Chapter 7
Revising a Draft

Simply go through what you have written and
try to curb the length of sentences,
question every passive verb and if possible make it active,
prune redundant words, and look for nouns used instead of verbs.

— BERNARD DIXON, *Sciwrite* (1973)

Every single word that I publish I write at least six times.

— PAUL R. HALMOS, Interview in *Paul Halmos: Celebrating* 50 *Years of Mathematics* (1991)

Two pages of the final manuscript . . .
Although they look like a first draft,
they had already been rewritten and retyped—like
almost every other page—four or five times.
With each rewrite I try to make what I have
written tighter, stronger and more precise,
eliminating every element that is not doing useful work.
Then I go over it once more, reading it aloud, and am always
amazed at how much clutter can still be cut.

— WILLIAM ZINSSER, *On Writing Well* (1990)

When one has finished a substantial paper there is commonly a
mood in which it seems that there is really nothing in it.
Do not worry, later on you will be thinking
"At least I could do something good then."

— J. E. LITTLEWOOD, *Littlewood's Miscellany*[14] (1986)

[14]See [34].

7.1. How to Revise

All writing benefits from revision. Your first attempt can always be made clearer, more concise, and more forceful. Effective revision is a skill that is acquired through practice. A prerequisite is an understanding of the principles of writing discussed in the previous chapters, together with the ability to spot when these principles have been violated and the skill and courage to rectify the weaknesses.

Put a draft aside for a few days, so that you can examine it with a fresh mind. Then analyse the draft in different ways. Read it aloud and listen to the rhythm. Read it at high speed. Does the text flow? Are the sentences of varying length? Read at the page level, focusing on the shape and the density of ink on the page (assuming the draft has been typeset). Is there a good balance between equations and text? Does the paper look inviting? Do the key ideas stand out? Knuth [164, p. 3] notes that "Many readers will skim over formulas on their first reading of your exposition. Therefore, your sentences should flow smoothly when all but the simplest formulas are replaced by 'blah' or some other grunting noise."

Most good writers work through several drafts—see the quotations at the beginning of the chapter. Some writers adopt the spiral or factorial technique of writing, whereby sections or chapters are written and revised in the order 1, 2, 1, 2, 3, 1, 2, 3, 4, ... Halmos [121] describes this as "the best way to start writing, perhaps the only way." But it is not always necessary to start at the beginning: for example, you can leave the introduction to last and begin by writing a numerical examples section or your conclusions. Every paper is different and versatility is the key.

Some writers have the ability to produce first drafts that need little revision. The amount of revision required depends partly on the stage in the research process at which the writing is begun. As mentioned in Chapter 1, writing can be an integral part of the research process.

I make no distinction between editing, revising and rewriting in my own work. I begin with the intention of making whatever changes are necessary to produce the best exposition I can. This may involve minor deletions, insertions and reorderings, or it may require complete rewrites of paragraphs and sections. As an extreme example, I deleted the original first paragraph of this chapter and started again from scratch. Although I made full use of computers in writing this book, I did most of my writing and revising at my desk by writing on a blank sheet or a printout of a draft; only occasionally did I compose at the keyboard. When marking corrections, I use a different colour for each pass through the document and my printouts usually finish up covered in coloured ink.

Write or print your drafts with wide margins and with a line spac-

ing big enough for you to write between the lines. (In LaTeX, when the `article` style is used, a suitable line spacing can be achieved by putting the command `\def\baselinestretch{1.3}` before the `\begin{document}` command.) This will give you a lot of space to mark revisions, which in turn will encourage you to make them.

Figure 7.1 contains a check-list of questions to ask yourself when revising your writing.

Many of the points discussed above and in previous chapters are illustrated in the extracts in the rest of this chapter, most of which are taken from the mathematics and computer science literature. Minor changes have been made to protect the guilty.

7.2. Examples of Prose

▷ GraphX—An experimental graphics development environment research project

Comments.
This title has the dubious distinction of having five nouns in succession, the first four used adjectivally. It contains three abstract *-al* or *-ent* words. This is not a title to arouse interest.

