

Procepts and Proceptual Solutions in Undergraduate Mathematics

Inderasan Naidoo

Department of Mathematical Sciences

University of South Africa

naidoi@unisa.ac.za

Belinda Huntley

St John's College, Johannesburg

huntley@stjohnscollege.co.za

This paper explores and promotes the notion of 'procept' in an undergraduate mathematics course in Linear Algebra for first year pure and engineering students. On the basis of students' preference for procedural to conceptual solutions to mathematical problems, this paper augments the role of certain concepts in pure and applied mathematics, particularly in the problem-solving approaches at the undergraduate level by providing novel solutions to problems solved in the usual traditional manner. The development of the concept of 'procept' and its applicability to mathematics teaching and learning is important to mathematics education research and tertiary pure and applied mathematics didactics in South Africa, welcoming the amalgamation of the theories developed at pre-tertiary level mathematics with theorems and proof at the undergraduate level.

This paper emanated from a series of semester lectures that we presented to first year pure and engineering mathematics students in an undergraduate Linear Algebra course at a large university in Gauteng in South Africa.

The novelty of this paper is twofold. First, it transcends the boundaries between elementary (and secondary) mathematics education research and tertiary pure and applied mathematics didactics, welcoming the amalgamation of the theories developed at pre-tertiary level mathematics education with an approach to teaching theorems and proof at the undergraduate level. The paper explores and promotes the notion of *procept* in an undergraduate mathematics context which emanated from the teaching unit of skew lines in 3-dimensional space, \mathbb{R}^3 . Secondly, it augments the role of certain concepts in pure and applied mathematics, particularly in the problem-solving approaches at the undergraduate level by providing novel solutions to problems solved in the usual traditional manner. The classical approach in determining whether two non-parallel lines in \mathbb{R}^3 are skew is to equate their parametric equations and then solve the resultant system of equations by Gaussian elimination. This procedural approach to solving such a problem has always been the traditional computational mode of instruction in an undergraduate level Linear Algebra course. The nature of the solution then determines the orientation of the lines, either they intersect or they are skew. We provide an immediate (vector) characterisation that determines when two non-parallel lines in \mathbb{R}^3 are skew. We also propose a novel approach to determine the point of intersection of two lines in \mathbb{R}^3 in contrast to the traditional, procedural approach using Gaussian elimination. Furthermore, we consider and propose a *proceptual solution* to the shortest vector between two skew lines in \mathbb{R}^3 .

Process, object and symbol are integrated in a common entity defined by Gray and Tall (1994), as an *elementary procept*. This amalgam of a process that produces a mathematical object, and a symbol that either represents the process or the object are considered for simple arithmetic in Gray and Tall (1991).

The notion of a procept emphasises the importance of having a single mathematical entity encompassing both a process and an object/concept. This duality between *process* and *concept* broadens to a wider audience in undergraduate mathematics. For instance, in differential calculus the partial derivative

$$\frac{\partial v}{\partial \lambda}(\lambda, \gamma) = \lim_{h \rightarrow 0} \frac{v(\lambda + h, \gamma) - v(\lambda, \gamma)}{h}$$

dually represents both *process* (as h gets smaller) and *concept* (the limit itself).

In integral calculus, the definite integral procept of a continuous function $f(x)$ on the interval $[a, b]$,

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i^*) \Delta x$$

is the limit of Riemann sums with x_i^* sample points in the subintervals $[x_{i-1}, x_i]$ of $[a, b]$, represents the area under f (for positive f) or the work done by a force f in moving an object from a to b .

The symbols $\frac{\partial}{\partial \lambda}$ and \int_a^b invoke both a process and a concept involved and the expressions above may be viewed in a range of different contexts. The definition of continuity of a function f at a point $x = a$, $\lim_{x \rightarrow a} f(x) = f(a)$ as a mathematical equation is an amalgam of two elementary procepts.

In undergraduate linear algebra, the symbol $\mathbf{a} \times \mathbf{b}$ for two vectors in \mathbb{R}^3 is a dichotomy of a process (a determinant calculation) and an outcome that produces a mathematical object (a vector cross product). The application of the cross product procept of two vectors in \mathbb{R}^3 tends to be fairly procedural in most traditional undergraduate texts (e.g. Anton & Rorres, 2005; Edwards & Penny, 1988; Salas & Hille, 1995).

The same can be said about the procedure of Gaussian elimination (or the Least-squares solution) and its applications. Undergraduate students are expected to perform routine calculations involving the cross product of two vectors in \mathbb{R}^3 , for example, finding the equation of a plane containing three points, finding the area of a parallelogram or triangle, or determining the volume of a parallelepiped or tetrahedron in applications of vectors in \mathbb{R}^3 . Most of the instruction at this level is aimed at equipping students with the necessary *procedural knowledge* in which the focus is on the *process* to compute or manipulate.

The algorithm of Gaussian elimination in finding the solution to a system of equations is another typical example of the application of procedural knowledge in solving particular undergraduate linear algebra problems, for example, determining the geometric orientation of three planes in \mathbb{R}^3 or finding bases for the eigenspaces of a 3×3 matrix.

Piaget (1972) highlighted cognitive development as involving the duality of process and concept:

...mathematical entities move from one level to another; an operation on such “entities” becomes in its turn an object of the theory, and this process is repeated until we reach structures that are alternately structuring or being structured by “stronger” structures. (p. 70)

Skemp (1987) proposed a general *varifocal theory* in which a *schema* seen as a whole is a concept and a concept seen in detail is a schema. Dubinsky (1991) speaks of *encapsulation* of processes as objects, Sfard (1991) of *reification* of processes as objects, and Gray and Tall (1991, 1994) see the symbol as a pivot between *process* and *concept* – the notion of *procept*. Gray and Tall (1994) introduced the idea of a procept as the symbolisation of an object that arises from processes carried out on other objects. Such procepts can be viewed in two distinct but related ways, as a process or as an object. A procept is considered as a *cognitive* construct in which the symbol can switch from a focus on a process to compute or manipulate, to a concept that may be thought of as an entity that can be manipulated. Thus, a procept refers to the amalgam of concept and process represented by the same symbol.

Tall et al. (2001) believe that

procepts are at the root of human ability to manipulate mathematical ideas in arithmetic, algebra and other theories involving manipulable symbols. They allow the biological brain to switch effortlessly from doing a process to thinking about a concept in a minimal way. (p. 81)

A student possessing a strong proceptual understanding of some mathematical idea is able to effortlessly switch between viewing it as a process (and doing the process) and seeing it in the condensed form as an object. Based on these ideas of a procept and proceptual thinking and understanding (a process of one kind and a concept of another) we propose the notion of a *proceptual solution* as the *type* of alternative solution described in this paper, a solution realised with the integration of a particular process and a particular concept.

