *are bilinear in* $\mathbf{x}_1$ *and* $\mathbf{x}_2$.

---

[2]By *power product* we mean *terms* with their coefficients deleted; for example, the power products of the polynomial $5x^2y + 3xz + 9y^2 + 4z = 0$ are the terms x^2y, xz, y^2 and z.

[3]Neither Li nor Raghavan and Roth disclosed the geometric interpretation of this fourth equation, first proposed by Lee and Liang (1988).

Proof.

$$\begin{aligned}\mathbf{f}\cdot\mathbf{f} &\equiv \|\mathbf{Q}_3(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5)\|^2 \\ &\equiv \|\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5\|^2 \\ &\equiv \sum_3^5 \|\mathbf{b}_i\|^2 + 2\mathbf{b}_3^T\mathbf{Q}_4(\mathbf{b}_4 + \mathbf{Q}_5\mathbf{b}_5) + 2\mathbf{b}_4^T\mathbf{Q}_5\mathbf{b}_5\end{aligned}$$

whose rightmost-hand side is clearly free of $\mathbf{x}_3$ and is bilinear in $\{\,\mathbf{x}_i\,\}_4^5$. Similarly,

$$\begin{aligned}\mathbf{f}\cdot\mathbf{h} &\equiv (\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5)^T\mathbf{Q}_3^T\mathbf{Q}_3\mathbf{Q}_4\mathbf{u}_5 \\ &\equiv \mathbf{b}_3^T\mathbf{Q}_4\mathbf{u}_5 + \mathbf{b}_4^T\mathbf{u}_5 + \mathbf{b}_5^T\mathbf{Q}_5^T\mathbf{u}_5\end{aligned}$$

whose rightmost-hand side is apparently bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$, except for the last term, which contains two factors that are linear in $\mathbf{x}_5$, and hence, can be suspected to be quadratic. However, $\mathbf{Q}_5\mathbf{b}_5$ is, in fact, $\mathbf{a}_5$, while $\mathbf{u}_5$ is the last column of $\mathbf{Q}_5$, the suspicious term thus reducing to a constant, namely, $b_5\cos\alpha_5$. Similar proofs for $\mathbf{g}\cdot\mathbf{g}$ and $\mathbf{g}\cdot\mathbf{i}$ will be given presently. Moreover, □

Fact 9.2.2. *Vector* $\mathbf{f}\times\mathbf{h}$ *is trilinear in* $\{\,\mathbf{x}_i\,\}_3^5$*, while its counterpart,* $\mathbf{g}\times\mathbf{i}$*, is bilinear in* $\{\,\mathbf{x}_i\,\}_1^2$.

Proof. If we want the cross product of two vectors in frame $\mathcal{A}$ but have these vectors in frame $\mathcal{B}$, then we can proceed in two ways: either (a) transform each of the two vectors into $\mathcal{A}$-coordinates and perform the cross product of the two transformed vectors; or (b) perform the product of the two vectors in $\mathcal{B}$-coordinates and then transform the product vector into $\mathcal{A}$-coordinates. Obviously, the two products will be the same, which allows us to write

$$\begin{aligned}\mathbf{f}\times\mathbf{h} &\equiv \mathbf{Q}_3\left[\mathbf{b}_3\times(\mathbf{Q}_4\mathbf{u}_5) + (\mathbf{Q}_4\mathbf{b}_4)\times(\mathbf{Q}_4\mathbf{u}_5) + (\mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5)\times(\mathbf{Q}_4\mathbf{u}_5)\right] \\ &\equiv \mathbf{Q}_3\{\mathbf{b}_3\times(\mathbf{Q}_4\mathbf{u}_5) + \mathbf{Q}_4(\mathbf{b}_4\times\mathbf{u}_5) + \mathbf{Q}_4\left[(\mathbf{Q}_5\mathbf{b}_5)\times\mathbf{u}_5)\right]\}\end{aligned}$$

whose rightmost-hand side is apparently trilinear in $\{\,\mathbf{x}_i\,\}_3^5$, except for the term in brackets, which looks quadratic in $\mathbf{x}_5$. A quick calculation, however, reveals that this term is, in fact, linear in $\mathbf{x}_5$ as well. Indeed, from the definitions given in Eqs. (4.3c and d) and (9.5) we have

$$(\mathbf{Q}_5\mathbf{b}_5)\times\mathbf{u}_5 \equiv \mathbf{a}_5\times\mathbf{u}_5 \equiv \begin{bmatrix} a_5\lambda_5 s_5 + b_5\mu_5 c_5 \\ -a_5\lambda_5 c_5 + b_5\mu_5 s_5 \\ -a_5\mu_5\end{bmatrix}$$

which is obviously linear in $\mathbf{x}_5$. The proof for the counterpart product, $\mathbf{g}\times\mathbf{i}$, parallels the foregoing proof, and will be given below. □

Fact 9.2.3. *Vector* $(\mathbf{f}\cdot\mathbf{f})\mathbf{h} - 2(\mathbf{f}\cdot\mathbf{h})\mathbf{f}$ *is trilinear in* $\{\,\mathbf{x}_i\,\}_3^5$*, its counterpart,* $(\mathbf{g}\cdot\mathbf{g})\mathbf{i} - 2(\mathbf{g}\cdot\mathbf{i})\mathbf{g}$*, being bilinear in* $\{\,\mathbf{x}_i\,\}_1^2$.

Proof. First, we write the (elongated or contracted) reflection of vector $\mathbf{h}$ in the form

$$(\mathbf{f}\cdot\mathbf{f})\mathbf{h} - 2(\mathbf{f}\cdot\mathbf{h})\mathbf{f} \equiv \mathbf{Q}_3\mathbf{v}$$

where

$$\begin{aligned}
\mathbf{v} \equiv{}& \left(\sum_3^5 \|\mathbf{b}_i\|^2\right)\mathbf{Q}_4\mathbf{u}_5 - 2[(\mathbf{u}_5^T\mathbf{Q}_4\mathbf{b}_3)\mathbf{b}_3 + (\mathbf{u}_5^T\mathbf{b}_4)\mathbf{b}_3 + (\mathbf{u}_5^T\mathbf{b}_4)\mathbf{Q}_4\mathbf{b}_4\\
&+ (\mathbf{u}_5^T\mathbf{Q}_5\mathbf{b}_5)\mathbf{b}_3 + (\mathbf{u}_5^T\mathbf{Q}_5\mathbf{b}_5)\mathbf{Q}_4\mathbf{b}_4 + (\mathbf{u}_5^T\mathbf{Q}_5\mathbf{b}_5)\mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5] + 2\mathbf{w}\\
={}& \left(\sum_3^5 \|\mathbf{b}_i\|^2\right)\mathbf{Q}_4\mathbf{u}_5 - 2[\mathbf{u}_5^T\mathbf{b}_4(\mathbf{b}_3 + \mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4)\\
&+ \mathbf{u}_5^T\mathbf{Q}_5\mathbf{b}_5(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5) + 2\mathbf{w}
\end{aligned}$$

with all terms on the right-hand side, except for $\mathbf{w}$, which will be defined presently, clearly bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$. Vector $\mathbf{w}$ is defined as

$$\mathbf{w} \equiv [\ \ ]_1 + [\ \ ]_2 + [\ \ ]_3$$

each of the foregoing brackets being expanded below:

$$\begin{aligned}
[\ \ ]_1 &\equiv \left[(\mathbf{b}_3^T\mathbf{Q}_4\mathbf{b}_4)\mathbf{Q}_4\mathbf{u}_5 - (\mathbf{u}_5^T\mathbf{Q}_4^T\mathbf{b}_3)\mathbf{Q}_4\mathbf{b}_4\right]\\
&\equiv \mathbf{Q}_4(\mathbf{u}_5\mathbf{b}_4^T\mathbf{Q}_4^T - \mathbf{b}_4\mathbf{u}_5^T\mathbf{Q}_4^T)\mathbf{b}_3\\
&\equiv \mathbf{Q}_4(\mathbf{u}_5\mathbf{b}_4^T - \mathbf{b}_4\mathbf{u}_5^T)\mathbf{Q}_4^T\mathbf{b}_3
\end{aligned}$$

which thus reduces to a product including a factor of the form $\mathbf{Q}_i\mathbf{A}\mathbf{Q}_i^T$, with $\mathbf{A}$ being the term in parentheses in the rightmost-hand side of the last equation. This is obviously a skew-symmetric matrix, and Lemma 9.2.1 applies, i.e., the rightmost-hand side of the last equation is linear in $\mathbf{x}_4$. This term is, hence, bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$. Furthermore,

$$\begin{aligned}
[\ \ ]_2 &\equiv \left[(\mathbf{b}_4^T\mathbf{Q}_5\mathbf{b}_5)\mathbf{Q}_4\mathbf{u}_5 - (\mathbf{u}_5^T\mathbf{b}_4)\mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5\right]\\
&\equiv \mathbf{Q}_4\left[(\mathbf{b}_5^T\mathbf{Q}_5^T\mathbf{b}_4)\mathbf{u}_5 - (\mathbf{u}_5^T\mathbf{b}_4)\mathbf{Q}_5\mathbf{b}_5\right]\\
&\equiv \mathbf{Q}_4(\mathbf{u}_5\mathbf{b}_5^T\mathbf{Q}_5^T - \mathbf{Q}_5\mathbf{b}_5\mathbf{u}_5^T)\mathbf{b}_4
\end{aligned}$$

which is apparently linear in $\mathbf{x}_4$, but it is not obvious that it is also linear in $\mathbf{x}_5$. To show that the second linearity also holds, we can proceed in two ways. First, note that the term in parentheses is the skew-symmetric matrix $\mathbf{u}_5\mathbf{a}_5^T - \mathbf{a}_5\mathbf{u}_5^T$, whose vector, $\mathbf{a}_5 \times \mathbf{u}_5$, was already proven to be linear in $\mathbf{x}_5$. Since the vector of a skew-symmetric matrix fully defines that matrix—see Sect. 2.3—the linearity of the foregoing term in $\mathbf{x}_5$ follows immediately. Alternatively, we can expand the aforementioned difference, thereby deriving

$$\mathbf{u}_5\mathbf{a}_5^T - \mathbf{a}_5\mathbf{u}_5^T = \begin{bmatrix} 0 & a_5\mu_5 & -a_5\lambda_5 c_5 + b_5\mu_5 s_5 \\ -a_5\mu_5 & 0 & -a_5\lambda_5 s_5 - b_5\mu_5 c_5 \\ a_5\lambda_5 c_5 - b_5\mu_5 s_5 & a_5\lambda_5 s_5 + b_5\mu_5 c_5 & 0 \end{bmatrix}$$

which is clearly linear in $\mathbf{x}_5$. Moreover, its vector can be readily identified as $\mathbf{a}_5 \times \mathbf{u}_5$, as calculated above. Finally,

$$\begin{aligned} [\ ]_3 &\equiv \left[(\mathbf{b}_3^T\mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5)\mathbf{Q}_4\mathbf{u}_5 - (\mathbf{u}_5^T\mathbf{Q}_4^T\mathbf{b}_3)\mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5\right] \\ &\equiv \mathbf{Q}_4(\mathbf{u}_5\mathbf{b}_5^T\mathbf{Q}_5^T - \mathbf{Q}_5\mathbf{b}_5\mathbf{u}_5^T)\mathbf{Q}_4^T\mathbf{b}_3 \\ &\equiv \mathbf{Q}_4(\mathbf{u}_5\mathbf{a}_5^T - \mathbf{a}_5\mathbf{u}_5^T)\mathbf{Q}_4^T\mathbf{b}_3 \end{aligned}$$

this bracket thus reducing to a product including the factor $\mathbf{Q}_i\mathbf{A}\mathbf{Q}_i^T$, with $\mathbf{A}$ skew-symmetric. Hence, the foregoing expression is linear in $\mathbf{x}_4$, according to Lemma 9.2.1. Moreover, the matrix in parentheses was already proven to be linear in $\mathbf{x}_5$, thereby completing the proof for vector $(\mathbf{f}\cdot\mathbf{f})\mathbf{h} - 2(\mathbf{f}\cdot\mathbf{h})\mathbf{f}$. The proof for vector $(\mathbf{g}\cdot\mathbf{g})\mathbf{i} - 2(\mathbf{g}\cdot\mathbf{i})\mathbf{g}$ parallels the foregoing proof and will be given presently. □

Finally, we have one more useful result:

Fact 9.2.4. *If a scalar, vector, or matrix equation is linear in* $\mathbf{x}_i$*, then upon substitution of* c_i *and* s_i *by their equivalent forms in terms of* $\tau_i \equiv \tan(\theta_i/2)$*, the foregoing equation becomes quadratic in* τ_i *after clearing denominators.*

Proof. We shall show that this result holds for a scalar equation, with the extension to vector and matrix equations following directly. The scalar equation under discussion takes the general form

$$Ac_i + Bs_i + C = 0$$

where the coefficients A, B, and C do not contain θ_i. Upon substituting c_i and s_i in terms of $\tau_i \equiv \tan(\theta_i/2)$, and multiplying both sides of that equation by $1 + \tau_i^2$, we obtain

$$A(1 - \tau_i^2) + 2B\tau_i + C(1 + \tau_i^2) = 0$$

which is clearly quadratic in τ_i, q.e.d.

Moreover, if a scalar, vector, or matrix equation is of degree k in $\mathbf{x}_i$, upon introducing the same trigonometric substitution, the said equation becomes of degree $2k$ in τ_i.