▷ This type of response surface determination produces information describing not only the variable coordinates for the optimum but also indicates the nature of the response surface in the domain of interest.

Comments.
The "not only" and "also" clauses do not match. The simplest solution is to delete "indicates". A possible improvement is "The information produced by this type of response surface determination describes not only the variable coordinates for the optimum but also the nature of the response surface in the domain of interest." I would also want to reword the clumsy "response surface determination".

▷ Thus there arises the question of whether a central database can be created in a manner which is consistent with the need for efficiency and breadth of scope.

Comments.
(1) *Thus there arises the question* can be replaced by the shorter and more direct *The question arises.*
(2) The which is a wicked one (see §4.14). In fact, the "which is" can be deleted.

▷ Can any words, phrases or sentences be deleted without loss of force or meaning (every such unit must earn its place)?

▷ Can you replace any long, woolly words with short, pithy ones (for example, require → need, utilize → use)?

▷ Are sentences and paragraphs in the best order, or should they be reordered or restructured?

▷ Are there any ambiguities or blatantly wrong words (such as *factor* instead of *feature*)?[a]

▷ Is there any unnecessary repetition?

▷ Can you convert a sentence from the passive to the active voice?

▷ Is every claim fully supported?

▷ Are the mathematical arguments and equations correct?

▷ Is the notation consistent? Can it be simplified or made more logical?

▷ Have quotations, references and numerical results been copied into the paper correctly?

▷ Is due reference made to the work of other authors (beware of "citation amnesia")?

▷ Are equations, results and the reference list properly numbered? Are all the cross-references and bibliographic citations correct?

[a]A draft of this book contained the phrase "chop sentences mercifully" instead of "chop sentences mercilessly"!

Figure 7.1. Check-list for revising.

▷ *Abstract*—This paper discusses the aims and methods of the FTEsol project and in this context discusses the architecture and design of the control system being produced as the focus of the project.

Comments.

(1) The phrases *in this context* and *as the focus of the project* serve no useful purpose and can be deleted.

(2) As mentioned in §6.6, it is not a good idea to begin an abstract with *This paper.* A complete rewrite is needed: "The aims and methods of the FTEsol project are discussed, together with the architecture and design of the control system being produced." This version avoids the repeated *discusses* in the original. A more direct alternative, if the use of *we* is allowed, is "We discuss the aims and methods of the FTEsol project ...."

▷ `Nsolve` finds numerically the 5 complex solutions to this fifth-order equation.

Comments.

There is a lack of parallelism in *5* and *fifth*, which the reader may find disturbing. I would write "the five complex solutions to this fifth-order equation".

▷ Tables give a systematic and orderly arrangement of items of information. Tabular layout has the particular virtue of juxtaposing items in two dimensions for easy comparison and contrast. Tables eliminate tedious repetition of words, phrases and sentence patterns that can instead be put at the tops of columns and the sides of rows in the table. Although tables do not make much impact by visual display, it is possible, by careful arrangements, to emphasize and highlight particular items or groups of information.

Comments.

This is the first paragraph of a section on tables from a manual on technical writing. The first sentence reads like a dictionary definition and is surely not telling readers anything they do not know already. The ideas in the third and fourth sentences are not clearly expressed. Revised version:

> Tables juxtapose items in two dimensions for easy comparison and contrast. Their row and column labels save the tedious repetition of words that would be necessary if the information were presented in textual form. Although tables lack the visual impact of

graphs, information can be grouped and highlighted by careful use of rules and white space.

▷ In order to pinpoint the requirements for an effective microchip development environment sufficiently to definitively obtain answers to the above questions, it is essential to be able to interview a wide variety of microchip developers who are knowledgeable and experienced in such matters.

Comments.

(1) *In order* is superfluous.

(2) *To definitively obtain* is a split infinitive (to-adverb-verb) that is more naturally written as *to obtain definitive answers.* It is probably better to replace the whole phrase by *to answer.*

(3) *Essential to be able to* is probably better replaced by *necessary to.*

▷ Due to the advances in computer graphics and robotics, a new interest in geometric investigations has now arisen within computer science focusing mainly on the computational aspects of geometry, forming the research field known as computational geometry.

Comments.

