The Role of Proof in Comprehending and Teaching Elementary Linear Algebra*

Frank Uhlig
Department of Mathematics
Auburn University
Auburn, AL 36849–5310, USA
uhligfd@auburn.edu
www.auburn.edu/~uhligfd

Abstract

We describe how elementary Linear Algebra can be taught successfully while introducing students to the concept and practice of "mathematical proof".
This is done badly with a sophisticated Definition–Lemma–Proof–Theorem–Proof–Corollary (DLPTPC) approach; badly – since students in elementary Linear Algebra courses have very little experience with proofs and mathematical rigor.
Instead, the subjects and concepts of Linear Algebra can be introduced in an exploratory and fundamentally reasoned way. One seemingly successful way to do this is to explore the concept of solvability of linear systems first via the row echelon form (REF). Solvability questions lead to row and column criteria for a REF that can be used repeatedly to: compute subspaces, settle linear (in)dependence, find inverses, perform basis change, compute determinants, analyze eigensystems etc. If these subjects are explained heuristically from the first principles of linear transformations, linear equations, and the REF, students experience the power of a concept–built approach and reap the benefit of deep math understanding. Moreover, early "salient point" proofs lead to an intuitive understanding of "math proof".
Once the basic concept of 'proof' is ingrained in students, more abstract proofs, even DLPTPC style expositions, on normal matrices, the SVD etc. become accessible and understandable to sophomore students. With the help of this gentle early approach, the concept and construct of a "math proof" becomes firmly embedded in the students' minds and helps with future math courses and general scientific reasoning.

Keywords : ternary education, mathematics education, teaching linear algebra, mathematical proof, exploration

1 Student background

A typical beginning college student in the US has encountered mathematical proofs only in the middle or early high school years with elementary proofs in planar geometry such as with congruent and similar triangles, transversals of parallel lines etc. A typical student will enter college without ever having seen

*To appear in *Educational Studies in Mathematics*

or thought of simple proofs such as that $\sqrt{2} \notin \mathbb{Q}$ or that there are infinitely many primes. Even if he/she has taken calculus in high school or college, there were likely no proofs in the course.

2 Proving in the history of mathematics

The recent history of mathematics is collected in our math journals. They span the last 200 to 300 years. Before these journals were founded, there were individual books and treatises, but very few in number. The math journals of the last 200 plus years particularly exhibit how mathematicians have grappled with "proofs" and mathematical rigor (besides the subject matter). Math papers prior to around 1850 – at least in my opinion – do not generally give tangible proofs. Old papers often flow on without statements of "Theorems", without giving clear definitions, and without full statements of the assumptions. They most often give correct but informal deductions and state the results rather vaguely. As an example, the matrix papers of Cauchy of around 1815 are extremely hard to read, to comprehend, and to assess. Gauss' 1799 thesis on the Fundamental Theorem of Algebra is almost incomprehensible as far as mathematical details are concerned.

By around 1870, mathematical journals begin to contain more and more papers with clear definitions, they begin to list assumptions, and contain – what we now call – 'formal proofs'. Rigor takes hold in the mathematical literature slowly. By around 1920, the modern formalities of "proving" are finally established. Almost universally from then on, a "Definition–Lemma–Proof–Theorem–Proof–Corollary" (DLPTPC) sequence of deductions is necessary for publishing mathematics research papers.

3 Elementary linear algebra in the undergraduate curriculum

Elementary linear algebra is usually studied in the second or third year at US colleges. It is intended mainly for science, math, and engineering majors. It generally follows after the calculus sequence and often parallels a first differential equations course. As our current calculus textbooks develop very few proofs – which are often skipped in class – the students of a first linear algebra course generally have had no experience with math proofs for many years. They enter our classes without a notion of what a 'proof' is, or what it means and does. Yet many upper division math teachers suggest, push for, and demand – probably quite rightfully so – that elementary linear algebra be a "proving course".

The first linear algebra course can answer this desire and serve as a transition and introduction to the modern culture of mathematics and its rigor. This transition should be effected in a gentle and subtle way, however. It requires reflection upon our teaching and respect for our students' state of mind when they enter our class. Trying to make linear algebra a proving course has led me to start off by exploring intuitively, and thereby to establish a mental and philosophical appreciation and a student-felt need, as well as an intellectual understanding of the necessity and beauty of proofs in the students' minds.