Expressions for the right-hand sides of Eqs. (9.23c–d) are given below:

$$\mathbf{g}\cdot\mathbf{g} = \sum_1^2 \|\mathbf{b}_i\|^2 + \|\boldsymbol{\rho}\|^2 - 2\boldsymbol{\rho}^T\mathbf{Q}_1(\mathbf{Q}_2\mathbf{b}_2 + \mathbf{b}_1) + 2\mathbf{b}_1^T\mathbf{Q}_2\mathbf{b}_2 \tag{9.24a}$$

$$\mathbf{g}\cdot\mathbf{i} = \boldsymbol{\sigma}_6^T(\boldsymbol{\rho} - \mathbf{Q}_1\mathbf{Q}_2\mathbf{b}_2 - \mathbf{Q}_1\mathbf{b}_1) \tag{9.24b}$$

$$\mathbf{g}\times\mathbf{i} = \mathbf{Q}_2^T\mathbf{Q}_1^T(\boldsymbol{\rho}\times\boldsymbol{\sigma}_6) - \mathbf{b}_2\times\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\sigma}_6 - \mathbf{Q}_2^T(\mathbf{b}_1\times\mathbf{Q}_1^T\boldsymbol{\sigma}_6) \tag{9.24c}$$

and

$$\begin{aligned}(\mathbf{g}\cdot\mathbf{g})\mathbf{i} - 2(\mathbf{g}\cdot\mathbf{i})\mathbf{g} = {} & \left(\sum_1^2 \|\mathbf{b}_i\|^2 + \|\boldsymbol{\rho}\|^2\right)\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\sigma}_6 \\ & -2[(\boldsymbol{\sigma}_6^T\boldsymbol{\rho})(\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\rho} - \mathbf{b}_2 - \mathbf{Q}_2^T\mathbf{b}_1) + (\boldsymbol{\sigma}_6^T\mathbf{Q}_1\mathbf{Q}_2\mathbf{b}_2)\mathbf{b}_2 \\ & +(\boldsymbol{\sigma}_6^T\mathbf{Q}_1\mathbf{b}_1)\mathbf{b}_2 + (\boldsymbol{\sigma}_6^T\mathbf{Q}_1\mathbf{b}_1)\mathbf{Q}_2^T\mathbf{b}_2] + 2\mathbf{w}'\end{aligned} \tag{9.24d}$$

In deriving and simplifying the above relations, we use the invariance relations—see Sect. 2.7—of the dot and cross products, i.e., for any arbitrary vectors $\mathbf{u}$ and $\mathbf{v}$, we have

$$(\mathbf{Q}_i\mathbf{u})^T(\mathbf{Q}_i\mathbf{v}) = \mathbf{u}^T\mathbf{v}$$
$$(\mathbf{Q}_i\mathbf{u})\times(\mathbf{Q}_i\mathbf{v}) = \mathbf{Q}_i(\mathbf{u}\times\mathbf{v})$$

All the terms on the right-hand sides of Eqs. (9.24a–d), except for $\mathbf{w}'$, are apparently bilinear in $\mathbf{x}_1$ and $\mathbf{x}_2$. This bilinearity also holds for the last term in Eq. (9.24d), i.e., $\mathbf{w}'$, which can be expressed in the form

$$\mathbf{w}' \equiv [\quad]_1' + [\quad]_2' + [\quad]_3' \tag{9.25}$$

Each of the above brackets is given as

$$\begin{aligned}[\quad]_1' & \equiv [(\boldsymbol{\sigma}_6^T\mathbf{Q}_1\mathbf{Q}_2\mathbf{b}_2)\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\rho} - (\boldsymbol{\rho}^T\mathbf{Q}_1\mathbf{Q}_2\mathbf{b}_2)\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\sigma}_6] \\ & = (\mathbf{Q}_2^T\mathbf{Q}_1^T)(\boldsymbol{\rho}\boldsymbol{\sigma}_6^T - \boldsymbol{\sigma}_6\boldsymbol{\rho}^T)(\mathbf{Q}_2\mathbf{Q}_1)\mathbf{b}_2\end{aligned} \tag{9.26a}$$

$$\begin{aligned}[\quad]_2' & \equiv [(\mathbf{b}_1^T\mathbf{Q}_2\mathbf{b}_2)\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\sigma}_6 - (\boldsymbol{\sigma}_6^T\mathbf{Q}_1\mathbf{Q}_2\mathbf{b}_2)\mathbf{Q}_2^T\mathbf{b}_1] \\ & = \mathbf{Q}_2^T[(\mathbf{Q}_1^T\boldsymbol{\sigma}_6)\mathbf{b}_1 - \mathbf{b}_1(\mathbf{Q}_1^T\boldsymbol{\sigma}_6)^T]\mathbf{Q}_2\mathbf{b}_2\end{aligned} \tag{9.26b}$$

$$\begin{aligned}[\quad]_3' & \equiv [(\boldsymbol{\sigma}_6^T\mathbf{Q}_1\mathbf{b}_1)\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\rho} - (\boldsymbol{\rho}^T\mathbf{Q}_1\mathbf{b}_1)\mathbf{Q}_2^T\mathbf{Q}_1^T\boldsymbol{\sigma}_6] \\ & = \mathbf{Q}_2^T[\mathbf{Q}_1^T(\boldsymbol{\rho}\boldsymbol{\sigma}_6^T - \boldsymbol{\sigma}_6\boldsymbol{\rho}^T)\mathbf{Q}_1]\mathbf{b}_1\end{aligned} \tag{9.26c}$$

According to Lemma 9.2.1, the terms in the right-hand sides of relations (9.26a–c) are all bilinear in $\mathbf{x}_1$ and $\mathbf{x}_2$.

It is noteworthy that the third components of vectors $\mathbf{f} \times \mathbf{h}$ and $(\mathbf{f} \cdot \mathbf{f})\mathbf{h} - 2(\mathbf{f} \cdot \mathbf{h})\mathbf{f}$, as well as $\mathbf{f} \cdot \mathbf{f}$ and $\mathbf{f} \cdot \mathbf{h}$, are all free of θ_3. Hence, among the 14 scalar equations, i.e., Eqs. (9.23a–f), six are free of θ_3. Casting all 14 equations in vector form results in the *fundamental closure equations*:

$$\mathbf{P}\mathbf{x}_{45} = \mathbf{R}\mathbf{x}_{12} \tag{9.27}$$

where $\mathbf{P}$ and $\mathbf{R}$ are 14×9 and 14×8 matrices, respectively. Moreover, the entries of $\mathbf{P}$ are linear in $\mathbf{x}_3$, while those of $\mathbf{R}$ are independent of the joint angles. In addition, the nine- and eight-dimensional vectors $\mathbf{x}_{45}$ and $\mathbf{x}_{12}$ are defined as

$$\mathbf{x}_{45} \equiv \begin{bmatrix} s_4s_5 & s_4c_5 & c_4s_5 & c_4c_5 & s_4 & c_4 & s_5 & c_5 & 1 \end{bmatrix}^T \tag{9.28a}$$

$$\mathbf{x}_{12} \equiv \begin{bmatrix} s_1s_2 & s_1c_2 & c_1s_2 & c_1c_2 & s_1 & c_1 & s_2 & c_2 \end{bmatrix}^T \tag{9.28b}$$

Various approaches have been reported to solve the fundamental closure equations for the unknown joint angles, but all methods fall into two categories: (a) *purely numerical approaches*, whereby no attempt is made to reduce the number of unknowns (Angeles 1985), or the reduction is rather limited, from six to four unknowns (Tsai and Morgan 1985); and (b) *elimination approaches*, whereby unknowns are eliminated *algebraically*, as opposed to *numerically*, until a reduced number of equations in a reduced number of unknowns is derived.

We focus here only on the second category. Of these, we have essentially two classes: (a) the *univariate-polynomial* approach and (b) the *bivariate-equation* approach. As the names indicate, the former aims at reducing the fundamental equations to one single equation in one unknown. Moreover, that single equation, being polynomial in form, is termed the *characteristic polynomial* of the problem at hand. The polynomial is derived upon substituting the cosine and sine functions of the unknown angle, say θ_x, by $(1-T^2)/(1+T^2)$ and $2T/(1+T^2)$, respectively, with $T \equiv \tan(\theta_x/2)$. This transformation is well known as the *tan-half trigonometric identities*. The second approach, in turn, aims at reducing all fundamental closure equations to a smaller system of *trigonometric*, as opposed to *polynomial*, equations in only two unknowns.

The transformation of the original problem given in terms of trigonometric functions of the unknown angles into a polynomial equation in T is essential from a conceptual viewpoint, for this transformation makes apparent that the problem under study admits a finite number of solutions, namely, the degree of the characteristic polynomial. On the other hand, the same transformation is not trouble-free. Indeed, the mapping from θ_x into T apparently includes a singularity at $\theta_x = \pi$, whereby $T \to \infty$. The outcome is that, if one of the solutions is $\theta_x = \pi$, then the characteristic polynomial admits at least one solution at infinity, which is reflected in a *deflation* of the polynomial. This phenomenon, called *polynomial deflation*, was made apparent in Example 4.4.3, where a quartic characteristic polynomial appeared as cubic because of one solution at infinity. The beginner may thus be misled to believing that, in the presence of a solution at infinity, the system at hand admits a smaller number of solutions than it actually does. Furthermore, in the neighborhood of $\theta_x = \pi$, one of the solutions is extremely large in absolute value, which thus gives rise to numerical inaccuracies, generically referred to as ill-conditioning. As a matter of fact, the problem of polynomial-root finding has been identified as ill-conditioned by numerical analysts for some time (Forsythe 1970).

In order to cope with the foregoing shortcomings of the tan-half identities, the author and his team devised an alternative means, the bivariate-equation approach, to solving the problem at hand and other similar ones in computational kinematics (Angeles and Etemadi Zanganeh 1992a,b). In this approach, the 14 equations are reduced to a system of bivariate trigonometric equations in the sines and cosines of two of the unknown angles. These equations are then plotted in the plane of the two unknowns, thus obtaining four contours, whose intersections yield the real values of the two unknowns. As a matter of fact, only two such equations would suffice; however, it turns out that the underlying reduction cannot be accomplished without the introduction of either extra equations or spurious roots, which must be detected in order to discard them. Notice that, for an intersection point to qualify as a solution, *all contours* must meet at that point. As illustrated with one example, even the use of extra contours does not guarantee a legitimate solution. *Spurious* solutions fail to allow for the computation of the remaining four joint angles.

9.3 The Univariate-Polynomial Approach

We describe here two procedures leading to one single univariate 16th-degree polynomial equation, which is the characteristic polynomial of the system at hand. The two procedures bear many similarities, but they also involve remarkable differences that warrant separate discussions.

9.3.1 The Raghavan–Roth Procedure

A sophisticated elimination procedure was proposed by Raghavan and Roth (1990, 1993). Their procedure is based on Eqs. (9.23a–f), but their 14 closure equations are different, as explained below.

At the outset, Raghavan and Roth define four vectors that will play a key role in the ensuing derivations, namely,

$$\tilde{\mathbf{f}} \equiv \tilde{\mathbf{f}}(\theta_4, \theta_5) = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \mathbf{Z}_3\mathbf{f} = \mathbf{X}_3(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5) \tag{9.29a}$$

$$\tilde{\mathbf{h}} \equiv \tilde{\mathbf{h}}(\theta_1) = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = \mathbf{Q}_2\mathbf{g} + \mathbf{a}_2 = \mathbf{Q}_1^T\boldsymbol{\rho} - \mathbf{b}_1 \tag{9.29b}$$

Table 9.1 Expressions for the components of vectors $\tilde{\mathbf{f}}$, $\tilde{\mathbf{h}}$, $\tilde{\mathbf{r}}$, and $\tilde{\mathbf{n}}$

Item	Expression	Item	Expression
f_1	$c_4\iota_1 + s_4\iota_2 + a_3$	r_1	$c_4m_1 + s_4m_2$
f_2	$-\lambda_3(s_4\iota_1 - c_4\iota_2)$ $+ \mu_3\iota_3$	r_2	$-\lambda_3(s_4m_1 - c_4m_2)$ $+ \mu_3m_3$
f_3	$\mu_3(s_4\iota_1 - c_4\iota_2)$ $+ \lambda_3\iota_3 + b_3$	r_3	$\mu_3(s_4m_1 - c_4m_2)$ $+ \lambda_3m_3$
ι_1	$c_5a_5 + a_4$	m_1	$s_5\mu_5$
ι_2	$-s_5\lambda_4a_5 + \mu_4b_5$	m_2	$c_5\lambda_4\mu_5 + \mu_4\lambda_5$
ι_3	$s_5\mu_4a_5 + \lambda_4b_5 + b_4$	m_3	$-c_5\mu_4\mu_5 + \lambda_4\lambda_5$
h_1	$c_1p + s_1q - a_1$	n_1	$c_1u + s_1v$
h_2	$-\lambda_1(s_1p - c_1q)$ $+ \mu_1(r - b_1)$	n_2	$-\lambda_1(s_1u - c_1v)$ $+ \mu_1w$
h_3	$\mu_1(s_1p - c_1q)$ $+ \lambda_1(r - b_1)$	n_3	$\mu_1(s_1u - c_1v)$ $+ \lambda_1w$
p	$-l_xa_6 - (m_x\mu_6 + n_x\lambda_6)b_6$ $+ p_x$	u	$m_x\mu_6 + n_x\lambda_6$
q	$-l_ya_6 - (m_y\mu_6 + qn_y\lambda_6)b_6$ $+ p_y$	v	$m_y\mu_6 + n_y\lambda_6$
r	$-l_za_6 - (m_z\mu_6 + n_z\lambda_6)b_6$ $+ p_z$	w	$m_z\mu_6 + n_z\lambda_6$

$$\tilde{\mathbf{r}} \equiv \tilde{\mathbf{r}}(\theta_4, \theta_5) = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \mathbf{Z}_3\mathbf{h} = \mathbf{X}_3\mathbf{Q}_4\mathbf{u}_5 \tag{9.29c}$$

$$\tilde{\mathbf{n}} \equiv \tilde{\mathbf{n}}(\theta_1) = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} = \mathbf{Q}_2\mathbf{i} = \mathbf{Q}_1^T\boldsymbol{\sigma}_6 \tag{9.29d}$$

Expressions for the components of the above four vectors are given in Table 9.1, where ι_i $(i = 1, 2, 3)$, p, q, r, u, v, and w are auxiliary variables. Using Eqs. (9.29a–d), (9.3), (9.4b), (9.10), and (9.9), we can rewrite Eqs. (9.21a and b) in terms of the foregoing vectors, namely,

$$\mathbf{Z}_3\tilde{\mathbf{r}}(\theta_4, \theta_5) = \mathbf{X}_2\mathbf{Z}_2\tilde{\mathbf{n}}(\theta_1) \tag{9.30a}$$

$$\mathbf{Z}_3\tilde{\mathbf{f}}(\theta_4, \theta_5) = \mathbf{X}_2[\mathbf{Z}_2\tilde{\mathbf{h}}(\theta_1) - \boldsymbol{\gamma}_2] \tag{9.30b}$$

where we have recalled definitions (9.9) for $i = 2$. These six scalar equations play a key role in deriving the Raghavan–Roth equations in five unknowns that are needed to solve the problem at hand.