A new interest ... has now arisen is a passive phrase that is easily improved. At the same time, we can remove the awkward and incorrect *due to* (see §4.14). The phrase *focusing ... aspects of geometry* seems unnecessary. My revised version:

> Advances in computer graphics and robotics have stimulated a new interest in geometric investigation within computer science, forming the research field known as computational geometry.

▷ In terms of fractals, a straight line has a dimension of one, an irregular line has a dimension of between one and two, and a line that is so convoluted as to completely fill a plane has a dimension approaching the dimension of the plane, namely a dimension of two. Fractal dimensions assign numbers to the degree of convolution of planar curves.

Comments.

This extract is taken from a model paper given in a book about writing a scientific paper. The previous two sentences had also been about fractals. The phrase "in terms of fractals" is unnecessary and the order of these two sentences should be

reversed. "Fractal dimensions" do not "assign numbers" but *are* numbers. It is not clear, grammatically, whether the "namely a dimension of two" applies to the convoluted line or the plane. The split infinitive "to completely fill" can be replaced by "to fill". My rewritten version:

> Fractal dimensions are numbers that measure the degree of convolution of planar curves. A straight line has dimension one, an irregular line has a dimension between one and two, and a line that is so convoluted as to fill a plane has a dimension approaching two, which is the dimension of the plane.

▷ Data flow analysis determines the treatment of every parameter and COMMON variable by every subprogram with sufficient precision that nonportable parameter passing practices can be detected.

Comments.

This sentence is difficult to understand on its first reading. The phrase *by every subprogram* delays the punch-line and the *every*s also delay comprehension. It is not clear whether "with sufficient precision" applies to the subprograms or the analysis. In fact, the latter was intended. A better version is "Data flow analysis determines the treatment of parameters and COMMON variables with sufficient precision that nonportable parameter passing practices can be detected."

▷ [12] reports an eigenproblem from the automobile industry. The eigenvalues of interest are those ones having real part greater than zero.

Comments.

A citation makes a weak start to a sentence and jolts the eye. In the second sentence *ones* is unnecessary and the sentence can be made more concrete. Revised version:

> Jones reports an eigenproblem from the automobile industry [12]. The eigenvalues of interest are those lying in the right half plane.

▷ Command names have been defined as two letter sequences because it is believed that users prefer to avoid verbosity.

Comments.

Who is the believer? *It is believed* is better replaced by *we believe*, or *evidence suggests*, or *our experience has shown*, preferably with a reference. A hyphen is needed in *two-letter sequences*, otherwise the meaning is *two sequences of letters.*

▷ When dealing with sets of simple figures, a basic problem is the determination of containment relations between elements of the set.

Comments.

When dealing is a dangling participle: we are not told who is dealing with the sets. A better version results on omitting *when dealing* and changing *the determination of* to *to determine.* Even better is

> A basic problem for sets of simple figures is to determine containment relations between elements of the set.

▷ In the function, $\exp(x)$ is first tested for overflow. If it does, then `inf` is returned.

Comments.

"If it does" refers to "tested for overflow" rather than "overflows", which the intended meaning requires. The second sentence can be replaced by "If overflow is detected then `inf` is returned."

▷ It is anticipated that the early versions of the system will provide definitive enough information that it will be reasonable to design with some assurance a variety of other systems which should be broader in scope.

Comments.

(1) *Anticipated* should be *expected.* This usage is described as "avoided by careful writers and speakers of English" in the *Collins English Dictionary* [60] but is so frequent that it may one day become accepted. Anticipate means to take action against (*The enemy had anticipated our move*).

(2) There are not degrees of definitiveness.

(3) *Reasonable* should probably be *possible.*

(4) There is a wicked which. The last phrase could be replaced by *systems of broader scope.*

▷ As far as the minimum eigenvalues of the other boundary element matrices are concerned they can only be small if the value of k is close to a corresponding element of S_Γ.

Comments.
"As far as ... is/are concerned" is a phrase more appropriate to speech than writing. This sentence can be shortened considerably: "The minimum eigenvalues of the other boundary element matrices can be small only if k is close to a corresponding element of S_Γ."