We highlight that this paper associates particularly the work of Gray and Tall (1991, 1994) to the mechanics of certain aspects of undergraduate mathematics. In the examples chosen in this paper, by doing the processes of Gaussian elimination, a Least-squares solution technique, and applying the procedure of Cramer's rule together with relating the concept of the cross product of two vectors in \mathbb{R}^3 , we develop an alternative solution to finding the point of intersection of two lines in \mathbb{R}^3 , as well as a similar type of solution in determining the shortest vector between two skew lines. These examples serve to illustrate the differences between the more traditional, fairly procedural approaches to methods of solution with the alternative proceptual solutions encompassing both the process and the concept.

In the next section we provide the necessary notation and preliminary material on vectors and lines in \mathbb{R}^3 that we require. Then in Examples 1 and 2 we provide a traditional solution via Gaussian elimination to find the point of intersection of two lines in \mathbb{R}^3 , as well as the shortest distance between two lines that are skew. Certain aspects of the solution presented in Examples 1 and 2 will be subsequently required. We then consider linear independence and the necessary theory on the rank of a matrix. A criterion for the existence of the intersection of two lines in \mathbb{R}^3 is given and an alternative solution to the point of intersection is presented, derived from using the rank of a matrix and Cramer's rule.

The point of intersection is considered as a Least-squares problem as stated in Theorems 5 and 6 where a vector solution is realised by integrating the *procedure* in solving a (unique) Least-squares problem with the *concepts* of the dot and cross product. The section on proceptual solutions concludes with arriving at a vector solution to the shortest vector between two skew lines in \mathbb{R}^3 and is similar to the approach of finding the point of intersection, save for the use of the process of solution by matrix inversion or that of Cramer's rule.

Preliminaries

Given the points $A_1(a_{11}, a_{12}, a_{13})$, $A_2(a_{21}, a_{22}, a_{23})$ and $A_3(a_{31}, a_{32}, a_{33})$ in \mathbb{R}^3 , we use the bold lowercase vector notation of their position vectors as \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 respectively. We will also make use of the following notations and matrix representations of given vectors. We will form the matrix with vectors as row vectors given by

$$\begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

or column vectors

$$\begin{pmatrix} \mathbf{a}_1^T & \mathbf{a}_2^T \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix}.$$

The vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are *linearly independent* provided that for scalars $\alpha_i \in \mathbb{R}$, the equation

$$\sum_{i=1}^3 \alpha_i \mathbf{a}_i = \mathbf{0}$$

has only the trivial solution $\alpha_i = 0$ for each $i = 1, 2, 3$. Otherwise, the vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are *linearly dependent*. Alternatively, a viable result disseminated (and proved) to undergraduate students using a

matrix determinant is that

$$\mathbf{a}_1, \mathbf{a}_2 \text{ and } \mathbf{a}_3 \text{ are linearly dependent if and only if } \det \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{pmatrix} = 0. \text{ (Edwards \& Penny, 1988, p. 134)}$$

Given \mathbf{a} and \mathbf{b} the position vectors of the points $A(a_1, a_2, a_3)$ and $B(b_1, b_2, b_3)$ the *joining vector* from A to B in \mathbb{R}^3 is $\mathbf{AB} = \mathbf{b} - \mathbf{a} = (b_1 - a_1, b_2 - a_2, b_3 - a_3)$. The *projection* of the vector \mathbf{a} in the direction of, or on \mathbf{b} is

$$\text{proj}_{\mathbf{b}} \mathbf{a} = \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \right) \mathbf{b} = \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2} \right) \mathbf{b}$$

and the *scalar component* of \mathbf{a} on \mathbf{b} is

$$\text{comp}_{\mathbf{b}} \mathbf{a} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|}$$

where $\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i$ is the usual dot product and $\|\mathbf{b}\| = \sqrt{b_1^2 + b_2^2 + b_3^2}$ is the Euclidean norm of \mathbf{b} .

The cross product of \mathbf{a} with \mathbf{b} is the vector $\mathbf{a} \times \mathbf{b}$ calculated as the determinant¹

$$\mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

where \mathbf{i}, \mathbf{j} and \mathbf{k} are the standard basis vectors of \mathbb{R}^3 .

A line ℓ in \mathbb{R}^3 is usually characterised by two points on the line. Alternatively, ℓ is determined by a point $A(a_1, a_2, a_3)$ on the line and a non-zero direction vector $\mathbf{d} = (d_1, d_2, d_3)$ for the line. The *vector equation* for ℓ is then given by

$$\ell: \mathbf{r} = \mathbf{a} + \lambda \mathbf{d}, \lambda \in \mathbb{R}$$

where $\mathbf{r} = (x, y, z)$ is the position vector for each point (x, y, z) on ℓ . Varying the scalar λ through \mathbb{R} , \mathbf{r} traces out the line ℓ . Using the algebra of vectors, one then realises the *parametric equations* for ℓ as

$$\begin{aligned} x &= a_1 + \lambda d_1 \\ y &= a_2 + \lambda d_2, \quad \lambda \in \mathbb{R} \\ z &= a_3 + \lambda d_3. \end{aligned}$$

The distinct lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda \mathbf{d}_1, \lambda \in \mathbb{R}$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma \mathbf{d}_2, \gamma \in \mathbb{R}$ are parallel in the case when their direction vectors \mathbf{d}_1 and \mathbf{d}_2 are parallel, i.e. $\mathbf{d}_1 = k \mathbf{d}_2$ for some scalar k or using the cross product $\mathbf{d}_1 \times \mathbf{d}_2 = \mathbf{0}$. The following result will be used in subsequent sections of this paper.

Lemma 1: Let \mathbf{u}, \mathbf{v} and \mathbf{w} be non-zero vectors in \mathbb{R}^3 , then

1. $\mathbf{u} \parallel \mathbf{v}$ if and only if $\|\mathbf{u} \times \mathbf{v}\| = 0$.
2. $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$ (*Lagrange's identity*)
3. $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$
4. $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u})$

¹ We all know that a 3×3 determinant is a mapping of the form $M_3(\mathbb{R}) \rightarrow \mathbb{R}$, which is certainly not an element of \mathbb{R}^3 , whereas $\mathbf{a} \times \mathbf{b} \in \mathbb{R}^3$ for all vectors \mathbf{a} and \mathbf{b} in \mathbb{R}^3 . What we should see/read in $\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$ is simply a

little memory aid, i.e. a 'recipe' to help us to remember how to calculate the vector product. In other words, it is *as if* you calculate a determinant, and treat the vectors \mathbf{i}, \mathbf{j} and \mathbf{k} as numbers for the time being. It should be emphasised that this is not really a determinant. We regularly fall into similar traps when we deal with determinants, by saying, for example, "multiply the second row of the determinant by 2", instead of "multiply the second row of the matrix of which the determinant is taken, by 2". If one is not properly informed and aware of the difference between the 'loose and informal' and the 'formal and strict' way of referring to these processes, the whole point of distinguishing between process and concept could be missed.

The expansion of the vector triple product (3) in the above lemma appears in Craig (1951) and elementary results on vector and matrix algebra may be accessed in Anton and Rorres (2005), or Edwards and Penny (1988).