Linear algebra is the first math course encountered by our students that is highly conceptual in nature. Compared to linear algebra, the calculus and elementary differential equations courses study mostly methods, recipes, and formulas, but not many concepts, though they could. Modern elementary qualitative differential equations textbooks such as [5] or [1], are beginning to bridge the gap towards conceptual understanding and teaching that the first linear algebra course has had to deal with from the start due to its inherently conceptual subject matter.

4 Ways to start proving in an elementary linear algebra course

One of the most important aspects and achievements of modern mathematics is the development of 'formal proof'. This 19^{th} century cultural advance of modern man is often not fully appreciateded, not even by mathematicians. Therefore one should not be surprised that trying to bring this relatively recent intellectual development to students today is a very delicate and difficult task for both the teacher and the students. Just think of the length and depth of the mathematical awareness process associated with understanding and practicing formal proofs. It took eminent mathematicians almost a century to achieve this consciously.

In a first linear algebra course that wants to bring students into a mathematical proving frame of mind, we consider it best to start like the mathematicians of the pre 1850s and only familiarize students with intuitive mathematical reasoning in the first half or two thirds of the course. This works hand in hand with the subject matter of the course as we shall see. Apparently one can not succeed by one day starting out with a rigorous DLPTPC sequence of math presentations. It is better to prepare the students slowly for proofs and mathematical rigor.

We have fared well with and suggest to begin with exploring the subjects intuitively with questions such as :

<u>W</u>hat happens if?

<u>W</u>hy does it happen?

<u>H</u>ow do different cases occur?

<u>W</u>hat is true here?

If these questions have been explored in their depth for one specific subject, then we can collect the gained knowledge in "<u>T</u>heorems". Such a WWHWT sequence of presentation quickly leads students to understand, construct, reason through, enjoy, and actually demand "salient point" type proofs. It prepares them mentally and emotionally for DLPTPC sequences of presentations later on in this class and in subsequent math classes. A subject specific "What, Why, and How?" sequence of exploratory questions generally gives students a deep conceptual understanding because this enforces the first principles of linear algebra and gives them the tools to master the subject matter, see [10]. Such an exploratory approach is also used in modern elementary qualitative differential equations textbooks, such as [5] or [1], with good success. A deep level of understanding can be achieved gently with WWHWT, while an early DLPTPC approach satisfies noone, neither student nor instructor.

What we have proposed and done in [9] for a "linear algebra and introduction to proving" course, is at first to return to the historical proving mode of old, where heuristics and intuition were deemed enough to establish a result. A good start for introducing 'proofs' by using salient point type arguments is to discuss the solvability of linear equations before actually teaching how to solve them. This presupposes a familiarity with the mechanics of row reduction and the properties of a row echelon form (REF).

A most enlightening discussion usually follows the decidedly naive question:

"*Do you know or can you find an equation, any mathematical equation, that cannot be solved by anyone?*"

Students will fall silent at first. Then they will want to know who 'anyone' is supposed to be. I usually answer: someone like you, or me, or Einstein, Gauss, or even God, as long as this 'anyone' follows the laws and rules of mathematics. Once they realize that we are looking for a simple equation that cannot be solved at all in the framework of mathematics, the suggestions begin to fly: $x^2 + 1 = 0$, ... until we arrive at the most elementary unsolvable equation $0 \cdot x = 1$ or $0 \cdot y = -7$. The students recognize this as a linear equation. Its augmented REF $(0 \mid \boxed{*}\,)$ has one pivot, while its 'system matrix' $A = (0)$ has none. Through this exploration students begin to realize that, with rank equated to the number of pivots in a REF, we have in general :

A system of linear equations $Ax = b$ is solvable if and only if $\mathit{rank}(A \mid b) = \mathit{rank}(A)$.

This lets students understand the solvability condition for $Ax = b$ in terms of the rows of $(A \mid b)$ and its REF, namely that there are no inconsistent rows.
Next (before actually solving any linear system at all) we study unique solvability and observe the column condition that

$Ax = b$ is uniquely solvable if and only if $Ax = b$ is solvable <u>and</u> every column in a REF of A contains a pivot.

The actual method of solving linear equations (by backsubstitution performed on the REF of $(A \mid b)$) is the last part of our teaching on linear equations. And in my opinion, it should be, if we want to foster a proof-depth understanding of the subject in our students.