▷ This object is achieved by utilizing a set of properties which the signal is known or is hypothesized as possessing.

Comments.
This passive sentence can be shortened and made active: "We achieve this aim by using properties that we know or hypothesize the signal to possess." This simplified form shows that the sentence says little. The paragraph in which the sentence appears should be rewritten.

▷ The *main* purpose of any scientific article is to convey in the fewest *number of* words the ideas, procedures and conclusions of an investigator *to the scientific community.* Whether *or not* this *admirable* aim is accomplished depends *to a large extent* on how skillful the author is in *assembling the words of* the *English* language.

Comments.
These are the opening two sentences of a medical journal editorial titled "Use, Misuse and Abuse of Language in Scientific Writing". The italics are those of Gregory [117], who points out that the italicized words can be omitted without loss of meaning.

▷ Mathematica was found to be a suitable environment in which to perform the computational experiments.

Comments.
The following rewrite is much shorter, avoids the passive voice, and takes for granted the "suitability": "We carried out our computational experiments in Mathematica."

▷ These observations simply imply that nearby orbits separate from the orbit of γ after many iterations of the map G. Hardy and Littlewood (1979) prove a classical theorem that is useful in this context.

Comments.

There is nothing intrinsically wrong with this extract, but the astute writer may wish to rewrite to remove two minor infelicities. First, the near repetition in the phrase "simply imply" could distract the reader. Second, the eye naturally tends to read "iterations of the map G. Hardy"—symbols can cause trouble at the end of a sentence as well as at the beginning!

7.3. Examples Involving Equations

▷ *Theorem.* Let $A \in \mathbb{R}^{m\times n}$ be a given matrix, and let $A = U\Sigma V^T$ be a singular value decomposition of A. Then the problem

$$\max\{Re\ \text{trace}(AQ) :\ Q \in \mathbb{R}^{n\times n} \text{ is orthogonal}\}$$

has the solution $Q = VU^T$, and the value of the maximum is $\sigma_1(A) + \cdots + \sigma_n(A)$, where $\{\sigma_i(A)\}$ is the set of singular values of A.

Comments.

The statement of this theorem is clear and readable but it can be improved. The first sentence can be shortened to "Let $A \in \mathbb{R}^{m\times n}$ have the singular value decomposition $A = U\Sigma V^T$." The second sentence can be rewritten as "Then

$$\max\{Re\ \text{trace}(AQ) :\ Q \in \mathbb{R}^{n\times n} \text{ is orthogonal}\} = \sum_{i=1}^{n} \sigma_i,$$

where the σ_i are the singular values of A, and the maximum is attained when $Q = VU^T$." This is shorter and, arguably, easier to read. Even better is to define the σ_i at the beginning by appending "where $\Sigma = \text{diag}(\sigma_i)$" to the first sentence.

▷ Writing $\theta_i = (E|x| + f)_i$ we have

$$\rho(x) \le \| \, |A^{-1}| \, \text{diag}(\theta_i) \, \|_\infty.$$

Using properties of the ∞-norm we obtain

$$\rho(x) \le \| \, |A^{-1}|(E|x| + f) \, \|_\infty.$$

Comments.

The presentation hides the fact that the upper bound in the second inequality is equal to that in the first. The following rewrite is more informative:

Writing $\theta_i = (E|x| + f)_i$, and using properties of the ∞-norm, we obtain

$$\begin{aligned}\rho(x) &\leq \| \, |A^{-1}| \, \mathrm{diag}(\theta_i) \, \|_\infty \\ &= \| \, |A^{-1}|(E|x| + f) \, \|_\infty.\end{aligned}$$

▷ Let A be $n \times n$. Show that if for any Hermitian matrix H, $\mathrm{trace}(HA) = 0$, then $A = 0$.

Comments.

This question is ambiguous because *any* can mean *whichever* or *at least one* in everyday English. As Halmos [121] recommends, *any* should always be replaced by *each* or *every* in mathematical writing.

▷ Suppose now that the assumption $a = 1$ fails.

Comments.