The point of intersection: Gaussian elimination

The lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda \mathbf{d}_1, \lambda \in \mathbb{R}$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma \mathbf{d}_2, \gamma \in \mathbb{R}$ intersect if and only if there are unique reals λ and γ for which $\mathbf{r}_1(\lambda) = \mathbf{r}_2(\gamma)$ (Salas & Hille, 1995). The two lines are skew lines if they are not parallel and do not intersect. The (traditional) approach of Salas and Hille to determine whether the distinct lines intersect (and then to find the point of intersection) is to consider the parametric equations for ℓ_1 and ℓ_2 . The two equations are equated, using the parametric form of $\mathbf{r}_1(\lambda) = \mathbf{r}_2(\gamma)$, to realise a 3×2 linear system of equations. Thereafter, the method of Gaussian elimination is used to resolve the non-square system. If no solution arises in the resulting algorithm, the conclusion is that the lines do not meet. Further, consulting the direction vectors it is noted that, if \mathbf{d}_1 and \mathbf{d}_2 are not parallel then the lines are skew.

In addition, the shortest distance between lines ℓ_1 and ℓ_2 is calculated using a projection vector,

$$D = \|\text{proj}_{\mathbf{d}_1 \times \mathbf{d}_2} \mathbf{AB}\| = \frac{|\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2)|}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \quad \dots \quad \text{Equation 1}$$

If the system has a unique solution given by λ_P and γ_P then the point of intersection P is determined using the value λ_P (or γ_P) in the parametric or vector equation for ℓ_1 (or ℓ_2), i.e. finding $\mathbf{r}_1(\lambda_P)$ in ℓ_1 or $\mathbf{r}_2(\gamma_P)$ in ℓ_2 . We illustrate this application of procedural knowledge in the two concrete examples below.

Example 1

Consider the lines with vector equations

$$\ell_1: \mathbf{r}_1 = (-8, 14, -7) + \lambda(2, 1, -5) \quad \text{and} \quad \ell_2: \mathbf{r}_2 = (41, -8, -7) + \gamma(1, 4, 1).$$

For the above lines determine if the lines intersect. If so, find the point of intersection. If not, then find the shortest distance between ℓ_1 and ℓ_2 .

Solution: We follow the classical process in Salas and Hille (1995) where ℓ_1 and ℓ_2 are parametrically represented as

$$\begin{array}{lcl} \ell_1: & x = -8 + 2\lambda & \ell_2: x = 41 + \gamma \\ & y = 14 + \lambda & \text{and} \quad y = -8 + 4\gamma \\ & z = -7 - 5\lambda & z = -7 + \gamma \end{array}$$

The above parametric equations are equated to give

$$\begin{array}{l} -8 + 2\lambda = 41 + \gamma \\ 14 + \lambda = -8 + 4\gamma \\ -7 - 5\lambda = -7 + \gamma \end{array}$$

The above then produces the 3×2 system of equations

$$\begin{array}{l} -\gamma + 2\lambda = 49 \\ -4\gamma + \lambda = -22 \\ -\gamma - 5\lambda = 0 \end{array}$$

The augmented matrix $\begin{pmatrix} -1 & 2 & : & 49 \\ -4 & 1 & : & -22 \\ -1 & -5 & : & 0 \end{pmatrix}$ of this system is then solved by Gaussian elimination.

Reduction produces the equivalent $\begin{pmatrix} 1 & -2 & : & -49 \\ 0 & -7 & : & -218 \\ 0 & 0 & : & -169 \end{pmatrix}$ from which the last row shows that the system is

inconsistent. We can therefore conclude that the lines do not intersect. Furthermore, since the direction vectors for ℓ_1 and ℓ_2 are not parallel, the lines are not parallel. Consequently, the two lines are skew.

To find the shortest distance between ℓ_1 and ℓ_2 , the point $A(-8, 14, -7)$ that lies on ℓ_1 is chosen and the point $B(41, -8, -7)$ that lies on ℓ_2 is used. The procedure is continued by forming the joining vector \mathbf{AB} . The shortest distance is then calculated using Equation 1 to give $D = \frac{170}{\sqrt{11}}$, where the direction vector for ℓ_1 is $\mathbf{d}_1 = (2, 1, -5)$ and that of ℓ_2 is $\mathbf{d}_2 = (1, 4, 1)$.

An example where the lines do intersect is given below. The same procedure that is followed in Example 1 is employed but with a different consequence.

Example 2

Determine whether the two lines intersect, and if so, find the point of intersection where

$$\ell_1: \mathbf{r}_1 = (1, 1, 5) + \lambda(2, -4, -1) \quad \text{and} \quad \ell_2: \mathbf{r}_2 = (4, -1, 4) + \gamma(-1, 6, 1)$$

If the lines do not intersect, then find the shortest distance between them.

Solution: Proceeding as in the previous example would result in the augmented matrix

$$\left(\begin{array}{ccc|ccc} -(-1) & \mathbf{2} & : & 3 & & \\ -(6) & -4 & : & -2 & & \\ -(1) & -1 & : & -1 & & \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & 2 & : & 3 & & \\ -6 & -4 & : & -2 & & \\ -1 & -1 & : & -1 & & \end{array} \right)$$

Performing the Gaussian algorithm on the augmented matrix gives

$$\left(\begin{array}{ccc|ccc} 1 & 0 & : & -1 & & \\ 0 & 1 & : & 2 & & \\ 0 & 0 & : & 0 & & \end{array} \right)$$

The system has the unique solution $\gamma_P = -1$ and $\lambda_P = 2$. The conclusion is then that the two lines ℓ_1 and ℓ_2 intersect in the unique point $P(5, -7, 3)$ realised by using either $\gamma = -1$ in the parametric equation for ℓ_2 or $\lambda = 2$ in the parametric equation for ℓ_1 .

The above calculations are the norm presented to students procedurally in solutions to problems of this type (e.g. Anton & Rorres, 2005; Edwards & Penny, 1988). We reflect on this technique and ascertain whether the solution methodology can be improved and the problem resolved by vector methods. In this procedural approach no existence criterion is presented to acknowledge that the lines do intersect or not and this formed the basis for the investigation in the next section.

The point of intersection: Cramer's rule

Consider the distinct lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda\mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma\mathbf{d}_2$ in \mathbb{R}^3 with position vectors $\mathbf{a} = (a_1, a_2, a_3)$ for the point $A(a_1, a_2, a_3)$ on ℓ_1 and $\mathbf{b} = (b_1, b_2, b_3)$ for the point $B(b_1, b_2, b_3)$ on ℓ_2 . The vectors $\mathbf{d}_1 = (d_{11}, d_{12}, d_{13})$ and $\mathbf{d}_2 = (d_{21}, d_{22}, d_{23})$ are the direction vectors for ℓ_1 and ℓ_2 respectively. It is immediate that the shortest distance given in Equation 1,

$$D = \|\text{proj}_{\mathbf{d}_1 \times \mathbf{d}_2} \mathbf{AB}\| = \left| \frac{\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2)}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \right|$$

between ℓ_1 and ℓ_2 provides a necessary and sufficient condition that determines an intersection. Certainly, if $D = 0$ then ℓ_1 and ℓ_2 must intersect, otherwise ℓ_1 and ℓ_2 are parallel or skew. This translates to a dot product criterion to determine whether ℓ_1 and ℓ_2 intersect or not. Since $D = 0$ if and only if $\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2) = 0$, we may conclude that ℓ_1 and ℓ_2 intersect in a unique point provided that the scalar triple product $\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2) = 0$.