Next, both sides of Eqs. (9.30a and b) are multiplied from the left by $\mathbf{X}_2^{-1} \equiv \mathbf{X}_2^T \equiv \mathbf{X}_2$; then, the two equations thus resulting are rearranged in the forms

$$\mathbf{X}_2\mathbf{Z}_3\tilde{\mathbf{f}} + \boldsymbol{\gamma}_2 = \mathbf{Z}_2\tilde{\mathbf{h}} \tag{9.31}$$

$$\mathbf{X}_2\mathbf{Z}_3\tilde{\mathbf{r}} = \mathbf{Z}_2\tilde{\mathbf{n}} \tag{9.32}$$

Now, four new vectors, the counterparts of those introduced in Eq. (9.22a–d), are defined as

$$\bar{\mathbf{f}} \equiv \mathbf{X}_2\mathbf{Z}_3\tilde{\mathbf{f}} + \boldsymbol{\gamma}_2 = \mathbf{X}_2(\mathbf{Z}_3\tilde{\mathbf{f}} + \mathbf{b}_2) \tag{9.33a}$$

$$\bar{\mathbf{g}} \equiv \mathbf{Z}_2\tilde{\mathbf{h}} \tag{9.33b}$$

$$\bar{\mathbf{h}} \equiv \mathbf{X}_2\mathbf{Z}_3\tilde{\mathbf{r}} \tag{9.33c}$$

$$\bar{\mathbf{i}} \equiv \mathbf{Z}_2\tilde{\mathbf{n}} \tag{9.33d}$$

Note that $\bar{\mathbf{f}}$ and $\bar{\mathbf{h}}$ are trilinear[4] in $\mathbf{x}_3$, $\mathbf{x}_4$, and $\mathbf{x}_5$, while the first two components of $\bar{\mathbf{g}}$ and $\bar{\mathbf{i}}$ are bilinear in $\mathbf{x}_1$ and $\mathbf{x}_2$, their third components being linear in $\mathbf{x}_1$ and free of θ_2. Similar to Eqs. (9.23a and b), six scalar equations are obtained:

$$\bar{\mathbf{f}} = \bar{\mathbf{g}} \tag{9.34a}$$

$$\bar{\mathbf{h}} = \bar{\mathbf{i}} \tag{9.34b}$$

Moreover, eight more scalar equations are obtained in the forms

$$\bar{\mathbf{f}} \cdot \bar{\mathbf{f}} = \bar{\mathbf{g}} \cdot \bar{\mathbf{g}} \tag{9.34c}$$

$$\bar{\mathbf{f}} \cdot \bar{\mathbf{h}} = \bar{\mathbf{g}} \cdot \bar{\mathbf{i}} \tag{9.34d}$$

$$\bar{\mathbf{f}} \times \bar{\mathbf{h}} = \bar{\mathbf{g}} \times \bar{\mathbf{i}} \tag{9.34e}$$

$$(\bar{\mathbf{f}} \cdot \bar{\mathbf{f}})\bar{\mathbf{h}} - 2(\bar{\mathbf{f}} \cdot \bar{\mathbf{h}})\bar{\mathbf{f}} = (\bar{\mathbf{g}} \cdot \bar{\mathbf{g}})\bar{\mathbf{i}} - 2(\bar{\mathbf{g}} \cdot \bar{\mathbf{i}})\bar{\mathbf{g}} \tag{9.34f}$$

The fourteen scalar equations (9.34a–f) are henceforth termed the *Raghavan–Roth (RR) equations*.

The third components of the two vectors on the right-hand sides of Eqs. (9.34e and f), and the terms on the right-hand sides of Eqs. (9.34c and d) are free of θ_2 and linear in $\mathbf{x}_1$. As proven by Raghavan and Roth in the above references, the eight foregoing equations have the same power products as $\bar{\mathbf{f}}$, $\bar{\mathbf{h}}$, $\bar{\mathbf{g}}$, and $\bar{\mathbf{i}}$. Now, the 14 RR equations (9.34a–f) are cast in the form

$$\overline{\mathbf{P}}\mathbf{x}_{45} = \overline{\mathbf{R}}\mathbf{x}_{12} \tag{9.35}$$

where $\overline{\mathbf{P}}$ and $\overline{\mathbf{R}}$ are 14×9 and 14×8 matrices, respectively. Moreover, the entries of $\overline{\mathbf{P}}$ are linear in $\mathbf{x}_3$, while those of $\overline{\mathbf{R}}$ are independent of the joint angles; moreover, $\overline{\mathbf{R}}$ has the structure:

[4]while the last row of $\mathbf{Z}_3$ is free of θ_3, the last row of $\mathbf{X}_2\mathbf{Z}_3$ is $[\mu_2 s_3,\ -\mu_2 c_3,\ \lambda_2]$.

$$
\overline{\mathbf{R}} \equiv \begin{bmatrix}
\times & \times & \times & \times & 0 & 0 & \times & \times \\
\times & \times & \times & \times & 0 & 0 & \times & \times \\
0 & 0 & 0 & 0 & \times & \times & 0 & 0 \\
\times & \times & \times & \times & 0 & 0 & \times & 0 \\
\times & \times & \times & \times & 0 & 0 & 0 & \times \\
0 & 0 & 0 & 0 & \times & \times & 0 & 0 \\
0 & 0 & 0 & 0 & \times & \times & 0 & 0 \\
0 & 0 & 0 & 0 & \times & \times & 0 & 0 \\
\times & \times & \times & \times & 0 & 0 & \times & 0 \\
\times & \times & \times & \times & 0 & 0 & 0 & \times \\
0 & 0 & 0 & 0 & \times & \times & 0 & 0 \\
\times & \times & \times & \times & 0 & 0 & \times & \times \\
\times & \times & \times & \times & 0 & 0 & \times & \times \\
0 & 0 & 0 & 0 & \times & \times & 0 & 0
\end{bmatrix} \tag{9.36}
$$

In the above display, all nonzero entries are denoted by $\times$ and rows are written according to the order of appearance in Eqs. (9.34a–f). This special structure of matrix $\overline{\mathbf{R}}$ is then exploited to eliminate the joint angles θ_1 and θ_2 in an efficient way.

Based on the structure of $\overline{\mathbf{R}}$, two groups of six and eight equations are defined:

$$\overline{\mathbf{P}}_u \mathbf{x}_{45} = \mathbf{C}\mathbf{x}_1 \tag{9.37a}$$

$$\overline{\mathbf{P}}_l \mathbf{x}_{45} = \mathbf{A}\tilde{\mathbf{x}}_{12} \tag{9.37b}$$

where $\mathbf{C}$ is a 6×2 constant matrix that is formed by the nonzero entries in rows 3, 6, 7, 8, 11, and 14 of matrix $\overline{\mathbf{R}}$. $\mathbf{A}$ is, in turn, an 8×6 matrix whose entries are all functions of the data, while $\mathbf{x}_1$ and $\mathbf{x}_{45}$ were defined in Eqs. (9.11) and (9.28a), respectively; $\tilde{\mathbf{x}}_{12}$ is, in turn, the six-dimensional vector defined as

$$\tilde{\mathbf{x}}_{12} \equiv \begin{bmatrix} s_1 s_2 & s_1 c_2 & c_1 s_2 & c_1 c_2 & s_2 & c_2 \end{bmatrix}^T \tag{9.38}$$

Furthermore, $\overline{\mathbf{P}}_u$ comprises the third, sixth, seventh, eighth, 11th and 14th rows of $\overline{\mathbf{P}}$, $\overline{\mathbf{P}}_l$ comprising the remaining eight rows. Notice that $\overline{\mathbf{P}}_u$ and $\overline{\mathbf{P}}_l$ are both linear in $\mathbf{x}_3$.

Any two of the six scalar equations in Eq. (9.37a) can now be used to solve for $\mathbf{x}_1$, the resulting expression then being substituted into the remaining four equations of the same group. This is done by first partitioning the six scalar equations as

$$\mathbf{C}_u \mathbf{x}_1 = \mathbf{d}_u \tag{9.39a}$$

$$\mathbf{C}_l \mathbf{x}_1 = \mathbf{d}_l \tag{9.39b}$$

where $\mathbf{C}_u$ and $\mathbf{C}_l$ are 2×2 and 4×2 submatrices of $\mathbf{C}$, respectively, with $\mathbf{d}_u$ and $\mathbf{d}_l$ being the corresponding two- and four-dimensional vectors that result from $\overline{\mathbf{P}}_u \mathbf{x}_{45}$;

these two vectors are trilinear in $\mathbf{x}_3$, $\mathbf{x}_4$ and $\mathbf{x}_5$. If Eq. (9.39a) is solved for $\mathbf{x}_1$ and the result is substituted into Eq. (9.39b), we obtain four equations free of θ_1 and θ_2, namely,

$$\overline{\mathbf{\Gamma}}_4\mathbf{x}_{45} \equiv \mathbf{C}_l\mathbf{C}_u^{-1}\mathbf{d}_u - \mathbf{d}_l = \mathbf{0}_4\,, \qquad \overline{\mathbf{\Gamma}}_4 \equiv \mathbf{C}_l\mathbf{C}_u^{-1}(\overline{\mathbf{P}}_u)_2 - (\overline{\mathbf{P}}_u)_4 \tag{9.40a}$$

in which $\overline{\mathbf{\Gamma}}_4$ is a 4×9 matrix whose entries are linear in $\mathbf{x}_3$, while $(\overline{\mathbf{P}}_u)_2$ and $(\overline{\mathbf{P}}_u)_4$ are 2×9 and 4×9 submatrices of matrix $\overline{\mathbf{P}}_u$, respectively. The above set of equations is now cast in the form

$$\mathbf{D}_1\mathbf{y}_3 = \mathbf{0}_4 \tag{9.40b}$$

with $\mathbf{D}_1$ defined as a 4×3 matrix whose entries are bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$, while $\mathbf{0}_4$ is the four-dimensional zero vector, and $\mathbf{y}_3$ is defined as

$$\mathbf{y}_3 \equiv \begin{bmatrix} c_3 & s_3 & 1 \end{bmatrix}^T \tag{9.41}$$

If $\mathbf{C}_u$ is chosen with nonzero entries in the third and sixth rows of matrix $\overline{\mathbf{R}}$, then we have

$$\mathbf{C}_u = \begin{bmatrix} \mu_1 p & -\mu_1 q \\ \mu_1 u & -\mu_1 v \end{bmatrix} \tag{9.42a}$$

with p, q, u, and v listed in Table 9.1. If $\mathbf{C}_u$ is nonsingular, $\mathbf{C}_u^{-1}$ is readily obtained as

$$\mathbf{C}_u^{-1} = \frac{1}{\mu_1(uq - pv)} \begin{bmatrix} -v & q \\ -u & p \end{bmatrix} \tag{9.42b}$$

However, if $\mathbf{C}_u$ turns out to be singular, then a different pair of Eqs. (9.37a), of the set associated with rows 3, 6, 7, 8, 11 and 14, should be selected.

Additional equations free of θ_1 and θ_2 can be derived from any six of the eight equations in Eq. (9.37b). Indeed, these six equations is all that is needed to solve for $\tilde{\mathbf{x}}_{12}$ in terms of θ_3, θ_4 and θ_5; the expressions thus resulting would then be substituted into the remaining two equations of the same set, to obtain two additional equations free of θ_1 and θ_2. However, this elimination process is not suitable for symbolic computations. Instead, Raghavan and Roth (1990) derived the two additional equations in a terser form. This is done by finding two independent linear combinations of the eight equations (9.37b) that render identically zero all terms in θ_1 and θ_2. The left-hand sides of these equations are given as

$$\begin{aligned}\phi_1(\theta_3,\theta_4,\theta_5) \equiv{}& \frac{\mu_1^2}{2a_1}[(\bar{\mathbf{f}}\cdot\bar{\mathbf{f}})\overline{h}_x - 2(\bar{\mathbf{f}}\cdot\overline{\mathbf{h}})\overline{f}_x] - \frac{\mu_1^2}{2a_1}\delta_1\overline{h}_x + \frac{\mu_1^2}{a_1}\delta_2\overline{f}_x \\ &-\lambda_1\mu_1(\bar{\mathbf{f}}\times\overline{\mathbf{h}})_x + \mu_1 w\overline{f}_y - \mu_1(r - b_1)\overline{h}_y\end{aligned} \tag{9.43a}$$

$$\phi_2(\theta_3, \theta_4, \theta_5) \equiv \frac{\mu_1^2}{2a_1}[(\bar{\mathbf{f}} \cdot \bar{\mathbf{f}})\overline{h}_y - 2(\bar{\mathbf{f}} \cdot \bar{\mathbf{h}})\overline{f}_y] - \lambda_1\mu_1(\bar{\mathbf{f}} \times \bar{\mathbf{h}})_y$$
$$-\mu_1 w\overline{f}_x + \mu_1(r - b_1)\overline{h}_x + \frac{\mu_1^2}{a_1}\delta_2\overline{f}_y - \frac{\mu_1^2}{2a_1}\delta_1\overline{h}_y \quad (9.43b)$$

while the right-hand sides are

$$\psi_1 \equiv \frac{\mu_1^2}{2a_1}[(\overline{\mathbf{g}} \cdot \overline{\mathbf{g}})\bar{\imath}_x - 2(\overline{\mathbf{g}} \cdot \bar{\mathbf{i}})\overline{g}_x] - \frac{\mu_1^2}{2a_1}\delta_1\bar{\imath}_x + \frac{\mu_1^2}{a_1}\delta_2\overline{g}_x$$
$$-\lambda_1\mu_1(\overline{\mathbf{g}} \times \bar{\mathbf{i}})_x + \mu_1 w\overline{g}_y - \mu_1(r - b_1)\bar{\imath}_y \quad (9.43c)$$

$$\psi_2 \equiv \frac{\mu_1^2}{2a_1}[(\overline{\mathbf{g}} \cdot \overline{\mathbf{g}})\bar{\imath}_y - 2(\overline{\mathbf{g}} \cdot \bar{\mathbf{i}})\overline{g}_y] - \lambda_1\mu_1(\overline{\mathbf{g}} \times \bar{\mathbf{i}})_y$$
$$-\mu_1 w\overline{g}_x + \mu_1(r - b_1)\bar{\imath}_x + \frac{\mu_1^2}{a_1}\delta_2\overline{g}_y - \frac{\mu_1^2}{2a_1}\delta_1\bar{\imath}_y \quad (9.43d)$$

On the other hand, $\overline{h}_x$, $\bar{\imath}_x$, $\overline{f}_x$ and $\overline{g}_x$ represent the first components of vectors $\bar{\mathbf{h}}$, $\bar{\mathbf{i}}$, $\bar{\mathbf{f}}$, and $\overline{\mathbf{g}}$, respectively, the other components being defined likewise. Furthermore, δ_1 and δ_2 are defined as

$$\delta_1 \equiv p^2 + q^2 + (r - b_1)^2 - a_1^2$$
$$\delta_2 \equiv pu + qv + (r - b_1)w$$