An assumption does not fail or succeed: it is either invalid or valid. Better wording might be *Suppose now that the condition $a = 1$ is not satisfied*, or *Suppose now that $a \neq 1$.*

▷ **Introduction**

Throughout this paper $\|A\|_F$ denotes the Frobenius norm $\left(\sum_{i,j} a_{ij}^2\right)^{1/2}$ of a real-valued matrix A and A^+ denotes the Moore–Penrose pseudo-inverse of A. We define the set

$$\mathcal{U}_+ = \{ \, U \in \mathbb{R}^{n \times n} : U^T U = I, \ \det(U) = 1 \, \}.$$

Given $A, B \in \mathbb{R}^{m \times n}$ we are interested in the *orthogonal Procrustes problem* in its pure rotation form:

$$\min_{U \in \mathcal{U}_+} \|A - BU\|_F.$$

Comments.

This is a weak start. It can be improved by stating the problem in the first sentence. We can delay the definition of A^+ until it is first used. "Real-valued" can be shortened to "real", but we can dispense with it altogether by using the $\mathbb{R}^{m \times n}$ notation. The "Given" phrase does not read well—interest in the mathematical problem is surely independent of being given matrices A and B. Revised first sentence:

The pure rotation form of the *orthogonal Procrustes problem* is

$$\min_{U \in \mathcal{U}_+} \|A - BU\|_F, \qquad A, B \in \mathbb{R}^{m \times n},$$

where

$$\mathcal{U}_+ = \{\, U \in \mathbb{R}^{n \times n} : U^T U = I,\ \det(U) = 1 \,\}$$

and $\|A\|_F = \left(\sum_{i,j} a_{ij}^2\right)^{1/2}$ is the Frobenius norm.

The writer should now go on to mention applications in which this problem arises, explain what is known about its solution, and state the purpose of the paper.

▷ *Theorem.* Let A be an $n \times p$ complex matrix with rank p. We define the $p \times p$ positive definite Hermitian matrix $S = A^* A$ and the $n \times p$ matrix $\overline{Q} = AS^{-\frac{1}{2}}$. Let $\mathcal{U}$ be the set of all $n \times p$ orthonormal matrices. Then the following is true:

$$\overline{Q} \in \mathcal{U} \qquad \text{and} \qquad \|A - \overline{Q}\|_2 \leq \min_{Q \in \mathcal{U}} \|A - Q\|_2.$$

Comments.

(1) The rank assumption implies $p \leq n$. It is clearer, though less concise, to say "Let A be an $n \times p$ complex matrix with $n \geq p = \operatorname{rank}(A)$."

(2) There is no need to introduce the symbol S. If the existence of $(A^*A)^{-1/2}$ is thought not to be obvious (note the preferable slashed fraction in the exponent), it can be established in the proof. The symbol $\mathcal{U}$ can also be dispensed with.

(3) The inequality is an equality. The second sentence onwards can be simplified as follows: Let $\overline{Q} = A(A^*A)^{-1/2}$. Then $\overline{Q}$ is orthonormal and

$$\|A - \overline{Q}\|_2 = \min\{\, \|A - Q\|_2 : Q^T Q = I_p \,\}.$$

▷ The optimality of the constant $\pi \log n/4$ in inequality (14.3) is due to Smith [10].

Comments.

This sentence suggests that Smith made the constant optimal. Better is *was first established by Smith.*

▷ (From an abstract:) The bound is derived in the case of k $(0 \leq k \leq p)$ explanatory variables measured with error.

Comments.
The intrusive inequalities and the all-purpose phrase "in the case of" can be removed, and the reader told, or reminded, what p is, by writing "The derivation of the bound allows for any k of the p explanatory variables to be measured with error."

7.4. Examples from My Writing

Here are some examples from my own writing of how I improved drafts.

(1) Original first two sentences of paper:

▷ Summation of floating point numbers is a ubiquitous and fundamental operation in scientific computing. It is required when evaluating inner products, means, variances, norms, and all kinds of functions in nonlinear problems.

Improved version (shorter, less passive, more direct):

> Sums of floating point numbers are ubiquitous in scientific computing. They occur when evaluating inner products, means, variances, norms, and all kinds of nonlinear functions.