Based on the given vector equations of the two lines, we immediately have the determinant criterion

$$\det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} = \det \begin{pmatrix} b_1 - a_1 & b_2 - a_2 & b_3 - a_3 \\ d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \end{pmatrix} = 0$$

which confirms that ℓ_1 and ℓ_2 intersect. If this determinant is non-zero and the direction vectors \mathbf{d}_1 and \mathbf{d}_2 are not parallel, then the lines are skew. This determinant criterion is equivalent to linear dependence. In our particular case of lines in \mathbb{R}^3 , this criterion for linear dependence may be contextualised in the following result.

Theorem 1: *The distinct non-parallel lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda\mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma\mathbf{d}_2$ in \mathbb{R}^3 intersect in a unique point if and only if the vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent.*

This also has a geometric interpretation. The scalar triple product gives the volume $V = |\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2)|$ of the parallelepiped defined by the vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 . Certainly, if the volume is zero then the three vectors must lie on the same plane. Consequently, the two lines must intersect. Another geometric interpretation is given by Smith and Henderson (1985), where it is shown that the skewness of two lines ℓ_1 and ℓ_2 necessitates the existence of a tetrahedron with volume $V = \frac{1}{6}|\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2)|$. Again, if the volume of the tetrahedron is zero, the vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 lie on the same plane so that the two lines intersect.

Once the fact that the two lines ℓ_1 and ℓ_2 do intersect is established, we can proceed with the traditional approach. However, on a closer look behind this traditional procedure, an interesting observation is made based on the existence of an intersection that has just been described. Suppose that $\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2) = 0$, then λ and γ exist such that

$$\begin{aligned} \mathbf{r}_1(\lambda) = \mathbf{r}_2(\gamma) &\Rightarrow && \mathbf{a} + \lambda\mathbf{d}_1 = \mathbf{b} + \gamma\mathbf{d}_2 \\ &\Rightarrow && \lambda\mathbf{d}_1 - \gamma\mathbf{d}_2 = \mathbf{b} - \mathbf{a} \\ &\Rightarrow && (\lambda d_{11} - \gamma d_{21}, \lambda d_{12} - \gamma d_{22}, \lambda d_{13} - \gamma d_{23}) = (b_1 - a_1, b_2 - a_2, b_3 - a_3) \end{aligned}$$

This then gives the linear system

$$\begin{aligned} \lambda d_{11} - \gamma d_{21} &= b_1 - a_1 \\ \lambda d_{12} - \gamma d_{22} &= b_2 - a_2 \\ \lambda d_{13} - \gamma d_{23} &= b_3 - a_3 \end{aligned}$$

which as a matrix equation, becomes

$$\begin{pmatrix} d_{11} & -d_{21} \\ d_{12} & -d_{22} \\ d_{13} & -d_{23} \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix}$$

In the notation of the defining vectors, the latter matrix equation is

$$\begin{pmatrix} \mathbf{d}_1^T & -\mathbf{d}_2^T \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = (\mathbf{AB}^T) \quad \dots \quad \text{Equation 2}$$

We note that based on the solutions of Example 1 and Example 2 in the previous section, the deliberate bold digits in the matrices makes for a shorter route to the solution as described by the matrix Equation 2. From the given equations of two lines in \mathbb{R}^3 ,

$$\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda\mathbf{d}_1 \text{ and } \ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma\mathbf{d}_2$$

we may directly form the augmented matrix

$$(\mathbf{d}_1^T \quad -\mathbf{d}_2^T : \mathbf{AB}^T) \text{ or } (-\mathbf{d}_2^T \quad \mathbf{d}_1^T : \mathbf{AB}^T)$$

and then pursue the Gaussian algorithm in attaining a solution where $\mathbf{AB} = \mathbf{b} - \mathbf{a}$. If ℓ_1 and ℓ_2 intersect (non-zero scalar triple product of the vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2), the solution for λ and γ in Equation 2 may be simplified with the use of Cramer's Rule as illustrated below. Since Equation 2 is consistent with a unique value for λ and γ , the rank of the coefficient matrix and that of the augmented matrix must be given by the Existence and Uniqueness Theorem of a solution of a system of linear equations in Theorem 2 below (Lipschutz & Lipson, 2001, p. 79).

Theorem 2: Consider a system of linear equations in n unknowns with augmented matrix $M = [A \quad B]$. Then

- a) the system has a solution if and only if $\text{rank}(A) = \text{rank}(M)$
- b) the solution is unique if and only if $\text{rank}(A) = \text{rank}(M) = n$.

The rank of a matrix A , symbolically $\text{rank}(A)$, is usually defined as the number of non-zero rows in an echelon form of A . The symbol $\text{rank}(A)$ is an amalgam of a process (finding an echelon form of A using Gaussian elimination or Theorem 3 below) to produce a mathematical object. Indeed, the rank of a matrix is also a process. Now the rank of a matrix may be determined by the following determinant criterion. The proof may be found in Beesack (1962).

Theorem 3: The rank of a matrix A is r if and only if A has some $r \times r$ submatrix with a non-zero determinant and all square submatrices of larger size have determinant zero.

In examining the matrix Equation 2, at least one of the three 2×2 submatrices of the coefficient matrix has to have non-zero determinant. Selecting that submatrix realises a 2×2 square system with appropriate omission of the corresponding component of the column vector given by \mathbf{AB} . Applying Cramer's rule to the resulting square system gives the unique solution to λ and γ without the use of Gaussian elimination. We illustrate this next by providing an alternative solution to Example 2 via Cramer's rule using Theorem 4:

Theorem 4: In Equation 2, if the coefficient matrix has at least one 2×2 submatrix S whose determinant is non-zero, then Equation 2 has the same solution as that of the (sub)augmented matrix determined by S if and only if $\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2) = 0$ if and only if \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent.

Proof: Suppose that the coefficient matrix of Equation 2 has a submatrix S such that $\det(S) \neq 0$ and that the solution to the subsystem determined by S is the same as that of Equation 2. Since $\det(S) \neq 0$, the subsystem determined by S has a unique solution, say λ_S and γ_S . These are then particular solutions to Equation 2 so that by Theorem 2, $\text{rank}(\mathbf{d}_1^T \quad -\mathbf{d}_2^T : \mathbf{AB}^T) = 2$. Consequently by Theorem 3, $\det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} = 0$, showing that the three vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent.