After a study of subspaces, such as the image and kernel of a linear map, the next crucial concepts are linear (in)dependence, span, and basis. The classical linear independence definition that is used and printed in every current Linear Algebra textbook is as follows:

A set of n vectors $\{x_i\} \subset \mathbb{R}^m$ is linearly independent if and only if $\sum\limits_{i=1}^{n} \alpha_i x_i = 0 \in \mathbb{R}^m$ implies that $\alpha_i = 0 \in \mathbb{R}$ for all i.

This definition, however, can only be grasped by someone with a fair amount of training in mathematical logic. It is not intuitive and – in my experience – completely beyond the comprehension of our sophomores by using a logical implication to define the common term.
Instead we can use an exploratory introduction to help define 'linear (in)dependence' concretely, thereby following the modern recommendations of [7]: For a matrix $A \in \mathbb{R}^{m,n}$ we compare the degree of freedom $(= n)$ in the domain of the linear map $x \in \mathbb{R}^n \mapsto Ax \in \mathbb{R}^m$ with the degree of freedom in its range $\mathrm{im}(A) := \{Ax \mid x \in \mathbb{R}^n\} \subset \mathbb{R}^m$. This leads students to think about redundant column vectors (linearly dependent ones) in the matrix A. What columns of A do not contribute independently to its range space? This translates to a question on the REF of A: The columns of A that correspond to a pivot column in any REF of A span the whole range. The others can be reconstructed from the pivot ones, since we can solve for these in terms of the pivot columns, i.e., we can express them as linear combinations of the columns associated with a pivot in the REF. Therefore our first linear independence definition instead is:

A set of n vectors $\{x_i\} \subset \mathbb{R}^m$ is linearly independent if and only if the REF of $A := \begin{pmatrix} | & & | \\ x_1 & \ldots & x_n \\ | & & | \end{pmatrix}$ has n pivots.

This definition is two steps closer to success than the classical logical implication definition of linear (in)dependence in terms of linear equations that is presented in most current textbooks. To see this, let us analyze the classical linear independence condition in its steps:

(1) First we interpret $\sum_{i=1}^{n} \alpha_i x_i = 0 \in \mathbb{R}^m$ as a homogeneous linear system $\begin{pmatrix} | & & | \\ x_1 & \dots & x_n \\ | & & | \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^m$.

(2) Then we rephrase "$\sum_{i=1}^{n} \alpha_i x_i = 0 \in \mathbb{R}^m$ *implies all* $\alpha_i = 0$" as a question on the unique solvability of the linear system in (1).

(3) Finally we row reduce $\begin{pmatrix} | & & | \\ x_1 & \dots & x_n \\ | & & | \end{pmatrix}$ to decide on the unique solvability of the system in (1), and thereby decide linear (in)dependence of the given set of vectors $x_1, ..., x_n \in \mathbb{R}^m$, all according to the classical definition.

Why teach three steps to students, when one, namely (3), suffices? The classical logical implication definition has its uses in theoretical proofs and should also be given, but only after the concrete one. Our concrete matrix based definition should be explored and practiced over and over again in class and with homework. We recommend both repeated hand computations of row echelon forms for specialized unimodular data that allows for integer LR factorizations throughout, <u>and</u> repeated computer checks via MATLAB in order to balance concrete 'number crunching' with conceptual understanding, see [9] and [10].

For example we consider problems such as

(a) Are the vectors $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 6 \\ -3 \end{pmatrix}$, and $\begin{pmatrix} 2 \\ 4 \\ -2 \end{pmatrix}$ linearly independent or not;

(b) Are the vectors $\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 6 \\ -3 \end{pmatrix}$, and $\begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}$ linearly independent or not; and

(c) For which vectors $u \in \mathbb{R}^4$ are the three vectors u, $u + e_2$, and $u + e_3 \in \mathbb{R}^4$ linearly independent, where e_i denotes the i^{th} unit vector in $\mathbb{R}^4$.

To answer these questions by using our concrete definition, students need only check (3), i.e., form the column vector matrix, followed by row reduction and counting pivots.

In case of (a) the associated column vector matrix $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 4 \\ -1 & -3 & -2 \end{pmatrix}$ row reduces to $\begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ with one pivot, signifying linear dependence of the three vectors. Note that we consciously avoid to mention that all three vectors in (a) are multiples of the first vector. Instead we replace this somewhat customary, but intrinsically geometric insight with the straightforward counting of pivots in a REF of the column vector matrix. This avoids student conceptual problems later, see below.