An alternative (avoids the dangling participle "when evaluating", at the cost of a more passive construction):

> Sums of floating point numbers occur everywhere in scientific computing: in the evaluation of inner products, means, variances, norms, and all kinds of nonlinear functions.

(2) Original:

▷ Here, instead of immediately feeding each correction e_i back into the summation, the corrections are accumulated as $e = \sum_{i=1}^{n} e_i$ (by recursive summation) and then the global correction e is added to the computed sum.

Improved version (omits the unnecessary mathematical notation):

Here, instead of immediately feeding each correction back into the summation, the corrections are accumulated by recursive summation and then the global correction is added to the computed sum.

(3) Original first sentence of abstract:

▷ If a stationary iterative method is used to solve a linear system $Ax = b$ in floating point arithmetic how small can the method make the error and the residual?

Improved version (avoids the misleading *if*, more direct):

How small can a stationary iterative method for solving a linear system $Ax = b$ make the error and the residual in the presence of rounding errors?

(4) Original first sentence of paper:

▷ A *block algorithm* in matrix computations is one that is defined in terms of operations on submatrices rather than matrix elements.

A copy editor removed the words *one that is*. This changes the meaning, since the sentence now states a property rather than gives a definition, but I felt that the shorter sentence was an improvement.

7.5. A Revised Proof

Gershgorin's theorem, a well-known theorem in numerical linear algebra, specifies regions in the complex plane in which the eigenvalues of a matrix must lie. Here is the theorem and a proof that is correct, but can be improved.

Theorem 1 (Gershgorin, 1931) *The eigenvalues of $A \in \mathbb{C}^{n\times n}$ lie in the union of the n disks in the complex plane*

$$D_i = \{\, z \in \mathbb{C} : |z - a_{ii}| \le \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \,\}, \qquad i = 1, \ldots, n.$$

Proof. The proof is by contradiction. Let λ be an eigenvalue of A and x an eigenvector associated with λ and assume that $\lambda \notin D_i$ for $i = 1, \ldots, n$. Then from

$$Ax = \lambda x$$

and

$$(\lambda I - A)x = 0$$

we have

$$(\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j, \qquad i = 1, \ldots, n.$$

Taking absolute values and applying the triangle inequality gives

$$|\lambda - a_{ii}||x_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}||x_j|, \qquad i = 1, \ldots, n.$$

Assume that $|x_k| = \max_i |x_i|$. Then, dividing the kth inequality by $|x_k|$, we have

$$|\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}||x_j|/|x_k| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}|,$$

showing that λ is contained in the disk $\{ \lambda : |\lambda - a_{kk}| \leq \sum_{j \neq k}^{n} |a_{kj}| \}$, which is a contradiction. □

Several different proofs of Gershgorin's theorem exist, and this one is the most elementary and direct. The presentation of the proof has several failings, though.

1. The proof is too detailed and contains some unnecessary equations and inequalities.

2. The proof by contradiction is unnecessary, since the assumption that is to be contradicted is not used in the proof.

3. The "assumption" on k is really a definition of k and is clearer if phrased as such.

4. In the last line of the proof the disk can be described by its name, D_k.

The following proof uses the same reasoning but is much more concise and no less clear.

Proof. Let λ be an eigenvector of A and x a corresponding eigenvector, and let $|x_k| = \max_i |x_i|$. From the kth equation in $Ax = \lambda x$ we have

$$\sum_{\substack{j=1 \\ j \neq k}}^{n} a_{kj} x_j = (\lambda - a_{kk}) x_k.$$

Hence

$$|\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |a_{kj}||x_j|/|x_k|,$$

and since $|x_j|/|x_k| \leq 1$ it follows that λ belongs to the kth disk, D_k. □

Of course a proof has to be written with the intended audience in mind. For an undergraduate text the revised proof is probably too concise and some intermediate steps could be added.

7.6. A Draft Article for Improvement

Below is a shortened version of an article that I wrote for an undergraduate mathematics magazine. I have introduced over twenty errors of various kinds, though most are relatively minor. How many can you spot? If you are an inexperienced writer, criticizing this "draft" will be a valuable exercise.