For the sufficiency, suppose that \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent and that the coefficient matrix of Equation 2 has a submatrix S with non-zero determinant. Since the three vectors are linearly dependent, the 3×3 matrix whose columns are that of \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 has zero determinant. Using determinant properties (column switching and factorisation) we see that

$$\det \begin{pmatrix} d_{11} & -d_{21} & b_1 - a_1 \\ d_{12} & -d_{22} & b_2 - a_2 \\ d_{13} & -d_{23} & b_3 - a_3 \end{pmatrix} = 0$$

Consequently, since $\det(S) \neq 0$ the rank of the coefficient matrix of Equation 2 is 2 and the rank of the augmented matrix, $\text{rank}(\mathbf{d}_1^T \quad -\mathbf{d}_2^T : \mathbf{AB}^T) = 2$. Thus Equation 2 has a unique solution. Suppose that $S = \begin{pmatrix} d_{11} & -d_{21} \\ d_{12} & -d_{22} \end{pmatrix}$. We show that the subsystem determined by S , namely

$$\begin{pmatrix} d_{11} & -d_{21} \\ d_{12} & -d_{22} \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix}$$

has the same solution as that of Equation 2. The solution to this subsystem, say λ_S and γ_S , may be resolved by Cramer's rule. Thereafter, as can be verified, a routine substitution of λ_S and γ_S into $\begin{pmatrix} \mathbf{d}_1^T \\ -\mathbf{d}_2^T \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix}$ satisfies Equation 2. Hence, since Equation 2 has a unique solution it is the same as that of the subsystem determined by S . The same conclusion is realised if S is any of the other two submatrices with non-zero determinant.

Returning to Example 2, we may use the method given by Theorem 1 in determining whether the two lines intersect and finding the point of intersection (by Theorem 4), if it exists. We have

$\mathbf{AB} = (3, -2, -1)$, $\mathbf{d}_1 = (2, -4, -1)$ and $\mathbf{d}_2 = (-1, 6, 1)$. Using the scalar triple product (as a matrix determinant), $\mathbf{AB} \cdot (\mathbf{d}_1 \times \mathbf{d}_2) = 0$. Thus ℓ_1 and ℓ_2 intersect in a unique point P that may be found by solving for λ and γ in the (2-rank) Equation 2, i.e.

$$\begin{pmatrix} 2 & -1 \\ -4 & 6 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ -1 \end{pmatrix}.$$

The coefficient matrix has rank 2 and all its 2×2 submatrices have non-zero determinant (easily verified). The unique solution may be derived using Cramer's rule by solving the subsystem, say

$$\begin{pmatrix} 2 & -1 \\ -4 & 6 \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

to get $\lambda = 2$ and $\gamma = -1$. Using $\lambda = 2$ in ℓ_1 (or $\gamma = -1$ in ℓ_2) we get the position vector for the point of intersection $\mathbf{OP} = (5, -7, 3)$. Selecting any of the other two subsystems will give this same solution. Hence, it all culminates in identifying and selecting a submatrix (that has non-zero determinant) of the coefficient matrix of Equation 2 in arriving at a solution without the use of the Gaussian algorithm.

Proceptual solutions

The point of intersection between two skew lines

According to Tall (1995), mathematical growth starts from perceptions of, and actions on, objects in the environment. Successful *actions on* objects use symbolic representations flexibly as procepts – processes to do and concepts to think about – in arithmetic and algebra. The concept image of a procept uses the symbol to link to suitable processes and relationships in the cognitive structure. Thus, although we may not have anything in our mind which is like a physical object, we have symbols that we can manipulate as if they were mental objects. Tall et al. (2001) consider the word *procedure* as a specific sequence of steps carried out one step at a time, while the term *process* is used in a more general sense to include any number of procedures with *the same effect*:

Those who are procedurally oriented are limited to a particular procedure, with attention focused on the steps themselves, whilst those who see symbolism as process or concept have a more efficient use of cognitive processing. (Tall et al., 2001, p. 90)

In this section we will recall some basic facts on the Least-squares problem. Thereafter, we consider the Least-squares procedure in solving an inconsistent system of equations together with the cross product concept in \mathbb{R}^3 to arrive at a *proceptual solution* to the point of intersection of two distinct lines in \mathbb{R}^3 .

Suppose that A is an $m \times n$ matrix with $m \geq n$ and $\text{rank}(A) = r$. Let $\mathbf{b} \in \mathbb{R}^m$ be any given vector. The Least-squares problem involves finding a vector $\tilde{\mathbf{x}} \in \mathbb{R}^n$ such that the Euclidean norm $\|\mathbf{b} - A\tilde{\mathbf{x}}\|$ is a minimum, i.e. $\|\mathbf{b} - A\tilde{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ for each $\mathbf{x} \in \mathbb{R}^n$. $\tilde{\mathbf{x}}$ is called a *Least-squares solution* of $A\mathbf{x} = \mathbf{b}$. It is well-known (e.g. Osborne, 1961; Edwards & Penny, 1988) that a Least-squares solution $\tilde{\mathbf{x}}$ satisfies the *normal equation* $A^T A \mathbf{x} = A^T \mathbf{b}$.

The algorithm to solve a Least-squares problem would involve the process of finding the product $A^T A$. Thereafter, computing $A^T \mathbf{b}$ and then performing Gaussian elimination on the augmented matrix for the normal equation $A^T A \mathbf{x} = A^T \mathbf{b}$. Naturally, there is only one Least-squares solution if the $n \times n$ matrix $A^T A$ is invertible. A useful criterion for the unique Least-squares solution is given by Theorem 5 in Lay (2005) which is particular to our situation.

Theorem 5: *The matrix $A^T A$ is invertible if and only if the columns of A are linearly independent. In this case, the equation $A\mathbf{x} = \mathbf{b}$ has only one Least-squares solution $\tilde{\mathbf{x}}$, and it is given by $\tilde{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$.*

We also have the following well-known theorem (see e.g. Edwards & Penny, 1988):

Theorem 6: *If the $m \times n$ matrix A has rank n , then the $n \times n$ matrix $A^T A$ is invertible.*

Returning to \mathbb{R}^3 , the lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda \mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma \mathbf{d}_2$ and our Equation 2

$$\begin{pmatrix} \mathbf{d}_1^T & -\mathbf{d}_2^T \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = (\mathbf{AB}^T),$$

we see that if the three vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent then the coefficient matrix $A = \begin{pmatrix} \mathbf{d}_1^T & -\mathbf{d}_2^T \end{pmatrix}$ of Equation 2 has rank 2 so that $A^T A$ is invertible. Furthermore, the direction vectors are not parallel so that the columns of A are linearly independent. By the results above, Equation 2 has a unique Least-squares solution given by

$$\begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = (A^T A)^{-1} (A^T C)$$

where $C = (\mathbf{AB}^T)$. However, since we are in \mathbb{R}^3 , $A^T A$ is a 2×2 matrix and in this unique solution case, $(A^T A)^{-1}$ is easily found using its determinant and its adjoint. The process in finding the unique value of λ and γ is continued by finding $A^T C$ and thereafter computing the final product $(A^T A)^{-1} (A^T C)$.