For (b) the associated matrix is $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 4 \\ -1 & -3 & -1 \end{pmatrix}$. It row reduces to $\begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ with two pivots, showing linear dependence of this set of three vectors.

These numerical examples prepare us for the more abstract problem (c). Its given vectors are represented in the 4 by 3 column vector matrix $\begin{pmatrix} | & | & | \\ u & u+e_2 & u+e_3 \\ | & | & | \end{pmatrix}$. Upon a bit of reasoning on the three given vectors $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$, $u + e_2 = \begin{pmatrix} u_1 \\ u_2 + 1 \\ u_3 \\ u_4 \end{pmatrix}$, and $u + e_3 = \begin{pmatrix} u_1 \\ u_2 \\ u_3 + 1 \\ u_4 \end{pmatrix}$, this matrix row reduces to the form $\begin{pmatrix} * & * & * \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ with $* \neq 0$ in case the first entry u_1 or the last entry u_4 in $u \in \mathbb{R}^4$ is nonzero, or to $\begin{pmatrix} 0 & 0 & 0 \\ * & * & * \\ * & * & * \\ 0 & 0 & 0 \end{pmatrix}$ with arbitrary second and third row entries marked by $*$ in case $u_1 = u_4 = 0$. Note that only the case $u_1 \neq 0$ or $u_4 \neq 0$ leads to three pivots. And this happens precisely when $u_1 \neq 0$ or $u_4 \neq 0$, or $u \notin \text{span}\{e_2,\ e_3\}$.

Homework for the reader : Try to solve (c) using the classical linear independence implication only.

We have consciously chosen a strict and concrete “pivot counting” approach to settle linear (in)dependence of vectors because for one it works very well in class, and for another it completely avoids the student confusion with linear (in)dependence, known as the “fog” in [2, p. 30] or as the “brick wall” in the promotional literature for [6]. Specifically there is strong evidence by teachers and by researchers in math education alike that a geometric start, i.e., 2–D or 3–D examples of linear (in)dependence, generally leads to some student confusion. Such a geometric approach violates two of Piaget's principles [8], namely those of 'Concreteness‘ and of 'Generalizability‘. See specifically [3, p. 265, 266] and [4]. The cited references disclose these difficulties to be student difficulties that are often not shared or felt by teachers. These snags are apparently due to several layers of added abstraction and an insufficient correlation between the intuitive geometry of low–D space and $\mathbb{R}^n$ for first time linear algebra students. Clearly, simple truths of $\mathbb{R}^2$ or $\mathbb{R}^3$ do not hold in $\mathbb{R}^n$.[1]

Given these caveats, we advise to explore and practice linear (in)dependence examples in low dimensions via our REF based definition (3) at first and not via the classical definition or geometrically.

Moreover, in a proof and comprehension oriented linear algebra course, the concrete REF based linear independence definition should supersede the classical logical implication definition since our aim is to reach and increase our students' level of sophistication. Its concreteness gives the students an easy and

[1] Compare also with the “artist's viewpoint” of F. Dostoevsky in *The Brothers Karamazov*, book 5, part 3:
... *But you must note this: if God exists and if He really did create the world, then, as we all know, He created it according to the geometry of Euclid and the human mind with the conception of only three dimensions in space.* ...
[spoken by Ivan Fyodorovich]

handy tool. Besides, when we explore the range space of a linear map via pivots of a REF, we unify our teaching and open the field for student understanding. In fact, putting 'linear transformations' at the core of our elementary linear algebra teaching, such as in [9], enables us to teach and comprehend the subject better and more deeply, see [10].