Upon substitution of $\overline{\mathbf{g}}$ and $\bar{\mathbf{i}}$, as given by Eqs. (9.33b and d), respectively, into Eqs. (9.43c and d), and introduction of the definitions given in Table 9.1, it turns out that both ψ_1 and ψ_2 vanish identically, i.e.,

$$\psi_1 = 0 \quad \text{and} \quad \psi_2 = 0$$

Also note that, in deriving expressions (9.43a and b) and (9.43c and d), we assume that $a_1 \neq 0$. However, a_1 vanishes in many industrial robots, those having their first two axes intersecting—usually at right angles—the foregoing procedure thus becoming inapplicable. One way of coping with this case is to go one step behind Raghavan and Roth's procedure and redefine, for $k = 1, 2$,

$$\phi_k(\theta_3, \theta_4, \theta_5) \longleftarrow a_1\phi_k(\theta_3, \theta_4, \theta_5);$$

and

$$\psi_k \longleftarrow a_1\psi_k$$

i.e.,

$$\phi_1(\theta_3,\theta_4,\theta_5) \equiv \frac{\mu_1^2}{2}[(\bar{\mathbf{f}}\cdot\bar{\mathbf{f}})\overline{h}_x - 2(\bar{\mathbf{f}}\cdot\bar{\mathbf{h}})\overline{f}_x] - \frac{\mu_1^2}{2}\delta_1\overline{h}_x + \mu_1^2\delta_2\overline{f}_x$$
$$-a_1\lambda_1\mu_1(\bar{\mathbf{f}}\times\bar{\mathbf{h}})_x + a_1\mu_1 w\overline{f}_y - a_1\mu_1(r-b_1)\overline{h}_y \tag{9.44a}$$

$$\phi_2(\theta_3,\theta_4,\theta_5) \equiv \frac{\mu_1^2}{2}[(\bar{\mathbf{f}}\cdot\bar{\mathbf{f}})\overline{h}_y - 2(\bar{\mathbf{f}}\cdot\bar{\mathbf{h}})\overline{f}_y] - a_1\lambda_1\mu_1(\bar{\mathbf{f}}\times\bar{\mathbf{h}})_y - a_1\mu_1 w\overline{f}_x$$
$$+a_1\mu_1(r-b_1)\overline{h}_x + \mu_1^2\delta_2\overline{f}_y - \frac{\mu_1^2}{2}\delta_1\overline{h}_y \tag{9.44b}$$

$$\psi_1 \equiv \frac{\mu_1^2}{2}[(\bar{\mathbf{g}}\cdot\bar{\mathbf{g}})\bar{\imath}_x - 2(\bar{\mathbf{g}}\cdot\bar{\mathbf{i}})\bar{g}_x] - \frac{\mu_1^2}{2}\delta_1\bar{\imath}_x + \mu_1^2\delta_2\bar{g}_x$$
$$-a_1\lambda_1\mu_1(\bar{\mathbf{g}}\times\bar{\mathbf{i}})_x + a_1\mu_1 w\bar{g}_y - a_1\mu_1(r-b_1)\bar{\imath}_y \tag{9.45a}$$

$$\psi_2 \equiv \frac{\mu_1^2}{2}[(\bar{\mathbf{g}}\cdot\bar{\mathbf{g}})\bar{\imath}_y - 2(\bar{\mathbf{g}}\cdot\bar{\mathbf{i}})\bar{g}_y] - a_1\lambda_1\mu_1(\bar{\mathbf{g}}\times\bar{\mathbf{i}})_y$$
$$-a_1\mu_1 w\bar{g}_x + a_1\mu_1(r-b_1)\bar{\imath}_x + \mu_1^2\delta_2\bar{g}_y - \frac{\mu_1^2}{2}\delta_1\bar{\imath}_y \tag{9.45b}$$

Under their new definitions, apparently, ψ_1 and ψ_2 also vanish. Once ϕ_1 and ϕ_2 are equated to zero, two equations are obtained that can be cast in the form

$$\overline{\boldsymbol{\Gamma}}_2\mathbf{x}_{45} = \mathbf{0}_2 \tag{9.46}$$

or equivalently,

$$\mathbf{D}_2\mathbf{y}_3 = \mathbf{0}_2 \tag{9.47}$$

where $\mathbf{0}_2$ is the two-dimensional zero vector, $\overline{\boldsymbol{\Gamma}}_2$ is a 2×9 matrix whose entries are linear in $\mathbf{x}_3$, $\mathbf{D}_2$ is a 2×3 matrix whose entries are bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$, and $\mathbf{y}_3$ was introduced in Eq. (9.41).

The two Eqs. (9.40a) and (9.46) thus involve a total of six scalar equations free of θ_1 and θ_2, and can be combined to yield a system of six equations trilinear in $\mathbf{x}_3$, $\mathbf{x}_4$, and $\mathbf{x}_5$, namely,

$$\boldsymbol{\Sigma}\mathbf{x}_{45} = \mathbf{0}_6 \tag{9.48a}$$

where $\boldsymbol{\Sigma}$ is a 6×9 matrix whose entries are linear in $\mathbf{x}_3$, and $\mathbf{0}_6$ is the six-dimensional zero vector. Now, the tan-half trigonometric identities relating s_i and c_i with $\tau_i \equiv \tan(\theta_i/2)$, for $i = 4, 5$, are substituted into Eq. (9.48a). Upon multiplying the two sides of the equation thus resulting by $(1+\tau_4^2)(1+\tau_5^2)$, Raghavan and Roth obtained

$$\boldsymbol{\Sigma}'\mathbf{x}'_{45} = \mathbf{0}_6 \tag{9.48b}$$

where $\boldsymbol{\Sigma}'$ is a 6×9 matrix that is linear in $\mathbf{x}_3$, while $\mathbf{x}'_{45}$ is defined as

$$\mathbf{x}'_{45} \equiv \left[\tau_4^2\tau_5^2 \ \tau_4^2\tau_5 \ \tau_4^2 \ \tau_4\tau_5^2 \ \tau_4\tau_5 \ \tau_4 \ \tau_5^2 \ \tau_5 \ 1\right]^T$$

If the same trigonometric identities, for $i = 3$, are now substituted into Eq. (9.48b), and then the first four scalar equations of this set are multiplied by $(1 + \tau_3^2)$ to clear denominators, the equation thus resulting takes the form

$$\boldsymbol{\Sigma}''\mathbf{x}'_{45} = \mathbf{0}_6 \tag{9.48c}$$

In the above equations, $\boldsymbol{\Sigma}''$ is a 6×9 matrix whose first four rows are quadratic in τ_3, while its last two rows are apparently rational functions of τ_3. However, as reported by Raghavan and Roth, the determinant of any 6×6 submatrix of $\boldsymbol{\Sigma}''$ is, in fact, an 8th-degree polynomial in τ_3 and not a rational function of the same. Moreover, in order to eliminate τ_4 and τ_5, they resort to dialytic elimination (Salmon 1964), introduced in this book in Sect. 5.4.1 and in Exercise 5.11. Dialytic elimination is further discussed in Sect. 9.3, in connection with the Li, Woernle, and Hiller method, and in Sect. 10.2 in connection with parallel manipulators.

In applying dialytic elimination, the two sides of the system of equations appearing in Eq. (9.48c) are first multiplied by τ_4; then, the system of equations thus obtained is adjoined to the original system, thereby deriving a system of 12 linear homogeneous equations in $\tilde{\mathbf{x}}_{45}$, namely,

$$\mathbf{S}\tilde{\mathbf{x}}_{45} = \mathbf{0}_{12} \tag{9.48d}$$

where $\mathbf{0}_{12}$ is the 12-dimensional zero vector, while the 12-dimensional vector $\tilde{\mathbf{x}}_{45}$ is defined as

$$\begin{aligned}\tilde{\mathbf{x}}_{45} \equiv [\, & \tau_4^3\tau_5^2 \ \tau_4^3\tau_5 \ \ \tau_4^3 \ \ \tau_4^2\tau_5^2 \ \tau_4^2\tau_5 \ \tau_4^2 \\ & \tau_4\tau_5^2 \ \tau_4\tau_5 \ \ \tau_4 \ \ \ \tau_5^2 \ \ \tau_5 \ 1\,]^T\end{aligned} \tag{9.48e}$$

Furthermore, the 12×12 matrix $\mathbf{S}$ is defined as

$$\mathbf{S} \equiv \begin{bmatrix}\mathbf{G}\\ \mathbf{K}\end{bmatrix}$$

its 6×12 blocks $\mathbf{G}$ and $\mathbf{K}$ taking on the forms

$$\mathbf{G} \equiv \left[\boldsymbol{\Sigma}'' \ \mathbf{O}_{63}\right], \qquad \mathbf{K} \equiv \left[\mathbf{O}_{63} \ \boldsymbol{\Sigma}''\right]$$

with $\mathbf{O}_{63}$ defined as the 6×3 zero matrix.

Now, in order for Eq. (9.48d) to admit a nontrivial solution, the determinant of its coefficient matrix must vanish, i.e.,

$$\det(\mathbf{S}) = 0 \tag{9.49}$$

which is the characteristic equation sought. The foregoing determinant turns out to be a 16th-degree polynomial in τ_3. Moreover, the roots of this polynomial give the values of τ_3 corresponding to the 16 solutions of the IDP. It should be noted that, using the same procedure, one can also derive this polynomial in terms of either τ_4 or τ_5 if the associated vector in Eq. (9.48d) is written as $\mathbf{x}_{35}$ or $\mathbf{x}_{34}$, respectively. Consequently, the entries of matrix $\mathbf{\Sigma}$ would be linear in either $\mathbf{x}_4$ or $\mathbf{x}_5$.

9.3.2 The Li–Woernle–Hiller Procedure

At the outset, the factoring of $\mathbf{Q}_i$ given in Eq. (4.1c) and the identities first used by Li (1990), namely, Eqs. (9.15b), are recalled. Additionally, Li defines a matrix $\mathbf{T}_i$ as

$$\mathbf{T}_i \equiv \begin{bmatrix} -\tau_i & 1 & 0 \\ 1 & \tau_i & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Hence,

$$\mathbf{T}_i\mathbf{C}_i \equiv \mathbf{U}_i = \begin{bmatrix} \tau_i & 1 & 0 \\ 1 & -\tau_i & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with $\mathbf{C}_i$ defined in Eq. (4.1b). Furthermore, we note that the left-hand sides of the four vector equations (9.23a, b, e and f) are of the form $\mathbf{Q}_3\mathbf{v}$, where $\mathbf{v}$ is a three-dimensional vector independent of θ_3. Upon multiplication of the above-mentioned equations from the left by matrix $\mathbf{T}_3$, Li and co-authors obtained a new set of equations, namely,

$$\mathbf{U}_3\hat{\mathbf{f}} = \mathbf{T}_3\mathbf{g} \tag{9.50a}$$

$$\mathbf{U}_3\hat{\mathbf{r}} = \mathbf{T}_3\mathbf{i} \tag{9.50b}$$

$$\mathbf{U}_3(\hat{\mathbf{f}} \times \hat{\mathbf{r}}) = \mathbf{T}_3(\mathbf{g} \times \mathbf{i}) \tag{9.50c}$$

$$\mathbf{U}_3\left[(\mathbf{f}\cdot\mathbf{f})\hat{\mathbf{r}} - 2(\mathbf{f}\cdot\mathbf{h})\hat{\mathbf{f}}\right] = \mathbf{T}_3\left[(\mathbf{g}\cdot\mathbf{g})\mathbf{i} - 2(\mathbf{g}\cdot\mathbf{i})\mathbf{g}\right] \tag{9.50d}$$

where $\hat{\mathbf{f}}$ and $\hat{\mathbf{r}}$ are defined as

$$\hat{\mathbf{f}} \equiv \mathbf{\Lambda}_3(\mathbf{b}_3 + \mathbf{Q}_4\mathbf{b}_4 + \mathbf{Q}_4\mathbf{Q}_5\mathbf{b}_5) \tag{9.51}$$

$$\hat{\mathbf{r}} \equiv \mathbf{\Lambda}_3(\mathbf{Q}_4\mathbf{u}_5) \tag{9.52}$$

with $\mathbf{\Lambda}_i$ defined, in turn, in Eq. (4.1c).

Because of the form of matrices $\mathbf{T}_3$ and $\mathbf{U}_3$, the third of each of the four vector equations (9.50a–d) is identical to its counterpart appearing in Eqs. (9.34a, b, e and f). That is, if we denote by either v_i or $(\mathbf{v})_i$ the ith component of any three-dimensional vector $\mathbf{v}$, the unchanged equations are

$$\hat{f}_3 = g_3 \tag{9.53a}$$

$$\hat{r}_3 = i_3 \tag{9.53b}$$

$$(\hat{\mathbf{f}} \times \hat{\mathbf{r}})_3 = (\mathbf{g} \times \mathbf{i})_3 \tag{9.53c}$$

$$(\mathbf{f} \cdot \mathbf{f})\hat{r}_3 - 2(\mathbf{f} \cdot \mathbf{h})\hat{f}_3 = (\mathbf{g} \cdot \mathbf{g})i_3 - 2(\mathbf{g} \cdot \mathbf{i})g_3 \tag{9.53d}$$

all of which are free of θ_3. Furthermore, six additional equations linear in τ_3 will be derived by multiplying both sides of Eqs. (9.53a–d) and of (9.23c and d) by τ_3, i.e.,

$$\tau_3\hat{f}_3 = \tau_3 g_3 \tag{9.54a}$$

$$\tau_3\hat{r}_3 = \tau_3 i_3 \tag{9.54b}$$

$$\tau_3(\hat{\mathbf{f}} \times \hat{\mathbf{r}})_3 = \tau_3(\mathbf{g} \times \mathbf{i})_3 \tag{9.54c}$$

$$\tau_3[(\mathbf{f} \cdot \mathbf{f})\hat{r}_3 - 2(\mathbf{f} \cdot \mathbf{h})\hat{f}_3]_3 = \tau_3[(\mathbf{g} \cdot \mathbf{g})i_3 - 2(\mathbf{g} \cdot \mathbf{i})g_3]_3 \tag{9.54d}$$

$$\tau_3(\mathbf{f} \cdot \mathbf{f}) = \tau_3(\mathbf{g} \cdot \mathbf{g}) \tag{9.54e}$$

$$\tau_3(\mathbf{f} \cdot \mathbf{h}) = \tau_3(\mathbf{g} \cdot \mathbf{i}) \tag{9.54f}$$