Now, amalgamating into the above process the vector concepts of the cross product and the dot product, produces the following direct solution. In the context of the vector equations of the lines, we realise a vector solution to the point of intersection without inference to matrix algebra. We first remark that

$$A^T A = \begin{pmatrix} \|\mathbf{d}_1\|^2 & -\mathbf{d}_1 \cdot \mathbf{d}_2 \\ -\mathbf{d}_1 \cdot \mathbf{d}_2 & \|\mathbf{d}_2\|^2 \end{pmatrix}$$

and by Lagrange's identity in Lemma 1,

$$\det(A^T A) = \|\mathbf{d}_1\|^2 \|\mathbf{d}_2\|^2 - (\mathbf{d}_1 \cdot \mathbf{d}_2)^2 = \|\mathbf{d}_1 \times \mathbf{d}_2\|^2.$$

Then

$$\begin{aligned} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} &= (A^T A)^{-1} (A^T C) \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \begin{pmatrix} \|\mathbf{d}_2\|^2 & \mathbf{d}_1 \cdot \mathbf{d}_2 \\ \mathbf{d}_1 \cdot \mathbf{d}_2 & \|\mathbf{d}_1\|^2 \end{pmatrix} \begin{pmatrix} \mathbf{d}_1 \\ -\mathbf{d}_2 \end{pmatrix} (\mathbf{AB}) \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \begin{pmatrix} \|\mathbf{d}_2\|^2 & \mathbf{d}_1 \cdot \mathbf{d}_2 \\ \mathbf{d}_1 \cdot \mathbf{d}_2 & \|\mathbf{d}_1\|^2 \end{pmatrix} \begin{pmatrix} \mathbf{AB} \cdot \mathbf{d}_1 \\ -\mathbf{AB} \cdot \mathbf{d}_2 \end{pmatrix} \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \begin{pmatrix} (\mathbf{d}_2 \cdot \mathbf{d}_2)(\mathbf{AB} \cdot \mathbf{d}_1) - (\mathbf{d}_1 \cdot \mathbf{d}_2)(\mathbf{AB} \cdot \mathbf{d}_2) \\ (\mathbf{d}_1 \cdot \mathbf{d}_2)(\mathbf{AB} \cdot \mathbf{d}_1) - (\mathbf{d}_1 \cdot \mathbf{d}_1)(\mathbf{AB} \cdot \mathbf{d}_2) \end{pmatrix} \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \begin{pmatrix} \mathbf{AB} \cdot ([\mathbf{d}_2 \cdot \mathbf{d}_2] \mathbf{d}_1 - [\mathbf{d}_1 \cdot \mathbf{d}_2] \mathbf{d}_2) \\ \mathbf{AB} \cdot ([\mathbf{d}_1 \cdot \mathbf{d}_2] \mathbf{d}_1 - [\mathbf{d}_1 \cdot \mathbf{d}_1] \mathbf{d}_2) \end{pmatrix} \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \begin{pmatrix} \mathbf{AB} \cdot [\mathbf{d}_2 \times (\mathbf{d}_1 \times \mathbf{d}_2)] \\ \mathbf{AB} \cdot [\mathbf{d}_1 \times (\mathbf{d}_1 \times \mathbf{d}_2)] \end{pmatrix} \quad \dots \quad \text{Lemma 1(3)} \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \begin{pmatrix} \frac{\mathbf{AB} \cdot [\mathbf{d}_2 \times (\mathbf{d}_1 \times \mathbf{d}_2)]}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \\ \frac{\mathbf{AB} \cdot [\mathbf{d}_1 \times (\mathbf{d}_1 \times \mathbf{d}_2)]}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \end{pmatrix} \\ &= \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \begin{pmatrix} \text{comp}_{\mathbf{d}_1 \times \mathbf{d}_2}(\mathbf{AB} \times \mathbf{d}_2) \\ \text{comp}_{\mathbf{d}_1 \times \mathbf{d}_2}(\mathbf{AB} \times \mathbf{d}_1) \end{pmatrix} \end{aligned}$$

which is the direct (proceptual) Least-squares solution to Equation 2. Thus, two distinct non-parallel lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda \mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma \mathbf{d}_2$ in \mathbb{R}^3 intersect in a unique point P whose position vector is $\mathbf{r}_1(\lambda_2)$ or $\mathbf{r}_2(\gamma_2)$, where

$$\lambda_2 = \frac{\text{comp}_{\mathbf{d}_1 \times \mathbf{d}_2}(\mathbf{AB} \times \mathbf{d}_2)}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \quad \text{and} \quad \gamma_2 = \frac{\text{comp}_{\mathbf{d}_1 \times \mathbf{d}_2}(\mathbf{AB} \times \mathbf{d}_1)}{\|\mathbf{d}_1 \times \mathbf{d}_2\|}$$

if and only if the vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent. It is noteworthy that λ_2 and γ_2 in the above solution can be easily resolved using a determinant calculation together with that of the vector $\mathbf{d}_1 \times \mathbf{d}_2$ as

$$\lambda_2 = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_2 \\ \mathbf{d}_1 \times \mathbf{d}_2 \end{pmatrix} \quad \text{and} \quad \gamma_2 = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_1 \times \mathbf{d}_2 \end{pmatrix}.$$

The shortest vector between two skew lines

We now consider the procedure of finding a solution to the shortest vector between two skew lines in \mathbb{R}^3 by the *process* of matrix inversion (or Cramer's rule) incorporating the *concept* of the cross product and thus realising a proceptual solution.

To this end, consider the two skew lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda \mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma \mathbf{d}_2$ in \mathbb{R}^3 . Apart from directly finding the shortest distance between ℓ_1 and ℓ_2 , using the norm of the projection of \mathbf{AB} in the direction of the vector $\mathbf{d}_1 \times \mathbf{d}_2$, one also procedurally determines the shortest vector between ℓ_1 and ℓ_2 as follows. The general vector joining ℓ_1 and ℓ_2 is found algebraically as

$$\mathbf{r}_2 - \mathbf{r}_1 = \mathbf{b} - \mathbf{a} + \gamma \mathbf{d}_2 - \lambda \mathbf{d}_1 = \mathbf{AB} + \gamma \mathbf{d}_2 - \lambda \mathbf{d}_1.$$

The shortest vector between ℓ_1 and ℓ_2 is then the one which is perpendicular to both ℓ_1 and ℓ_2 . In effect this gives the two equations $(\mathbf{r}_2 - \mathbf{r}_1) \cdot \mathbf{d}_1 = 0$ and $(\mathbf{r}_2 - \mathbf{r}_1) \cdot \mathbf{d}_2 = 0$ realising a 2×2 system of linear equations in λ and γ . The approach is to solve for these two variables using Cramer's rule. Once λ and γ are found, the shortest vector is calculated. Although tedious, this approach also provides a way to the shortest distance between the two lines.