After introducing matrix inverses and their computation using the well known $(A \mid I) \rightarrow (I \mid A^{-1})$ row reduction scheme, the next – and in my opinion possibly last – subjects that need and benefit from an early intuitive WWHWT approach are basis change and matrix representations with respect to different bases. The coordinate vector $x_{\mathcal{U}}$ of a point $x \in \mathbb{R}^n$, when expressed with respect to the basis $\mathcal{U} = \{u_1, ..., u_n\}$ of $\mathbb{R}^n$, is the solution of the linear system $\begin{pmatrix} | & & | \\ u_1 & ... & u_n \\ | & & | \end{pmatrix} x_{\mathcal{U}} =: Ux_{\mathcal{U}} = x$. Basis change from $\mathcal{U}$ coordinates $x_{\mathcal{U}}$ to $\mathcal{V}$ coordinates $x_{\mathcal{V}}$ is achieved by $X_{\mathcal{V}\leftarrow\mathcal{U}} = V^{-1}U$ since $Ux_{\mathcal{U}} = x = Vx_{\mathcal{V}}$. Consequently, we can express a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ that has the standard matrix representation $A_{\mathcal{E}}$ with respect to the standard unit vector basis $\mathcal{E} = \{e_i\}$ in terms of another basis $\mathcal{U}$ by the matrix $A_{\mathcal{U}} = U^{-1}A_{\mathcal{E}}U$. Here the right most matrix factor U of $A_{\mathcal{U}}$ maps $\mathcal{U}$ coordinate vectors to standard basis vectors as seen above, then $A_{\mathcal{E}}$ processes those as the linear transformation specifies, before U^{-1} sends the resulting standard basis images back to their $\mathcal{U}$ coordinate vectors. Thus $A_{\mathcal{U}} = U^{-1}A_{\mathcal{E}}U$ describes the action of the linear map on $\mathbb{R}^n$ entirely in terms of the basis $\mathcal{U}$. This exploration is intuitive and sits well with students. Being self-evident, it needs no formal proof and it establishes the main result elegantly. Besides, it naturally prepares students to accept rigouros DLPTPC type deductions in further expositions, such as with eigenvalues and eigenvectors. These are now introduced in terms of a diagonal matrix representation $A_{\mathcal{U}} = \text{diag}(\lambda_i) = U^{-1}A_{\mathcal{E}}U$ for a given square matrix $A = A_{\mathcal{E}}$ and a specific basis $\mathcal{U}$.

5 DLPTPC type approach in the late stages of an elementary linear algebra course

The reason for looking at eigenvalues is to find revealing matrix representations of a given linear transformation. If $A_{\mathcal{U}}$ is diagonal for a certain basis $\mathcal{U} = \{u_1, ..., u_n\}$ of $\mathbb{R}^n$, then $A_{\mathcal{E}}u_i = \lambda_i u_i$. That is, we are naturally led to study the eigenvalue/eigenvector equation. Again we best proceed via exploration: Do eigenvalues and eigenvectors exist for each square matrix A? Why do they exist? How can they be found?

With determinants, the answer is obvious by noting that $Au = \lambda u$ or $(A - \lambda I)u = 0$ for some $u \neq 0$ is equivalent to $A - \lambda I$ being singular, or to $\det(\lambda I - A) = 0$. The Fundamental Theorem of Algebra, used on the characteristic polynomial of A, then lets us – at least in theory – find the eigenvalues of A.

Proving eigenvalue existence is even easier, when we do not teach or use determinants: We simply take a nonzero vector $y \in \mathbb{R}^n$ and look at the $n+1$ vectors $y, Ay, ..., A^n y \in \mathbb{R}^n$. They are necessarily linearly dependent. Therefore there is a polynomial $p(x)$ with $p(A)y = 0 \in \mathbb{R}^n$. p has degree n at most and its coefficients can be derived from the first linear dependence encountered in the row reduction of $\begin{pmatrix} | & | & & | \\ y & Ay & ... & A^n y \\ | & | & & | \end{pmatrix}$. Now we factor this polynomial p, called the vanishing polynomial for A and y, into first degree factors $p(x) = c \prod_{i=1}^{k} (x - \lambda_i)$ with $k \leq n$ and $\lambda_i \in \mathbb{C}$. And we observe that not every matrix factor $(A - \lambda_i I)$ of $p(A) = c \prod_{i=1}^{k} (A - \lambda_i I)$ can be invertible, unless $0 \neq y \in \ker(p(A)) = \{0\}$ which is a contradiction. Thus at least one of the matrices $A - \lambda_i I$ is singular, leading to an eigenvalue/eigenvector pair. And in fact, each one of the matrix factors $A - \lambda_i I$ of $p(A)$ is singular.

From here on, symmetric and normal matrices, the Schur normal form, the SVD, the Jordan normal form, and all of modern matrix theory can be studied safely in the more formal DLPTPC way, since definitions, concepts, and proofs have become accessible both mentally and emotionally, and they have become part of the students' experience and desire for certainty.

Note that with our early WWHWT and our later DLPTPC approach combined, we have been able to transit in one semester from almost 'no proofs' through 'salient point proofs' to 'formal, rigorous proofs'.