We now have 20 scalar equations that are linear in τ_3, namely, the 12 Eqs. (9.50a–d) plus the six equations (9.54a–f) and the two scalar equations (9.23c and d). Moreover, the left-hand sides of the foregoing 20 equations are trilinear in τ_3, $\mathbf{x}_4$, and $\mathbf{x}_5$, while their right-hand sides are trilinear in τ_3, $\mathbf{x}_1$, and $\mathbf{x}_2$. These 20 equations can thus be written in the form

$$\mathbf{A}\mathbf{x} = \boldsymbol{\beta} \tag{9.55a}$$

where the 20×16 matrix $\mathbf{A}$ is a function of the data only, while the 20-dimensional vector $\boldsymbol{\beta}$ is trilinear in τ_3, $\mathbf{x}_1$, and $\mathbf{x}_2$, the 16-dimensional vector $\mathbf{x}$ being defined, in turn, as

$$\begin{aligned}\mathbf{x} \equiv [&\tau_3 c_4 c_5 \quad \tau_3 c_4 s_5 \quad \tau_3 s_4 c_5 \quad \tau_3 s_4 s_5 \quad \tau_3 c_4 \quad \tau_3 s_4 \quad \tau_3 c_5 \quad \tau_3 s_5 \\ &c_4 c_5 \quad c_4 s_5 \quad s_4 c_5 \quad s_4 s_5 \quad c_4 \quad s_4 \quad c_5 \quad s_5]^T\end{aligned} \tag{9.55b}$$

Next, matrix $\mathbf{A}$ and vector $\boldsymbol{\beta}$ are partitioned as

$$\mathbf{A} \equiv \begin{bmatrix} \mathbf{A}_U \\ \mathbf{A}_L \end{bmatrix}, \quad \boldsymbol{\beta} \equiv \begin{bmatrix} \boldsymbol{\beta}_U \\ \boldsymbol{\beta}_L \end{bmatrix} \tag{9.56}$$

where $\mathbf{A}_U$ is a nonsingular 16×16 matrix, $\mathbf{A}_L$ is a 4×16 matrix, vector $\boldsymbol{\beta}_U$ is 16-dimensional, and vector $\boldsymbol{\beta}_L$ is 4-dimensional. Moreover, the two foregoing matrices are functions of the data only. Thus, we can solve for $\mathbf{x}$ from the first 16 equations of Eq. (9.55a) in the form

$$\mathbf{x} = \mathbf{A}_U^{-1} \boldsymbol{\beta}_U$$

Upon substituting the foregoing value of $\mathbf{x}$ into the four remaining equations of Eq. (9.55a), we derive four equations free of $\mathbf{x}$, namely,

$$\mathbf{A}_L \mathbf{A}_U^{-1} \boldsymbol{\beta}_U = \boldsymbol{\beta}_L \tag{9.57}$$

In Eq. (9.57) the two matrices involved are functions of the data only, while the two vectors are trilinear in τ_3, $\mathbf{x}_1$, and $\mathbf{x}_2$. These equations are now cast in the form

$$(A_i c_2 + B_i s_2 + C_i)\tau_3 + D_i c_2 + E_i s_2 + F_i = 0, \quad i = 1, 2, 3, 4 \tag{9.58a}$$

where all coefficients $A_i, \ldots, F_i$ are linear in $\mathbf{x}_1$. Next, Li and co-authors substitute c_2 and s_2 in the foregoing equations by their equivalents in terms of $\tau_2 \equiv \tan(\theta_2/2)$, thereby obtaining, for $i = 1, 2, 3, 4$,

$$C_{ii}\tau_2^2\tau_3 + 2B_i\tau_2\tau_3 + A_{ii}\tau_3 + F_{ii}\tau_2^2 + 2E_i\tau_2 + D_{ii} = 0 \tag{9.58b}$$

with the definitions

$$A_{ii} \equiv A_i + C_i \tag{9.58c}$$

$$C_{ii} \equiv C_i - A_i \tag{9.58d}$$

$$D_{ii} \equiv D_i + F_i \tag{9.58e}$$

$$F_{ii} \equiv F_i - D_i \tag{9.58f}$$

Further, τ_2 and τ_3 are both eliminated dialytically from the four equations (9.58a). To this end, both sides of all four equations (9.58b) are multiplied by τ_2, which yields

$$C_{ii}\tau_2^3\tau_3 + 2B_i\tau_2^2\tau_3 + A_{ii}\tau_2\tau_3 + F_{ii}\tau_2^3 + 2E_i\tau_2^2 + D_{ii}\tau_2 = 0 \tag{9.58g}$$

We have now eight equations that are linear homogeneous in the eight-dimensional nonzero vector $\mathbf{z}$ defined as

$$\mathbf{z} \equiv \left[\tau_2^3\tau_3 \;\; \tau_2^2\tau_3 \;\; \tau_2^3 \;\; \tau_2\tau_3 \;\; \tau_2^2 \;\; \tau_3 \;\; \tau_2 \;\; 1\right]^T \tag{9.58h}$$

and hence, the foregoing eight-dimensional system of equations takes the form

$$\mathbf{Mz} = \mathbf{0} \tag{9.59}$$

where the 8×8 matrix $\mathbf{M}$ is simply

$$\mathbf{M} \equiv \begin{bmatrix} 0 & C_{11} & 0 & 2B_1 & F_{11} & A_{11} & 2E_1 & D_{11} \\ 0 & C_{22} & 0 & 2B_2 & F_{22} & A_{22} & 2E_2 & D_{22} \\ 0 & C_{33} & 0 & 2B_3 & F_{33} & A_{33} & 2E_3 & D_{33} \\ 0 & C_{44} & 0 & 2B_4 & F_{44} & A_{44} & 2E_4 & D_{44} \\ C_{11} & 2B_1 & F_{11} & A_{11} & 2E_1 & 0 & D_{11} & 0 \\ C_{22} & 2B_2 & F_{22} & A_{22} & 2E_2 & 0 & D_{22} & 0 \\ C_{33} & 2B_3 & F_{33} & A_{33} & 2E_3 & 0 & D_{33} & 0 \\ C_{44} & 2B_4 & F_{44} & A_{44} & 2E_4 & 0 & D_{44} & 0 \end{bmatrix}$$

Now, since $\mathbf{z}$ is necessarily nonzero, Eq. (9.59) should admit nontrivial solutions, and hence, matrix $\mathbf{M}$ should be singular, which leads to the condition below:

$$\det(\mathbf{M}) = 0 \tag{9.60}$$

Thus, considering that all entries of $\mathbf{M}$ are linear in $\mathbf{x}_1$, $\det(\mathbf{M})$ is *octic* in $\mathbf{x}_1$, and hence, Eq. (9.60) is equally octic in $\mathbf{x}_1$. By virtue of Fact 9.2.2, then, Eq. (9.60) is of 16th degree in τ_1; this equation takes the form

$$\sum_0^{16} a_k \tau_1^k = 0 \tag{9.61}$$

which is the characteristic equation sought, its roots providing up to 16 real values of θ_1 for the IDP at hand.

9.4 The Bivariate-Equation Approach

The difference between this approach and those leading to the univariate polynomial, as outlined in Sect. 9.3, lies in three aspects: (a) only four, out of the six original unknowns, are eliminated; (b) the tan-half identities are avoided, in order to avoid polynomial deflation at or around values of π, and to allow for finding *all* real roots; and (c) direct polynomial-root finding is avoided, rough estimates of all roots being found, first, by inspection, then refined by means of a Newton procedure.

Now, to derive the bivariate equations, we have to eliminate three of the five unknowns from the 14 fundamental closure equations. To this end, we resort to Eqs. (9.40b), which are trilinear in $\{\mathbf{x}_i\}_3^5$. Furthermore, from definition (9.41),

$\mathbf{y}_3 \neq \mathbf{0}$, and hence, the 4×3 matrix $\mathbf{D}_1$ of Eq. (9.40b) must be rank-deficient, which means that every one of its four—the number of combinations of four objects taking three at a time—3×3 determinants, obtained by deleting one of its four rows, should vanish. We need, in principle, only two of these determinants to obtain two independent equations in θ_4 and θ_5. To be on the safe side regarding *spurious* roots and *formulation* singularities,[5] we impose the vanishing of all four possible determinants, which yields, correspondingly, four contours in the θ_4–θ_5 plane; the intersections of all contours then yield the real (θ_4, θ_5) pairs of values which render $\mathbf{D}_1$ rank-deficient. Each of the four equations thus derived describes a contour $\mathcal{C}_i$, for $i = 1, 2, 3, 4$, in the θ_4–θ_5 plane:

$$\mathcal{C}_i : \quad F_i(\theta_4, \theta_5) = 0, \quad i = 1, 2, 3, 4. \tag{9.62}$$

Note that, by plotting the four contours in a square of the θ_4–θ_5 plane, of side 2π, we ensure that no real solutions will be missed.

The intersection points can be detected visually by the user or, automatically, by a suitable *graphical user interface (GUI)*.[6] Regardless of the detection method, numerical code can be employed to refine each pair (θ_4, θ_5) of intersection coordinates to the desired accuracy. The well-known Newton–Raphson method for nonlinear-equation solving, outlined in Sect. B.3, can be used here. However, this method works for solving systems of as many equations as unknowns. In our case, we end up with four nonlinear equations in only two unknowns. While, in principle, any two of those four equations can be used to solve for the two unknowns, numerical roundoff error and the *numerical conditioning* of the problem at hand, to be discussed in Sect. 9.4.1, will invariably lead to different numerical solutions for different choices of those two equations. The question then is which of the four distinct solutions to pick up. In order to avoid this quandary, we suggest here to regard all four equations as independent, entailing possible contradictions—roundoff errors may render the four equations independent, which they aren't. With this approach, then, rather than one *solution* to the four equations, what we seek is their *least-square approximation*, which can be done using a method known as *Newton–Gauss* (Dahlquist and Björck 1974), as outlined in Sect. B.4. Alternatively, Matlab's function `lsqnonlin` can be used to find the same least-square approximation. In any event, the problem is solved *iteratively*. Within the Newton–Gauss method, a linear overdetermined system of equations is solved at each iteration, using one of the methods of Sect. B.1.

[5] *Formulation* singularities occur when, in the absence of a kinematic singularity—characterized by the vanishing of $\det(\mathbf{J})$, for $\mathbf{J}$ defined as in Eq. (5.10b)—two or three contours $\mathcal{C}_i$ are tangent at an intersection. When this is the case, and a pair of functions (9.62) is chosen to find their roots, whose contours are tangent, the numerical computation of the coordinates of the intersection point becomes impossible.

[6] The intersection points appearing in Figs. 9.3 and 9.4 were obtained using the Matlab GUI developed by Dr. Stephane Caro, a postdoctoral fellow at McGill University's Robotic Mechanical Systems Laboratory. The GUI is available in the CD accompanying this edition

In this way, two of the unknown joint angles, θ_4 and θ_5, are computed accurately, the remaining four unknowns being determined uniquely, as described in Sect. 9.6. Notice, however, that spurious solutions to the IDP are likely to occur. These are intersections of the four contours which, although verifying the four equations (9.62), fail to produce a full set of solutions $\{\,\theta_i\,\}_1^6$. The computation of all remaining joint variables, θ_1, θ_2, θ_3 and θ_6, once θ_4 and θ_5 are available, is the subject of Sect. 9.6.3.

9.4.1 Numerical Conditioning of the Solutions

We recall here the concept of *condition number* of a square matrix (Golub and Van Loan 1989), as introduced in Sect. 5.8. In this subsection we stress the relevance of the concept in connection with the *accuracy* of the computed solutions of the general IDP.

The concept of condition number of a square matrix is of the utmost importance because it measures the roundoff-error amplification upon solving a system of linear equations having that matrix as coefficient. The condition number of a matrix, discussed in Sect. 5.8, can be computed in many possible ways. For the purpose at hand, it will prove convenient to work with the condition number defined in terms of the Frobenius norm, as given in Eqs. (5.80a and b).

In the context of the bivariate-equation approach, we can intuitively argue that the accuracy in the computation of a solution is dictated by the angle at which two contours giving a solution intersect. Thus, the solutions computed most accurately are those determined by contours intersecting at right angles. On the contrary, the solutions computed least accurately are those obtained by tangent contours. We shall formalize this observation in the discussion below.

We distinguish between the condition number of a matrix and the conditioning of a solution of a nonlinear system of equations. We define the latter as the condition number of the Jacobian matrix of the system, evaluated at that particular solution. More concretely, let

$$f_1(x_1, x_2) = 0$$

$$f_2(x_1, x_2) = 0$$

be a system of two nonlinear equations in the two unknowns x_1 and x_2. Moreover, the Jacobian matrix of this system is defined as

$$\mathbf{F} \equiv \begin{bmatrix} (\nabla f_1)^T \\ (\nabla f_2)^T \end{bmatrix} \tag{9.63}$$

where ∇f_k denotes the gradient of function $f_k(x_1,\, x_2)$, defined in turn as

$$\nabla f_k \equiv \begin{bmatrix} \partial f_k/\partial x_1 \\ \partial f_k/\partial x_2 \end{bmatrix} \tag{9.64}$$

It is to be noted that multiplying each of the two given equations by a scalar other than zero does not affect its solutions, each Jacobian row being, then, correspondingly multiplied by the same scaling factor. To ease matters, we will assume henceforth that each of the above equations has been properly scaled so as to render its gradient a unit vector in the plane of the two unknowns. In order to calculate the condition number of $\mathbf{F}$, which determines the conditioning of the solutions, we calculate first $\mathbf{F}\mathbf{F}^T$ and its inverse, namely,

$$\mathbf{F}\mathbf{F}^T = \begin{bmatrix} 1 & \nabla f_1 \cdot \nabla f_2 \\ \nabla f_1 \cdot \nabla f_2 & 1 \end{bmatrix} \equiv \begin{bmatrix} 1 & \cos\gamma \\ \cos\gamma & 1 \end{bmatrix}$$

and

$$\left(\mathbf{F}\mathbf{F}^T\right)^{-1} = \frac{1}{\sin^2\gamma} \begin{bmatrix} 1 & -\cos\gamma \\ -\cos\gamma & 1 \end{bmatrix}$$

where γ is the angle at which the contours intersect. The condition number κ_F of $\mathbf{F}$, based on the Frobenius norm, can then be computed as

$$\kappa_F = \frac{1}{|\sin\gamma|} \qquad -\pi \le \gamma \le \pi \tag{9.65}$$

which means that for the best possible solutions from the numerical conditioning viewpoint, the two contours cross each other at right angles, whereas at singular configurations, the contours are tangent to each other. The reader may have experienced that, when solving a system of two linear equations in two unknowns with the aid of drafting instruments,[7] the solution becomes fuzzier as the two lines representing those equations become closer to parallel.

9.5 Implementation of the Solution Method

Whatever method is chosen to solve the IDP, the solution procedure will eventually require numerical computations. Indeed, both the univariate-polynomial and the bivariate-equation approaches ultimately resort to a numerical procedure to find either the roots of a polynomial equation that can be of up to 16th degree or, correspondingly, the solutions of a system of trigonometric equations. Now, formulas for the roots of polynomial equations are available only for the quadratic, the cubic and the quartic polynomials[8]; those for the cubic and quartic equations are

[7]Graphical methods of mechanism analysis rely on this form of linear-equation solving.