To develop a proceptual solution to the above problem, we now solve the dot product equations above for the arbitrary skew lines ℓ_1 and ℓ_2 . Lemma 1 in the preliminary section will be required. Now,

$$\begin{aligned} (\mathbf{r}_2 - \mathbf{r}_1) \cdot \mathbf{d}_1 = 0 &\Rightarrow \mathbf{AB} \cdot \mathbf{d}_1 + \gamma \mathbf{d}_2 \cdot \mathbf{d}_1 - \lambda \mathbf{d}_1 \cdot \mathbf{d}_1 = 0 \\ &\Rightarrow \mathbf{AB} \cdot \mathbf{d}_1 = \lambda \mathbf{d}_1 \cdot \mathbf{d}_1 - \gamma \mathbf{d}_2 \cdot \mathbf{d}_1 \end{aligned}$$

and

$$\begin{aligned} (\mathbf{r}_2 - \mathbf{r}_1) \cdot \mathbf{d}_2 = 0 &\Rightarrow \mathbf{AB} \cdot \mathbf{d}_2 + \gamma \mathbf{d}_2 \cdot \mathbf{d}_2 - \lambda \mathbf{d}_1 \cdot \mathbf{d}_2 = 0 \\ &\Rightarrow -\mathbf{AB} \cdot \mathbf{d}_2 = -\lambda \mathbf{d}_1 \cdot \mathbf{d}_2 + \gamma \mathbf{d}_2 \cdot \mathbf{d}_2 \end{aligned}$$

produces the matrix equation

$$\begin{pmatrix} \mathbf{d}_1 \cdot \mathbf{d}_1 & -\mathbf{d}_2 \cdot \mathbf{d}_1 \\ -\mathbf{d}_1 \cdot \mathbf{d}_2 & \mathbf{d}_2 \cdot \mathbf{d}_2 \end{pmatrix} \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = \begin{pmatrix} \mathbf{AB} \cdot \mathbf{d}_1 \\ -\mathbf{AB} \cdot \mathbf{d}_2 \end{pmatrix} \quad \dots \quad \text{Equation 3}$$

On closer inspection, the above Equation 3 is precisely

$$(A^T A) \begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = A^T \cdot (\mathbf{AB}^T) \quad \dots \quad \text{Equation 4}$$

where $A = \begin{pmatrix} \mathbf{d}_1^T & -\mathbf{d}_2^T \end{pmatrix}$ as before. Since ℓ_1 and ℓ_2 are skew their direction vectors, \mathbf{d}_1 and \mathbf{d}_2 respectively, are not parallel. Thus $\|\mathbf{d}_1 \times \mathbf{d}_2\| \neq 0$. If one considers the coefficient matrix in the above Equation 4, we have by Lagrange's identity

$$\det(A^T A) = \det \begin{pmatrix} \mathbf{d}_1 \cdot \mathbf{d}_1 & -\mathbf{d}_2 \cdot \mathbf{d}_1 \\ \mathbf{d}_1 \cdot \mathbf{d}_2 & \mathbf{d}_2 \cdot \mathbf{d}_2 \end{pmatrix} = -\|\mathbf{d}_1 \times \mathbf{d}_2\|^2 \neq 0.$$

Hence, we have the *same* unique solution

$$\begin{pmatrix} \lambda \\ \gamma \end{pmatrix} = (A^T A)^{-1} A^T (\mathbf{AB}^T) = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|} \begin{pmatrix} \text{comp}_{\mathbf{d}_1 \times \mathbf{d}_2}(\mathbf{AB} \times \mathbf{d}_2) \\ \text{comp}_{\mathbf{d}_1 \times \mathbf{d}_2}(\mathbf{AB} \times \mathbf{d}_1) \end{pmatrix}.$$

Cramer's rule could also be used with the use of Lemma 1 to produce the same result. Alternatively, the shortest vector between the skew lines ℓ_1 and ℓ_2 may be found using differential calculus by minimising the norm of the general joining vector $\mathbf{r}_2 - \mathbf{r}_1$.

In finding just the shortest distance between the skew lines, Srinivasan (1999) illustrates this procedure in an example, but presents no underlying theory. We now examine the underlying theory governing this procedure as follows:

Let $v(\lambda, \gamma) = \|\mathbf{r}_2 - \mathbf{r}_1\|^2$. Expanding,

$$v(\lambda, \gamma) = \|\mathbf{AB}\|^2 - 2\lambda(\mathbf{AB} \cdot \mathbf{d}_1) + 2\gamma(\mathbf{AB} \cdot \mathbf{d}_2) + \lambda^2\|\mathbf{d}_1\|^2 + \gamma^2\|\mathbf{d}_2\|^2 - 2\lambda\gamma(\mathbf{d}_1 \cdot \mathbf{d}_2).$$

Minimising v would produce λ and γ that will give the shortest vector between ℓ_1 and ℓ_2 . To this end we generate two equations

$$\frac{\partial v}{\partial \lambda} = 0 \quad \text{and} \quad \frac{\partial v}{\partial \gamma} = 0.$$

Then

$$\begin{aligned} \frac{\partial v}{\partial \lambda} = 0 &\Rightarrow -2(\mathbf{AB} \cdot \mathbf{d}_1) + 2\lambda\|\mathbf{d}_1\|^2 - 2\gamma(\mathbf{d}_1 \cdot \mathbf{d}_2) = 0 \\ &\Rightarrow \lambda(\mathbf{d}_1 \cdot \mathbf{d}_1) - \gamma(\mathbf{d}_1 \cdot \mathbf{d}_2) = \mathbf{AB} \cdot \mathbf{d}_1 \end{aligned}$$

and

$$\frac{\partial v}{\partial \gamma} = 0 \Rightarrow \lambda(\mathbf{d}_1 \cdot \mathbf{d}_2) - \gamma(\mathbf{d}_2 \cdot \mathbf{d}_2) = \mathbf{AB} \cdot \mathbf{d}_2$$

produces the same matrix Equation 3, the solution of which is determined above. Looking at the solution for the point of intersection (of the two lines), it turns out that the values of

$$\lambda_2 = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_2 \\ \mathbf{d}_1 \times \mathbf{d}_2 \end{pmatrix} \quad \text{and} \quad \gamma_2 = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_1 \times \mathbf{d}_2 \end{pmatrix}$$

are exactly the same as that of the parameters for the shortest vector between the two skew lines. In all, the linear dependence of the three vectors \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 is crucial in the orientation of the two non-parallel lines $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda\mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma\mathbf{d}_2$ in \mathbb{R}^3 . If these three vectors are linearly dependent, then the two lines intersect in a unique point. Otherwise the two lines are skew.

We summarise these findings in our concluding theorem.