6 Conclusion

The described transition of our teaching methods from WWHWT to DLPTPC type explanations in one semester of linear algebra is of great benefit to our students. It gives a satisfying growth experience to both the students and the instructor. Moreover, if at any of the later stages of the course, the formality of a DLPTPC presentation does not sit well with a student or a group of students, then the teacher can readily revert to the intuitive WWHWT approach of teaching and thereby help stragglers catch up. This is generally not necessary since – in my experience – fellow students will jump in and explain the points of a late formal proof much more 'student–like' than I usually can, and they will readily do so, too.

Even with brilliant classes and students, once I have started using WWHWT style explorations, I never miss the DLPTPC method of presentation (in which I was originally taught long ago in Germany). The reason for this is that math classes are generally somewhat heterogenous. While the best and fastest students would quickly learn to adopt and argue in a DLPTPC way, they are, however, also very satisfied and get excited to be led through mathematical explorations in the WWHWT approach. And in a way, the exploratory WWHWT method nicely evens the playing field: In its explorations it does not matter so much what a student already knows and can do. Instead, the impressions and perceptions of our math explorations bring knowledge to all of our students minds.

Our approach in [9] to teaching a first linear algebra course differs from that of most, if not all other textbooks. Besides introducing proofs gently via exploration as described above, [9] bases and introduces all subjects of such a course on linear algebra's fundamental concept of *linear transformation*, rather than on its elemental tool of row reduction or Gaussian elimination, as most other textbooks currently do. In particular, our approach starts by classifying all linear transformations of $\mathbb{R}^n$ as matrix $\times$ vector multiplication and then uses linear transformations to illustrate all subsequent concepts such as linear equations, subspaces, matrix inverses, linear (in)dependence, basis, eigenvalues, etc. This unifies the whole course and helps students understand, retain, and apply the earlier results throughout the later parts of the course, thereby improving student understanding overall. For more details see [10].

7 Drawbacks

With this measured and slow introduction of formal mathematical rigor, there is little time to work on the fine points of mathematical proofs, such as the structure of direct versus indirect proofs, or the mechanics of induction proofs. Specifically in elementary linear algebra, we often use 'reduction proofs', rather than formal 'induction proofs'. An example is the constructive proof of the Schur Normal Form

in which an arbitrary matrix $A \in \mathbb{C}^{n,n}$ is unitarily reduced to the form $\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$ in the first stage and then the induction principle is invoked on $B \in \mathbb{C}^{n-1,n-1}$.
The finesses of "proof" must simply wait for later math classes.

Acknowledgement

I thank my referees and colleagues for their many astute and helpful comments. Indeed, the vector example in section 4 was inspired by the comments of one referee.

References

[1] P. Blanchard, R. L. Devaney, and G. R. Hall, (1996) *Differential Equations*, PWS (paperback); 2^{nd} ed., Brooks/Cole, 2002.

[2] D. Carlson, (1993), Teaching linear algebra: must the fog always roll in?, *The College Mathematics Journal*, 24, p. 29–40.

[3] J.–L. Dorier and A. Sierpinska, (2001), Research into the teaching and research of linear algebra, in *The Teaching and Learning of Mathematics at University Level, An ICMI Study*, D. Holton (ed.), Kluwer, p. 255-273.

[4] G. Harel, (2000), Three principles of learning and teaching mathematics: Particular reference to linear algebra: Old and new observations. In Jean-Luc Dorier (Ed.), *On the Teaching of Linear Algebra*, Kluwer, p. 177-190.

[5] E. J. Kostelich and D. Armbruster, (1996), *Introductory Differential Equations, from linearity to chaos*, Addison–Wesley, 645 p..

[6] D. Lay, (1993), *Linear Algebra and its Applications*, 2^{nd} ed. 2000, Addison–Wesley, 485+ p..

[7] The Linear Algebra Curriculum Study Group recommendations for a first course in linear algebra, *The College Mathematics Journal*, 24 (1993), 41–46.

[8] J. Piaget and R. Garcia, (1983), *Psychogenesis and the History of Science*, Columbia U Press: New York.

[9] F. Uhlig, (2002), *Transform Linear Algebra*, Upper Saddle River: Prentice–Hall, 504 + xx p.

[10] F. Uhlig, (2002), A new unified, balanced, and conceptual approach to teaching linear algebra, to appear in Linear Algebra and its Applications, vol. (Haifa 2001 Conference issue) , 11p.
Or see `http://www.auburn.edu/~uhligfd/TLA/download/tlateach.pdf`