[8]The Italian mathematicians Niccolò Tartaglia—meaning the “stammerer,” his real name believed to have been Fontana—(1535) and Girolamo Cardano (1545), independently, or so each claimed, found the formula for the three roots of the cubic equation, now known as *Cardan’s formula*.

so cumbersome that in practice they are seldom applied. The Italian mathematician Ruffini gave a sketch of a proof in 1799 showing that formulas for the roots of polynomials of fifth or higher degree are not possible in general (Wells 1986). Then, the Norwegian mathematician Abel, in 1826, provided a more rigorous proof of the same result. It was the genius of the French Evariste Galois (1811–1832) that, aided by Galois' own *theory of groups* (Livio 2005), led to an elegant theory on the solvability of polynomial equations that closed an important chapter in the history of mathematics.

Now, when numerically solving the equations involved, whether polynomial or trigonometric, intermediate computations can yield coefficients with absolute values of disparate orders of magnitude, which is prone to numerical instabilities—*ill-conditioning*. These occur naturally in the neighborhood of singularities, and cannot be avoided. Another source of ill-conditioning lies in the data themselves. When working with two different sets of equations, one representing point displacements, the other angular displacements, we end up with a mixture of equations with physical units of length and equations that are dimensionless. Such a mixture is a source of ill-conditioning, which can be avoided without too much effort: we recommend to start by rendering the point-displacement equations *dimensionless*, which can be done by dividing the DH parameters $\{ a_i, b_i \}_1^6$ introduced in Sect. 4.2 and the position vector $\mathbf{p}$ of the EE operation point by the characteristic length L introduced in Sect. 5.8. This stage, which can be termed *normalization*, is done in the numerical examples included in Sect. 9.7.

Furthermore, when refining the rough estimates of the contour intersections, as occurring in the implementation of the bivariate-equation approach, we are confronted with computing the least-square approximation to an overdetermined system of nonlinear equations. This is a well-researched problem in the realm of *numerical analysis* (Dahlquist and Björck 1974). While effective methods exist that solve the problem without resorting to gradients, we have used in the solutions an in-house developed package of C routines and Matlab functions, ODA, for a broad class of problems occurring in mathematical programming.[9] In this library, we have a routine, `LSSNLS`, that implements the Newton–Gauss algorithm described in Sect. B.4. `LSSNLS` requires an initial guess $\mathbf{x}_0$ for the unknown vector $\mathbf{x}$ as well as information on the dimensions n of $\mathbf{x}$, the number of unknowns, and of $\mathbf{f}(\mathbf{x})$, the number of equations, $m > n$. Then, `LSSNLS` returns an *optimum value* $\mathbf{x}^*$ that *best* approximates the overdetermined system of equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ in the *least-square* sense, and that is dependent on $\mathbf{x}_0$. In the absence of ill-conditioning, $\mathbf{x}^*$ is the *local optimum* of the problem *closest* to the initial guess $\mathbf{x}_0$. However, the Matlab GUI that was developed by Dr. Caro—see footnote 6—to automate the refining of the visual

---

Ferrari's formula—so named after the Italian mathematician Ludovico Ferrari, a disciple of Cardano's—provides the four roots of a quartic polynomial.

[9]The ODA library is available on www.mcgill.ca/~/rmsl/Angeles_html/courses/MECH577/.

estimates relies on Matlab's `lsqnonlin` function. The method implemented in this function is *direct*, in that it is based solely on function evaluations, thus obviating gradient computations.

9.6 Computation of the Remaining Joint Angles

So far we have reduced the system of displacement equations to either one single univariate polynomial in the tangent of half one of the joint angles—the univariate-polynomial approach—or a system of bivariate trigonometric equations in the sines and cosines of two joint angles—the bivariate-equation approach. In either case, we still need a procedure to compute the remaining joint angles, which is the subject of the balance of this section.

9.6.1 The Raghavan–Roth Procedure

The most straightforward means of computing θ_4 and θ_5 in this procedure is Eq. (9.48d), which can be interpreted as an eigenvalue problem associated with the 12×12 matrix $\mathbf{S}$, and has one known eigenvalue, namely, 0, for its sole variable, θ_3, was computed so as to render $\mathbf{S}$ singular. Now, every scientific package offers eigenvalue calculations, whereby the eigenvectors are usually produced in a normalized form, i.e., with all eigenvectors computed as unit vectors. Let, for example, $\boldsymbol{\sigma}$ be the 12-dimensional eigenvector of $\mathbf{S}$ corresponding to the zero eigenvalue. In this case, $\|\boldsymbol{\sigma}\| = 1$, but $\tilde{\mathbf{x}}_{45}$, the solution sought, is obviously of magnitude greater than unity, for its 12th component, σ_{12}, is exactly 1, according to its definition, Eq. (9.48e). In order to produce $\tilde{\mathbf{x}}_{45}$ from $\boldsymbol{\sigma}$, then, all we need is a suitable scaling of this vector that will yield $(\tilde{\mathbf{x}}_{45})_{12} = 1$. We thus have that $\sigma_{12} \neq 0$—otherwise, Eqs. (9.48d) would be inconsistent—and hence,

$$\tilde{\mathbf{x}}_{45} = \frac{1}{\sigma_{12}}\boldsymbol{\sigma}$$

The outcome will be a set of unique values of θ_4 and θ_5 for each of the 16 possible values of θ_3.

Next, θ_1 and θ_2 are computed from Eq. (9.35), which is rewritten below in a more suitable form:

$$\overline{\mathbf{R}}\mathbf{x}_{12} = \overline{\mathbf{x}}_{345} \tag{9.66a}$$

with the 14-dimensional vector $\overline{\mathbf{x}}_{345}$ defined as

$$\overline{\mathbf{x}}_{345} \equiv \overline{\mathbf{P}}\mathbf{x}_{45} \tag{9.66b}$$

Since $\overline{\mathbf{R}}$ is a 14×8 matrix, Eq. (9.66a) comprises 14 linear equations in the eight unknown components of $\mathbf{x}_{12}$. Although any eight of the 14 equations (9.66a) suffice, in principle, to determine $\mathbf{x}_{12}$, we should not forget that these computations will most likely be performed with finite precision, and hence, roundoff-error amplification is bound to occur. In order to keep roundoff errors as low as possible, we recommend to use all the foregoing 14 equations and calculate $\mathbf{x}_{12}$ as the *least-square approximation* of the overdetermined system (9.66a). This approximation will be, in fact, the solution of the given system because all 14 equations are compatible. The least-square solution of this system yields, *symbolically*,

$$\mathbf{x}_{12} = (\overline{\mathbf{R}}^T \overline{\mathbf{R}})^{-1} \overline{\mathbf{R}}^T \overline{\mathbf{x}}_{345} \tag{9.66c}$$

In practice, the foregoing least-square approximation is computed using an orthogonalization procedure (Golub and Van Loan 1989), the explicit or the numerical inversion of the product $\mathbf{H}^T \mathbf{H}$ being advised against because of its frequent ill-conditioning. Appendix B outlines the robust numerical computation of the least-square approximation of an overdetermined system of equations using orthogonalization procedures. The only remaining unknown is θ_6, which is computed below: This unknown is readily computed from Eq. (4.9a). Indeed, the first of the three vector equations represented by this matrix equation yields

$$\mathbf{Q}_1 \mathbf{Q}_2 \mathbf{Q}_3 \mathbf{Q}_4 \mathbf{Q}_5 \mathbf{p}_6 = \mathbf{q} \tag{9.67a}$$

where $\mathbf{q}$ denotes the first column of $\mathbf{Q}$, while, according to Eq. (9.5), $\mathbf{p}_6$ denotes the first column of matrix $\mathbf{Q}_6$, i.e.,

$$\mathbf{p}_6 \equiv \begin{bmatrix} \cos\theta_6 \\ \sin\theta_6 \\ 0 \end{bmatrix}, \quad \mathbf{q} \equiv \begin{bmatrix} q_{11} \\ q_{21} \\ q_{31} \end{bmatrix} \tag{9.67b}$$

Thus, Eq. (9.67a) can be readily solved for $\mathbf{p}_6$, i.e.,

$$\mathbf{p}_6 = \mathbf{Q}_5^T \mathbf{Q}_4^T \mathbf{Q}_3^T \mathbf{Q}_2^T \mathbf{Q}_1^T \mathbf{q} \tag{9.68}$$

thereby obtaining a unique value for θ_6 for every set of values of $\{\theta_k\}_1^5$. This completes the solution of the IDP under study.

9.6.2 The Li–Woernle–Hiller Procedure

Once θ_1 is available, the remaining angles are computed from linear equations: Equations (9.59) are first rearranged in nonhomogeneous form, namely,

$$\mathbf{N}\mathbf{z}' = \mathbf{n} \tag{9.69}$$

with the 8×7 matrix $\mathbf{N}$ and the seven- and eight-dimensional vectors $\mathbf{z}'$ and $\mathbf{n}$ defined as

$$\mathbf{N} \equiv \begin{bmatrix} 0 & C_{11} & 0 & 2B_1 & F_{11} & A_{11} & 2E_1 \\ 0 & C_{22} & 0 & 2B_2 & F_{22} & A_{22} & 2E_2 \\ 0 & C_{33} & 0 & 2B_3 & F_{33} & A_{33} & 2E_3 \\ 0 & C_{44} & 0 & 2B_4 & F_{44} & A_{44} & 2E_4 \\ C_{11} & 2B_1 & F_{11} & A_{11} & 2E_1 & 0 & D_{11} \\ C_{22} & 2B_2 & F_{22} & A_{22} & 2E_2 & 0 & D_{22} \\ C_{33} & 2B_3 & F_{33} & A_{33} & 2E_3 & 0 & D_{33} \\ C_{44} & 2B_4 & F_{44} & A_{44} & 2E_4 & 0 & D_{44} \end{bmatrix}$$

and

$$\mathbf{z}' \equiv - \begin{bmatrix} \tau_2^3 \tau_3 \\ \tau_2^2 \tau_3 \\ \tau_2^3 \\ \tau_2 \tau_3 \\ \tau_2^2 \\ \tau_3 \\ \tau_2 \end{bmatrix}, \qquad \mathbf{n} \equiv \begin{bmatrix} D_{11} \\ D_{22} \\ D_{33} \\ D_{44} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Now, Eq. (9.69) represents an overdetermined linear algebraic system of eight equations, but only seven unknowns. Again, we recommend here a least-square approach to cope with ill-conditioning. In this way, the solution of Eq. (9.69) can be expressed *symbolically* in the form

$$\mathbf{z}' = (\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T \mathbf{n}$$

With $\mathbf{z}'$ known, both τ_2 and τ_3, and hence, θ_2 and θ_3, are known uniquely. Further, with θ_1, θ_2, and θ_3 known, the right-hand side of Eq. (9.55a), $\boldsymbol{\beta}$, is known. Since the coefficient matrix $\mathbf{A}$ of that equation is independent of the joint angles, $\mathbf{A}$ is known, and that equation can be solved for vector $\mathbf{x}$ *uniquely*. Once $\mathbf{x}$ is known, the two angles θ_4 and θ_5 are uniquely determined, with θ_6 the sole remaining unknown; this can be readily determined, also uniquely, as discussed in connection with the Raghavan–Roth method.

9.6.3 The Bivariate-Equation Approach

After all common intersections of the four foregoing contours have been determined, we have already two of the unknowns, θ_4 and θ_5, the remaining four unknowns being calculated uniquely as described presently. First, we calculate one of the remaining joint variables, θ_3, using Eq. (9.40b). For this purpose, we evaluate matrix $\mathbf{D}_1$ *for all intersection points*. Then, we rewrite the same equation in the form

$$\mathbf{H}\mathbf{x}_3 = \boldsymbol{\tau} \tag{9.70a}$$

the 4×2 matrix $\mathbf{H}$ being obtained from $\mathbf{D}_1$ by excluding its last column, which is denoted by $-\boldsymbol{\tau}$. Moreover, matrix $\mathbf{H}$ and the four-dimensional vector $\boldsymbol{\tau}$ are both bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$ and hence, known. Again, we use all four equations (9.70a) at our disposal to compute the two-dimensional vector $\mathbf{x}_3$ using a least-square approach. If $\mathbf{H}$ is of full rank—its two columns are linearly independent—then the solution can be expressed *symbolically* in the form

$$\mathbf{x}_3 = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\boldsymbol{\tau} \tag{9.70b}$$

However, if $\mathbf{H}$ is rank-deficient, i.e., if its two columns are linearly dependent, then the inverse appearing in Eq. (9.70b) cannot be computed, the solution (θ_4, θ_5) being spurious. In fact, even if $\mathbf{H}$ is of full rank, the computed $\mathbf{x}_3$ may fail to be a unit vector, as required by its definition. The outcome here is that

$$\cos^2\theta_3 + \sin^2\theta_3 \neq 1$$

which means that the value of $\mathbf{x}_3$ computed from Eq. (9.70a) will yield a complex value of θ_3. In this case, the intersection (θ_4, θ_5) at stake is spurious as well.

When $\mathbf{H}$ is of full rank and the computed $\mathbf{x}_3$ is of unit Euclidean norm, Eq. (9.70b) determines θ_3 *uniquely* for the given values of θ_4 and θ_5.

With θ_3, θ_4 and θ_5 known, we can now calculate θ_1 and θ_2 simultaneously from Eq. (9.27), which we reproduce below in a more suitable form

$$\mathbf{R}\mathbf{x}_{12} = \mathbf{x}_{345} \tag{9.71}$$

where $\mathbf{R}$ is a 14×8 matrix depending only on the problem data, while $\mathbf{x}_{345}$, defined as

$$\mathbf{x}_{345} \equiv \mathbf{P}\mathbf{x}_{45} \tag{9.72}$$

is a 14-dimensional vector trilinear in $\mathbf{x}_3$, $\mathbf{x}_4$, and $\mathbf{x}_5$, and is hence, known. Moreover, matrices $\mathbf{P}$ and $\mathbf{R}$ as well as vectors $\mathbf{x}_{12}$ and $\mathbf{x}_{45}$ were defined in Eqs. (9.27) and (9.28a and b). Again, we have an overdetermined system, of 14 equations, in eight unknowns this time, which can best be solved for $\mathbf{x}_{12}$ using a least-square approach with an orthogonalization procedure. The unique solution of the overdetermined system at hand can thus be expressed as

$$\mathbf{x}_{12} = (\mathbf{R}^T\mathbf{R})^{-1}\mathbf{R}^T\mathbf{x}_{345} \tag{9.73}$$

Note that the solution thus obtained determines $\mathbf{x}_1$ and $\mathbf{x}_2$ uniquely, the only remaining unknown being θ_6, which is computed as in Eq. (9.68).