Theorem 7: Let $\ell_1: \mathbf{r}_1 = \mathbf{a} + \lambda\mathbf{d}_1$ and $\ell_2: \mathbf{r}_2 = \mathbf{b} + \gamma\mathbf{d}_2$ be two non-parallel lines in \mathbb{R}^3 . Then

a) ℓ_1 and ℓ_2 intersect in a unique point P whose position vector is given by $\mathbf{r}_1(\lambda_2)$ or $\mathbf{r}_2(\gamma_2)$ if and

only if $\det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} = 0$ if and only if \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly dependent

b) ℓ_1 and ℓ_2 are skew and the shortest vector between ℓ_1 and ℓ_2 is given by

$$\mathbf{v}_2 = \mathbf{AB} + \gamma_2 \mathbf{d}_2 - \lambda_2 \mathbf{d}_1$$

if and only if $\det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} \neq 0$ if and only if \mathbf{AB} , \mathbf{d}_1 and \mathbf{d}_2 are linearly independent

where

$$\lambda_2 = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_2 \\ \mathbf{d}_1 \times \mathbf{d}_2 \end{pmatrix} \quad \text{and} \quad \gamma_2 = \frac{1}{\|\mathbf{d}_1 \times \mathbf{d}_2\|^2} \det \begin{pmatrix} \mathbf{AB} \\ \mathbf{d}_1 \\ \mathbf{d}_1 \times \mathbf{d}_2 \end{pmatrix}.$$

Returning to the Least-squares problem, when a Least-squares solution $\tilde{\mathbf{x}}$ is found for $A\mathbf{x} = \mathbf{b}$, the Euclidean norm $\|\mathbf{b} - A\tilde{\mathbf{x}}\|$ is the *Least-squares error*. In the inconsistent system Equation 2 (two skew

lines) the Least-squares error is the norm of the shortest vector $\|\mathbf{AB} + \gamma_2 \mathbf{d}_2 - \lambda_2 \mathbf{d}_1\|$ (the shortest distance between the two skew lines). Indeed, if the Least-squares error is zero then the lines intersect in a unique point.

Concluding remarks

In South Africa, the general perception is that secondary school teaching of mathematics tends to be fairly procedural and that students that enter university are better equipped to deal with procedural problems than with conceptual problems (Engelbrecht, Harding, & Potgieter, 2005). Teaching for procedural knowledge means teaching definitions, symbols and isolated skills in an expository manner without first focusing on building deep, connected meaning to support those concepts (Skemp, 1987). Teaching for conceptual knowledge, on the other hand, begins with posing problems that require students to reason flexibly. Through the solution process, students make connections to what they already know, thus allowing them to extend their prior knowledge and transfer it to new situations (National Council of Teachers of Mathematics, 2000).

This paper introduces proceptual solutions as a type of alternative solution to problems at tertiary level mathematics, the particular case of solving for the point of intersection of two lines and the shortest vector between two skew lines. Such proceptual solutions require a more efficient use of cognitive processing, as they require a cognitive switch between doing a process and thinking about a concept. We also propose and encourage a proceptual approach to mathematics instruction at tertiary level. Such a proceptual view, which in our example amalgamates the Gaussian elimination, Cramer's Rule and Least-squares solution processes and the dot product and cross product concepts, encourages a cognitive switch from a focus on the process of algorithmic manipulations of symbols to a deeper, conceptual understanding and thinking about the concepts. As proceptual thinking grows in conceptual richness, procepts can be manipulated as simple symbols at higher levels or scaffolded to perform computations, to be decomposed or recomposed at will. The unique features of the proceptual solution as described in this paper, suggest that it might be of benefit for mathematics educators to look for such solutions in the teaching and learning of mathematics.

We also mention that the direct (vector) solutions presented allow greater access to such type of problems and deeper applications of vector concepts in the sciences and engineering undergraduate Linear Algebra courses.

References

- Anton, H., & Rorres, C. (2005). *Elementary linear algebra, applications versions* (9th ed.). New York, NY: Wiley.
- Beesack, P. (1962). On the rank of a matrix. *Mathematics Magazine*, 35(2), 73-77.
- Craig, H. (1951). Vector analysis. *Mathematics Magazine*, 25(2), 67-86.
- Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced mathematical thinking* (pp. 95-123). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Edwards Jr., C.H., & Penny, D.E. (1988). *Elementary linear algebra*. Upper Saddle River, NJ: Prentice Hall.
- Engelbrecht, J., Harding, A., & Potgieter M. (2005). Undergraduate students' performance in procedural and conceptual mathematics. *International Journal of Mathematical Education in Science and Technology*, 36(7), 701-712. doi: 10.1080/00207390500271107
- Gray, E. M., & Tall, D. O. (1991). Duality, ambiguity and flexibility in successful mathematical thinking. In F. Furinghetti (Ed.), *Proceedings of 15th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 72-79). Assisi, Italy.
- Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity and flexibility: A "proceptual" view of simple arithmetic. *Journal for Research in Mathematics Education*, 25(2), 116-140. Available at <http://www.jstor.org/stable/749505>
- Lay, D. C. (2005). *Linear algebra and its applications* (3rd ed. update). Upper Saddle River, NJ: Pearson.
- Lipschutz, S., & Lipson, M. L. (2001). *Schaum's outlines of theory and problems of linear algebra* (3rd ed.). Columbus, OH: McGraw-Hill.

- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- Osborne, E. E. (1961). On least squares solutions of linear equations. *Journal of the ACM*, 8(4), 628-636. doi: 10.1145/321088.321100
- Piaget, J. (1972). *The principles of genetic epistemology*. (W. Mays trans.) London: Routledge & Keagan Paul.
- Salas, S. L., & Hille, E. (1995). *Calculus, one and several variables* (6th ed.). Hoboken, NJ: Wiley.
- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on process and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22(1), 1-36. doi: 10.1007/BF00302715
- Skemp, R. R. (1987). *The psychology of learning mathematics*. Hillsdale, NJ: Erlbaum.
- Smith, J., & Henderson, M. (1985). Tetrahedra, skew lines and volumes. *The College Mathematics Journal*, 16(2), 138-140. Available at <http://www.jstor.org/stable/2686218>
- Srinivasan, V. K. (1999). Skew lines in \mathbb{R}^3 . *International Journal of Mathematical Education in Science and Technology*, 30(5), 739-746. doi: 10.1080/002073999287734
- Tall, D. (1995). Cognitive growth in elementary and advanced mathematical thinking. In L. Meira & D. Carraher (Eds.), *Proceedings of the 19th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 61-75). Recife, Brazil: Universidade Federal de Pernambuco.
- Tall, D., Gray, E., Bin Ali, M., Crowley, L., Demarois, P., McGowen, M., ... Yusof, Y. (2001). Symbols and the bifurcation between procedural and conceptual thinking. *Canadian Journal of Science, Mathematics and Technology Education*, 1(1), 81-104. doi: 10.1080/14926150109556452