Numerical Linear Algebra in the Sky

In aerospace computations, transformations between different co-ordinate systems are accomplished using the *direction cosine matrix* (DCM), which is defined as the solution to a time dependent matrix differential equation. The DCM is 3×3 and exactly orthogonal, but errors in computing it lead to a loss of orthogonality. A simple remedy, first suggested in a research paper in 1969 is to replace the computed DCM by the nearest orthogonal matrix every few steps. These computations are done in real-time by an aircrafts on-board computer so it is important that the amount of computation be kept to a minimum. One suitable method for computing a nearest orthogonal matrix is described in this Article. We begin with the case of 1×1 matrices—scalars.

(a) Let x_1 be a nonzero real number and define the sequence:

$$x_{k+1} = \frac{1}{2}(x_k + 1/x_k), \qquad k = 1, 2, \ldots, .$$

If you compute the first few terms on your calculator for different x_1 (e.g. $x_1 = 5$, $x_1 = -3$) you'll find that x_k converges to ± 1; the sign depending on the sign of x_1. Prove that this will always be the case (*Hint:* relate $x_{k+1} \pm 1$ to $x_k \pm 1$ and then divide this two relations). This result can be interpreted as saying that the iteration computes the nearest real number of modulus one to x_1.

(b) This scalar iteration can be generalized to matrices without loosing it's best approximation property. For a given nonsingular $X_1 \in \mathbb{R}^{n \times n}$ define

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-T}) \tag{1}$$

(This is one of those very rare situations where it really is necessary to compute a matrix inverse!) Here, X_k^{-T} denotes the transpose of the inverse of X_k. Natural questions to ask are: Is the iteration well defined (i.e., is X_k always nonsingular)? Does it converge? If so, to what matrix?

To investigate the last question suppose that $X_k \to X$. Then X will satisfy $X = \frac{X+X^{-T}}{2}$, or $X = X^{-T}$, or

$$X^T X = I,$$

thus X is orthogonal! Moreover, X is not just any orthogonal matrix. It is the nearest one to X_1 as shown by the following

Theorem 1

$$\|X - X_1\| = \min\{\|Q - X_1\| : Q \in R^{nxn},\ Q^T Q = I_n\}$$

where the norm is denoted by

$$\|X\| = \left(\sum_{i}^{n} \sum_{j=1}^{n} x_{ij}^2 \right)^{1/2}.$$

This is the matrix analogue of the property stated in (a).

Returning to the aerospace application, the attractive feature of iteration (1) is that if we don't wait to long before "re-orthogonalising" our computed iterates then just one or two applications of the iteration (1) will yield the desired accuracy with relatively little work. □

Here are the corrections I would make to the article. (In repeating the exercise myself some time after preparing this section, I could not find all the errors!)

1. First paragraph: hyphenate *time-dependent*; comma after 1969; *aircraft's*; comma after *on-board computer*; *article* in lower case.

2. Second paragraph: no colon after *sequence*. In display, $\frac{1}{x_k}$ instead of $1/x_k$, and replace "...,." by "....". Comma after *e.g.* and instead of semicolon after ± 1. *These two relations*; *modulus* 1.

3. Third paragraph: *losing*; *its*. Right parenthesis in display (1) is too large and full stop needed at end of display.

4. Fourth paragraph: third X should be in mathematics font, not roman; $(X+X^{-T})/2$. The equation $X^TX = I$ should not bc displayed and it should be followed by a semicolon instead of a comma. (The spacing in X^TX should be tightened up—see page 192 for how to do this in LaTeX.) Comma after X_1; *following theorem*.

5. No need to number the theorem as it's the only one. It should begin with words: *The matrix* X_1 *satisfies*. $\mathbb{R}^{n\times n}$ instead of R^{nxn} (two changes). Comma at end of first display. "I_n" is inconsistent with "I" earlier: make both I. *Denoted* should be *defined*. In first sum of second display, $i = 1$. The parentheses are too large.

6. Last paragraph: *too long*. Wrong opening quotes. For consistency with *generalized* (earlier in article), spell as *re-orthogonalizing*. Logical omission: I haven't shown that the iteration converges, or given or referred to a proof of the theorem.