9.7 Examples

We solve the examples below using the bivariate-equation approach with the purpose of both helping the reader visualize the real solutions and avoiding the formulation singularities brought about by the tan-half identities.[10]

Example 9.7.1. Find all inverse-displacement solutions of the Fanuc Arc Mate S series manipulator of 1990 for the end-effector pose given below:

$$\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{p} = \begin{bmatrix} 130 \\ 850 \\ 1540 \end{bmatrix}$$

in which $\mathbf{p}$ is given in mm, the DH parameters of the robot being given in Table 5.2.

Solution: For starters, we divide the DH parameters $\{a_i, b_i\}_1^6$ and vector $\mathbf{p}$ by $L = 351.23$ mm, the characteristic length of this manipulator found in Sect. 5.8. In following the bivariate-equation approach, we plot the four contours in the θ_4–θ_5 plane guaranteeing that matrix $\mathbf{D}_1$ of Eq. (9.40b) is rank-deficient. The four contours are superimposed in Fig. 9.3, where, apparently, we can detect eight intersections.

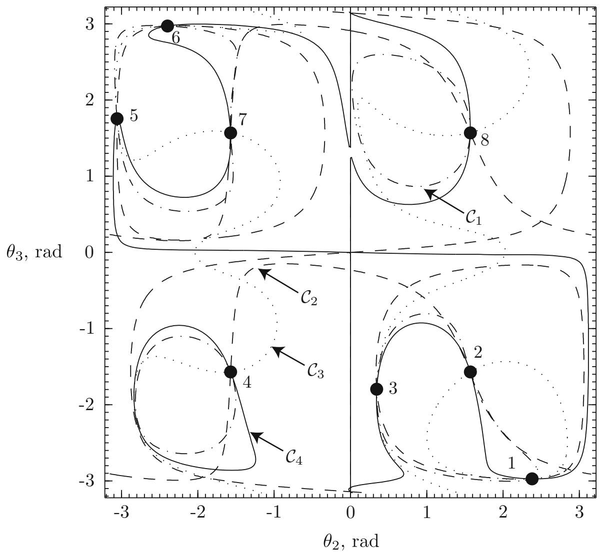

Fig. 9.3 Contours $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$, and $\mathcal{C}_4$ for the Fanuc Arc Mate S series manipulator of 1990

[10]The accompanying CD includes a GUI allowing the user to automate the computation of accurate values of the joint variables by clicking at the visual estimates of the intersections of all four contours.

Table 9.2 Rough estimates of the coordinates of the intersection points of the Fanuc Arc Mate S series manipulator of 1990

	Sol'n no.							
	1	2	3	4	5	6	7	8
θ_4 (rad)	2.38	1.57	0.34	−1.57	−3.06	−2.39	−1.57	1.57
θ_5(rad)	−2.97	−1.57	−1.79	−1.57	1.75	2.97	1.57	1.57

Table 9.3 Refined estimates of the coordinates of the eight intersection points of Fig. 9.3

Inters'n no.	θ_4 (rad)	θ_5 (rad)
1	2.381637539	−2.973100582
2	1.570796327	−1.570796327
3	0.344592933	−1.797513978
4	−1.570796327	−1.570796327
5	−3.060229795	1.755742649
6	−2.397512950	2.972311705
7	−1.570796327	1.570796327
8	1.570796327	1.570796327

Table 9.4 Legitimate solutions of the inverse displacement of the Fanuc Arc Mate S series manipulator of 1990 at the given pose

Sol'n no.	1	3	5	6
θ_4	136.457°	19.743°	−175.338°	−137.367°
θ_5	−170.346°	−102.989°	100.596°	170.300°

The coordinates (θ_4, θ_5) of each intersection point are first estimated by inspection, as listed in Table 9.2. Further, we submit each of these eight values as an initial guess to the Newton–Gauss procedure—or Matlab's function `lsqnonlin`—to find the least-square approximation of the overdetermined system of four equations in two unknowns of Eq. (9.62). The eight solutions thus found are then used to compute $\mathbf{x}_3$ of Eq. (9.70a). As it turned out, solutions 2, 4, 7 and 8 led to a rank-deficient $\mathbf{H}$, and were, thus, discarded as spurious. For the record, we include all eight least-square solutions found in radians, in Table 9.3.

The legitimate solutions are displayed in Table 9.4, in degrees for easier visualization. The robot thus admits four real inverse displacement solutions at the given pose.

The values of the remaining angles are recorded in Table 9.5.

Example 9.7.2. Here we include an example of a manipulator admitting 16 real inverse displacement solutions. This manipulator was proposed by Li (1990), its Denavit–Hartenberg parameters appearing in Table 9.6.

Solution: The foregoing procedure was applied to this manipulator for an end-effector pose given as

Table 9.5 Remaining angles corresponding to the solutions of Table 9.4

Sol'n no.	θ_1	θ_2	θ_3	θ_6
1	83.366°	90.974°	−8.004°	43.134°
3	70.781°	15.151°	151.077°	175.387°
5	85.417°	16.156°	153.212°	−0.859°
6	83.447°	87.898°	9.268°	−42.221°

Table 9.6 DH parameters of Li's manipulator

i	a_i (m)	b_i (m)	α_i	θ_i
1	0.12	0	−57°	θ_1
2	1.76	0.89	35°	θ_2
3	0.07	0.25	95°	θ_3
4	0.88	−0.43	79°	θ_4
5	0.39	0.50	−75°	θ_5
6	0.93	−1.34	−90°	θ_6

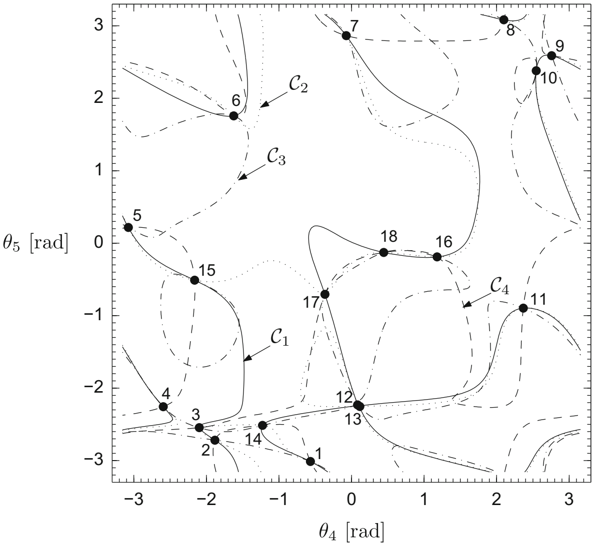

Fig. 9.4 Contours $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$, and $\mathcal{C}_4$ for the Li manipulator

$$
\mathbf{Q} = \begin{bmatrix} -0.357279 & -0.850000 & 0.387106 \\ 0.915644 & -0.237000 & 0.324694 \\ -0.184246 & 0.470458 & 0.862973 \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} 0.798811 \\ -0.000331 \\ 1.200658 \end{bmatrix}
$$

where $\mathbf{p}$ is given in meters. Again, we start by dividing $\{a_i, b_i\}_1^6$ and vector $\mathbf{p}$ by the characteristic length L, that was found to be $L = 890.1$ mm.

The four contours obtained with the bivariate-equation approach are superimposed in Fig. 9.4, where, apparently, we can detect 18 intersections. This means that at least two are spurious, for the number of inverse-displacement solutions can be, at most, 16. In this figure, intersections 12 and 13 appear as one single point. A zoom-in revealed two neighboring solutions in a region around the said single point.

Table 9.7 Rough estimates of the coordinates of the intersection points of Li's manipulator

Sol'n no.	1	2	3	4	5	6	7	8	9
θ_4 (rad)	−0.56	−1.88	−2.09	−2.59	−3.07	−1.62	−0.72	2.09	2.75
θ_5 (rad)	−3.01	−2.71	−2.54	−2.25	0.21	1.75	2.86	3.08	2.5
Sol'n no.	10	11	12	13	14	15	16	17	18
θ_4 (rad)	2.54	2.36	0.08	0.11	−1.22	−2.16	−0.36	1.17	0.44
θ_5 (rad)	2.37	−0.89	−2.23	−2.25	−2.51	−0.50	−0.70	−0.18	−0.12

Table 9.8 Refined estimates, to 14 digits, of the coordinates of the 18 intersection points of Fig. 9.4

Inters'n no.	θ_4 (rad)	θ_5 (rad)
1	−0.5656865073441	−3.0127341939867
2	−1.8817916819320	−2.7181441928227
3	−2.0982054358488	−2.5458222487325
4	−2.5943879129109	−2.2563308501840
5	−3.0760703821644	0.2173802902678
6	−1.6227073591253	1.7564609766664
7	−0.7268312801527	2.8637219341062
8	2.0991093946626	3.0822214487834
9	2.7591234998160	2.5875200635823
10	2.5458806726888	2.3797734576690
11	2.3681644908739	−0.8961886662259
12	0.0834264321499	−2.2306893314165
13	0.1144843294210	−2.2536422392721
14	−1.2262527241259	−2.5145351139614
15	−2.1620940322382	−0.5098897084087
16	−0.3665297041826	−0.7057880105554
17	1.1793192137176	−0.1889758121252
18	0.4440232934648	−0.1282084013846

The coordinates (θ_4, θ_5) of each intersection point are first estimated by inspection, as listed in Table 9.7. Further, we submit each of these 18 values as the initial guess for the Newton–Gauss procedure—or Matlab's function `lsqnonlin`—to find the least-square approximation of the overdetermined system of four equations in two unknowns of Eqs. (9.62). We used ODA to compute the least-square approximation sought, and verified the result with `lsqnonlin`. For the record, we include all 18 solutions found, with 14 digits, in radians, in Table 9.8. The 18 solutions thus found were then used to compute $\mathbf{x}_3$ of Eq. (9.70a). As it turned out, solutions 6 and 14 led to values of Eucledian norm of vector $\mathbf{x}_3$ greater than unity, and were, thus, discarded as spurious. The robot thus admits 16 real inverse displacement solutions at the given pose. The legitimate solutions are displayed in Table 9.9, the values of the remaining angles being displayed in Table 9.10, in degrees for easier visualization.

Table 9.9 Legitimate solutions of the inverse displacement of Li's manipulator at the given pose

Sol'n no.	θ_4	θ_5	Sol'n no.	θ_4	θ_5
1	$-107.81°$	$-155.73°$	9	$-21.00°$	$-40.43°$
2	$120.25°$	$176.59°$	10	$6.55°$	$-129.12°$
3	$4.77°$	$-127.80°$	11	$135.68°$	$135.68°$
4	$158.09°$	$148.26°$	12	$25.44°$	$-7.34°$
5	$-4.16°$	$164.07°$	13	$-176.07°$	$11.57°$
6	$-120.21°$	$-145.86°$	14	$67.57°$	$-10.82°$
7	$-32.41°$	$-172.61°$	15	$-123.87°$	$-29.21°$
8	$145.86°$	$136.35°$	16	$-148.64°$	$-148.64°$

Table 9.10 Remaining angles corresponding to the solutions of Table 9.9

Sol'n no.	θ_1	θ_2	θ_3	θ_6
1	$174.083°$	$-163.302°$	$-164.791°$	$141.281°$
2	$-159.859°$	$-159.324°$	$-111.347°$	$21.654°$
3	$164.800°$	$-154.290°$	$-85.341°$	$-101.359°$
4	$-148.749°$	$-179.740°$	$-78.505°$	$55.719°$
5	$-16.480°$	$-10.747°$	$-58.894°$	$5.677°$
6	$-46.014°$	$-19.256°$	$-46.988°$	$-114.768°$
7	$-22.260°$	$-22.431°$	$-32.024°$	$-17.155°$
8	$-53.176°$	$26.165°$	$9.103°$	$127.978°$
9	$-173.928°$	$150.697°$	$47.811°$	$-92.284°$
10	$-41.684°$	$-29.130°$	$52.360°$	$25.091°$
11	$-137.195°$	$-156.920°$	$68.306°$	$147.446°$
12	$-139.059°$	$128.112°$	$96.052°$	$-119.837°$
13	$-22.696°$	$29.214°$	$98.631°$	$170.303°$
14	$-83.094°$	$57.022°$	$130.976°$	$-110.981°$
15	$1.227°$	$-7.353°$	$142.697°$	$149.208°$
16	$177.538°$	$-148.178°$	$159.429°$	$110.984°$

Example 9.7.3. In this example, we discuss the IDP of DIESTRO, the *isotropic* six-axis orthogonal manipulator shown in Fig. 5.15 (Williams et al. 1993). For a meaning of kinematic isotropy, we refer the reader to Sect. 5.8. This manipulator has the DH parameters given in Table 5.1. The pose of the end-effector leading to an *isotropic posture*, i.e., one whose Jacobian matrix is isotropic, is defined by the orthogonal matrix $\mathbf{Q}$ and the position vector $\mathbf{p}$ displayed below:

$$\mathbf{Q} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{p} = \begin{bmatrix} 0 \\ -50 \\ 50 \end{bmatrix}$$

with $\mathbf{p}$ given in mm. Compute all real inverse displacement solutions at the given pose.

Solution: The characteristic length of DIESTRO was found in Sect. 5.8 to be equal to the common value $a_i = b_i = 50$ mm, for $i = 1, \ldots, 6$. This manipulator, at the given pose of the EE, exhibits a *self-motion*, proper of redundant manipulators, but not expected in a six-revolute robot. A self-motion occurs when a manipulator has the ability to move all its joints while keeping its EE fixed at one given pose. This feature makes the procedure of Sect. 9.4 difficult to apply.[11] We resort, hence, to an alternative approach: We go back to Eq. (9.27) and partition it into two sets of equations:

$$\mathbf{P}_u \mathbf{x}_{45} = \mathbf{R}_u \mathbf{x}_{12} \tag{9.74a}$$

$$\mathbf{P}_l \mathbf{x}_{45} = \mathbf{R}_l \mathbf{x}_{12} \tag{9.74b}$$

where $\mathbf{P}_u$ and $\mathbf{P}_l$ are 6×9 and 8×9 submatrices of $\mathbf{P}$. Likewise, $\mathbf{R}_u$ and $\mathbf{R}_l$ are the corresponding 6×8 and 8×8 submatrices of $\mathbf{R}$. In the above partitioning, the equations must be grouped such that $\mathbf{R}_l$ be nonsingular. Using Eqs. (9.74a and b), six scalar equations free of θ_1 and θ_2 can be derived, namely,

$$\mathbf{\Gamma x}_{45} = \mathbf{0}_6 \,, \qquad \mathbf{\Gamma} \equiv \mathbf{P}_u - \mathbf{R}_u(\mathbf{R}_l^{-1}\mathbf{P}_l) \tag{9.75}$$

where $\mathbf{0}_6$ is the six-dimensional zero vector. Since the entries of the 6×9 matrix $\mathbf{\Gamma}$ are all linear in $\mathbf{x}_3$, the entry in the ith row and jth column of the foregoing matrix, γ_{ij}, can be expressed in the form

$$\gamma_{ij} = a_{ij}c_3 + b_{ij}s_3 + c_{ij} \,; \quad i = 1, \ldots, 6; \; j = 1, \ldots, 9 \tag{9.76}$$

In the above expression, coefficients a_{ij}, b_{ij}, and c_{ij} are independent of the joint variables. Using Eq. (9.76), we can expand Eq. (9.75) and then rearrange the terms in the ith equation, thus obtaining

$$A_i c_3 + B_i s_3 + C_i = 0 \,; \qquad i = 1, \ldots, 6 \tag{9.77a}$$

where, for $i = 1, \ldots, 6$, we have

$$A_i \equiv a_{i1}s_4s_5 + a_{i2}s_4c_5 + a_{i3}c_4s_5 + a_{i4}c_4c_5 + a_{i5}s_4 + a_{i6}c_4 + a_{i7}s_5 + a_{i8}c_5 + a_{i9} \tag{9.77b}$$

$$B_i \equiv b_{i1}s_4s_5 + b_{i2}s_4c_5 + b_{i3}c_4s_5 + b_{i4}c_4c_5 + b_{i5}s_4 + b_{i6}c_4 + b_{i7}s_5 + b_{i8}c_5 + b_{i9} \tag{9.77c}$$

$$C_i \equiv c_{i1}s_4s_5 + c_{i2}s_4c_5 + c_{i3}c_4s_5 + c_{i4}c_4c_5 + c_{i5}s_4 + c_{i6}c_4 + c_{i7}s_5 + c_{i8}c_5 + c_{i9} \tag{9.77d}$$

[11] The self-motion is not readily detected by contour-intersection using this procedure.

Now the six scalar equations (9.77a) are cast in vector form as

$$\mathbf{D}\mathbf{y}_3 = \mathbf{0}_6 \tag{9.78}$$

In the above equation, $\mathbf{D}$ is a 6×3 matrix whose entries are bilinear in $\mathbf{x}_4$ and $\mathbf{x}_5$, while $\mathbf{y}_3$ was defined in Eq. (9.41). Now, to eliminate θ_3, we realize that, from its definition, $\mathbf{y}_3 \neq \mathbf{0}$, and hence, $\mathbf{D}$ *must be rank-deficient*. This means that every one of its 20 3×3 determinants, obtained by picking up three of its six rows at a time, should vanish—the number of combinations of six objects taking three at a time is 20. We need, in principle, only two of these determinants to obtain two independent equations in θ_4 and θ_5. Actually, to be on the safe side, we should impose the vanishing of all 20 possible determinants, which would yield, correspondingly, 20 contours in the θ_4–θ_5 plane; the intersections of all contours would then yield the real (θ_4, θ_5) pairs of values which render $\mathbf{D}$ rank-deficient. Nevertheless, the visualization of the intersections of all 20 contours would be physically impossible, and so, we have to compromise with a smaller number. As we have experienced, only two of the above determinants are prone to yield spurious solutions, for which reason we pick up a reduced number of determinants and derive three equations in θ_4 and θ_5.

We produce the three desired equations by first partitioning the 6×3 matrix $\mathbf{D}$ of Eq. (9.78) into two 3×3 blocks, $\mathbf{D}_u$ being the upper, $\mathbf{D}_l$ the lower block, which thus yields

$$\Delta_1 = \det(\mathbf{D}_u), \quad \Delta_2 = \det(\mathbf{D}_l)$$

Now, since the determinant is not additive, i.e., $\det(\mathbf{D}_u + \mathbf{D}_l) \neq \det(\mathbf{D}_u) + \det(\mathbf{D}_l)$, we choose Δ_3 as

$$\Delta_3 \equiv \det(\mathbf{D}_u + \mathbf{D}_l)$$

which is apparently independent of Δ_1 and Δ_2, thereby obtaining three determinants,[12] which, when equated to zero, yield three independent equations in θ_4 and θ_5. Each of these equations describes a contour $\mathcal{C}_i$, for $i = 1, 2, 3$, in the θ_4–θ_5 plane, i.e.,

$$\mathcal{C}_i : \quad F_i(\theta_4, \theta_5) = 0, \quad i = 1, 2, 3 \tag{9.79}$$

Note that, by plotting the three contours in the $-\pi \leq \theta_i \leq \pi$ region, for $i = 4, 5$, we ensure that no real solutions will be missed.

[12]This idea was proposed by Dr. Kourosh Etemadi Zanganeh, CANMET (Nepean, Ontario, Canada).

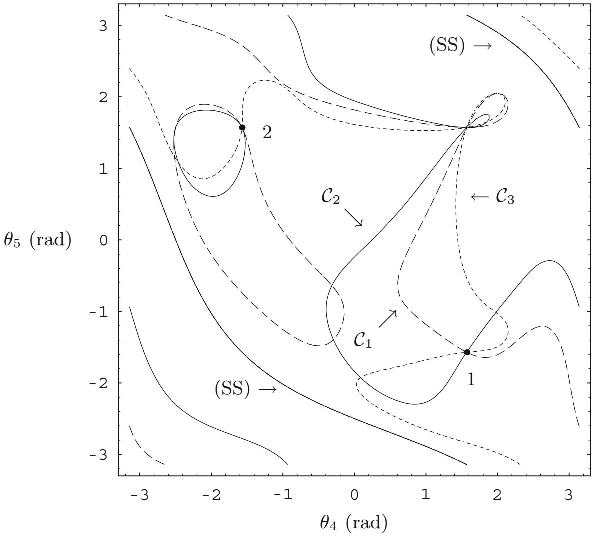

Fig. 9.5 Contours $\mathcal{C}_1$, $\mathcal{C}_2$, and $\mathcal{C}_3$ for the DIESTRO manipulator at given pose

Table 9.11 Inverse displacement solutions of the DIESTRO manipulator

Solution no.	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
1	0°	90°	−90°	90°	−90°	180°
2	180°	−90°	90°	−90°	90°	0°

The three contours thus obtained are plotted in Fig. 9.5a. As this figure shows, the three contours intersect at two isolated points, those labeled 1 and 2. The contours also intersect along a curve labeled SS in the same figure, which thus represents a *manifold* of singular solutions; this means that DIESTRO admits a set of self-motions. These motions can be explained by noticing that when the end-effector is located at the given pose and the manipulator is postured at joint-variable values determined by any point on the SS curve, the six links form a Bricard mechanism (Bricard 1927). The degree of freedom of a Bricard mechanism cannot be determined from the well-known Chebyshev–Grübler-Kutzbach formula (Angeles 2005), which yields a dof $= 0$. Here, the single-dof motion of the mechanism occurs because the six revolute axes are laid out in such a way that if they are grouped in two alternating triads, then these triads intersect.

Furthermore, contours $\mathcal{C}_1$ and $\mathcal{C}_2$ intersect at right angles at solution 1, which corresponds to the isotropic posture of the robot. The numerical values of the joint variables for the isolated solutions are given in Table 9.11.

This example shows interesting features of the manipulator IDP which are not present in manipulators with simpler architectures, such as those with intersecting or parallel consecutive axes.

Moreover, the point of coordinates $\theta_4 = \theta_5 = \pi/2$ of Fig. 9.5 appears to be an intersection of the three contours, and hence, a solution of the IDP at hand. A close-up of this point, as displayed in Fig. 9.6a, shows that this point is indeed an

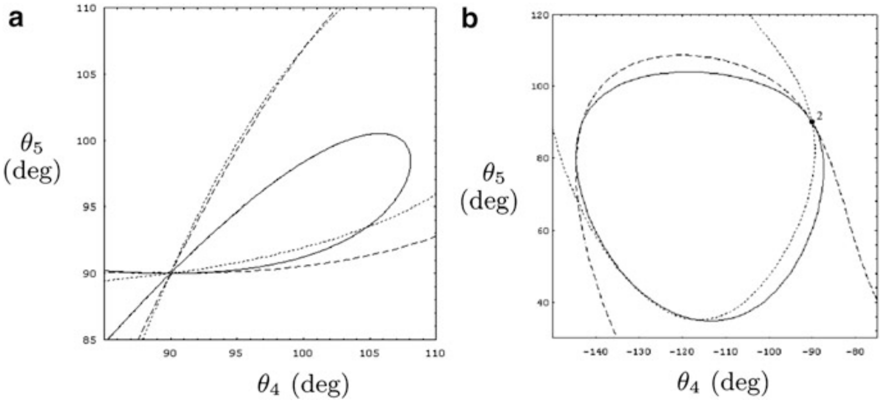

Fig. 9.6 A close-up of: (**a**) the apparent contour intersection at the point of coordinates $\theta_4 = \theta_5 = \pi/2$ (90°); and (**b**) the apparent contour intersection southwest of solution 2

intersection of all three contours, *but* this point is, in fact, a *double point*, i.e., a point at which each contour crosses itself; this gives the point a special character: When verifying whether this point is a solution of the problem under study, we tried to solve for $\mathbf{x}_3$ from eq. (9.70a), but then found that matrix $\mathbf{H}$ of that equation vanishes, and hence, does not allow for the calculation of $\mathbf{x}_3$. An alternative approach to testing the foregoing values of θ_4 and θ_5 is described in Exercise 9.5. In following this approach, it was found that these values do not yield a solution, and hence, the intersection point is discarded.

One more point that appears as an intersection of the three contours is that southwest of solution 2. A close-up of this point, as shown in Fig. 9.6b, reveals that the three contours do not intersect in that region. In summary, then, the manipulator at hand admits two isolated inverse-displacement solutions at the given pose and an infinity of solutions along the curve SS.

9.8 Exercises

9.1 Show that the left-hand side of Eq. (9.23f) represents a pure reflection of vector $\mathbf{h}$ about a plane of unit normal $\mathbf{f}/\|\mathbf{f}\|$, if multiplied by $\|\mathbf{f}\|^2$. Also show that the right-hand side of the same equation represents a pure reflection of vector $\mathbf{i}$ about a plane of unit normal $\mathbf{g}/\|\mathbf{g}\|$, if multiplied by $\|\mathbf{g}\|^2$.

9.2 Show that ψ_1 and ψ_2, as defined in Eqs. (9.43c and d) both vanish.

9.3 In this exercise, we will try to gain insight into the consequence of the double point at $\theta_4 = \theta_5 = \pi/2$ of Fig. 9.5 of Example 9.7.3. To this end, show that, for this combination of values, matrix $\mathbf{H}$ of Eq. (9.70a) becomes zero, and hence, $\mathbf{x}_3$ cannot be computed from this equation. As a result, none of the remaining angles can be computed recursively.

9.4 As an alternative approach to the 14 fundamental equations derived in Sect. 9.2, we recall Eqs. (9.16a and b), if written in a more convenient form, so as to have a minimum number of matrix multiplications, namely,

$$\mathbf{Q}_3\mathbf{Q}_4\mathbf{Q}_5 = \mathbf{Q}_2^T\mathbf{Q}_1^T\mathbf{Q}\mathbf{Q}_6^T$$

$$\mathbf{Q}_2^T\mathbf{Q}_1^T(\mathbf{a}_1 - \mathbf{p}) + \mathbf{Q}_2^T\mathbf{a}_2 + \mathbf{a}_3 + \mathbf{Q}_3\mathbf{a}_4 + \mathbf{Q}_3\mathbf{Q}_4\mathbf{a}_5 + \mathbf{Q}_3\mathbf{Q}_4\mathbf{Q}_5\mathbf{a}_6 = \mathbf{0}$$

Now equate the four linear invariants of the two sides of the first of the two foregoing equations. The result is a set of four scalar equations. When the translational equations are expanded, and appended to the first four equations, a system of seven trigonometric equations in the six unknown angles is derived. Obtain that system of seven equations and comment on their suitability to solve the IDP.

9.5 In Sect. 9.6 we realized that, upon applying the Raghavan–Roth elimination method, and once θ_3 is computed, θ_4 and θ_5 can be computed at once by finding the eigenvector of $\mathbf{S}$ associated with its zero eigenvalue. While this calculation can be performed with the eigenvalue-computation module of any scientific package, computing the eigenvalues of a 12×12 matrix like $\mathbf{S}$ requires an iterative procedure, which can be time-consuming, especially if this computation is only a part of a more complex procedure.

In order to find $\tilde{\mathbf{x}}_{45}$, and hence, θ_4 and θ_5, from Eq. (9.48d), we need not resort to a full eigenvalue problem. Instead, a vector $\mathbf{v}$ can be computed *directly*, as opposed to *iteratively*, that spans the null space of $\mathbf{S}$, for a given computed value of θ_3, if a change of variables is introduced that will yield $\mathbf{S}$ in upper-triangular form. In fact, since $\mathbf{S}$ is *a fortiori* singular, its last row is bound to have zero entries in that form. Devise an algorithm that will render $\mathbf{S}$ in upper-triangular form and hence, compute vector $\tilde{\mathbf{x}}_{45}$ under the conditions that this vector (a) lie in the null space of $\mathbf{S}$ and (b) its 12th entry be unity. *Hint: Apply an orthogonalization procedure, as described in Appendix B.*

9.6 With reference to Example 9.7.3, keep the EE of DIESTRO fixed to the manipulator base at the given pose, thereby forming a 6R closed kinematic chain. Find the singularity locus SS of Fig. 9.5 by means of a kinematic input–output analysis of the closed chain, which turns out to be a Bricard mechanism.

9.7 Using the rough estimates displayed in Table 9.2, Example 9.7.1, compute refined estimates of the coordinates of intersection point 4 upon solving the four equations (9.62) pairwise by means of the Newton–Raphson method. Compute the condition number of each solution based on the Frobenius norm of the 2×2 Jacobian $\mathbf{F}$ of Eq. (9.63). Comment on your result.

9.8 Write a procedure to compute matrix $\mathbf{S}$ of the Raghavan–Roth method. Then, evaluate this matrix at solutions 7 and 8 of Example 9.7.1.

9.9 Derive expressions for vectors $\bar{\mathbf{f}}$, $\bar{\mathbf{g}}$, $\bar{\mathbf{h}}$ and $\bar{\mathbf{i}}$ of Eqs. (9.33a–d).

9.10 Derive an expression for $\mathbf{\Gamma}_4$, and hence, one for $\mathbf{\Gamma}_4\mathbf{x}_{45}$ of Eq. (9.40a).