# SOME THOUGHTS ON A FIRST COURSE IN LINEAR ALGEBRA AT THE COLLEGE LEVEL

Ed Dubinsky

Purdue University and Education Development Center

**Running head:** Thoughts on Linear Algebra

## Abstract

This paper has two purposes: to react to the recommendations of the Linear Algebra Curriculum Study Group (LACSG) and of David Carlson (one of the organizers of the LACSG), and to begin consideration of an alternative approach to helping students learn linear algebra.

The LACSG efforts to bring linear algebra into the overall undergraduate curriculum reform movement are laudable. There are some concerns with their approach, however, and this paper attempts to raise some questions concerning the lack of a body of research establishing a need for curriculum reform in linear algebra, the apparent identification, in the LACSG recommendations, of abstract with useless (and hence, by negation, the equating of concrete with useful) and the assumption that performing calculations with matrices (even intriguing operations) is the same as applying mathematical ideas so as to meet the needs of client departments.

The alternative proposed here — not in the spirit of replacement but in the sense of letting one more flower bloom — is an extension to linear algebra of an overall curriculum reform project. This project involves learning mathematics by programming in a mathematical-oriented programming language, extensive use of cooperative learning, and the development of alternatives to the lecture method.

The discussion of the alternative in this paper is far from complete. It is, rather, an attempt to make a beginning of mapping out a project that will apply to linear algebra an approach that is having some success in precalculus, calculus, discrete mathematics and abstract algrebra.

Work on this project has been partially supported by the National Science Foundation.

## INTRODUCTION

The Linear Algebra Curriculum Study Group (LACSG) has made an important contribution to curriculum reform in mathematics. In general, they have placed linear algebra on the agenda and in particular, they have developed a set of recommendations for a first course in linear algebra [4]. By holding conferences and generating publications (such as the present volume) they have initiated and maintained a level of activity which is necessary if the mathematical community is going to engender revision of the linear algebra curriculum that will be accepted by the community as a whole.

The LACSG curriculum recommendations are an essential first step in curriculum reform, but I believe that important additions and even changes in directions from this first approximation are called for. The main thrust of this article is to point out what I consider to be some omissions in the discussion so far and to introduce some points of view that are different from what appear to be espoused by LACSG and by Carlson.

In the first place, there seems to be a "conventional wisdom" (in which I concur) that the existing curriculum in linear algebra is not successful in terms of what is learned by students who take the first course. There is plenty of anecdotal evidence. David Carlson [3] says that his students become "confused and disoriented" when faced with ideas such as subspaces, spanning, and linear independence, and Harel points out that both high school and university students encounter difficulty with the basic notions of linear algebra [12].

There is not, however, a body of research that provides evidence that would convince a skeptic of the lack of success of linear algebra courses. Unlike calculus and some other topics, we do not have data about failure rates or attrition, analyses of exam questions and results, or documentation of complaints from faculty who teach courses for which linear algebra is a prerequisite[1]. I think that as we tool up for serious reconsideration of the curriculum in linear algebra, some people should be engaged in careful studies that document the inadequacy of the present courses — or show that we are under somewhat of a misapprehension about their adequacy.

The second point I wish to consider in this paper has to do with the LACSG-proposed first course in linear algebra as described in [4]. In my opinion, the main issues which this proposal addresses are: the major difficulties students have with linear algebra and their causes, and the principle that the course should be oriented towards applications and operations with matrices. Although I do not reject the LACSG proposal completely, I would like to raise certain questions about it relating to these ideas and their implementation in the recommended syllabus.

Finally, although I will not propose anything so complete as what we have from

[1] See [5] for a summary of reports that provide such information for calculus.

LACSG, I would like to suggest an alternative program of research and curriculum development that could lead to a course that is quite different from what LACSG has proposed. This is not a bad situation. As with calculus reform, I think there will be general agreement that, if the present linear algebra courses are as unsuccessful as some of us think they are, then real curriculum reform requires that, at least at the early stages, there should be a number of different approaches to innovation in helping students learn linear algebra.

Indeed, there already are a number of beginnings and we will discuss some of them briefly below. They represent a variety of strategies and this is as it should be. Although I think it is important that we look at these initiatives with a critical eye, we should continue for a long time to be open to any seriously considered approach to improvement in learning linear algebra.

So these are the two main considerations of the rest of this paper. I will provide a critique of what the LACSG and David Carlson have presented and then I will talk about an alternative approach. But before beginning the two main topics, I should say something about my background relative to the issues that I am considering.

I am involved with a loosely connected team of individuals throughout the United States and Mexico who are working, accroding to a very specific paradigm, on research and development aimed at a complete revision of all mathematics courses in the first two years of college. The group is called *Research in Undergraduate Mathematics Education Community* or *RUMEC*. We are doing research and have and/or are preparing textbooks and teaching packages for courses in precalculus, calculus, discrete mathematics Laplace Transforms, and abstract algebra. We intend to begin work in linear algebra and differential equations fairly soon. A related group has published a report on cooperative learning ([11]). Our approach makes use of a theoretical perspective on how mathematics can be learned and makes use of students programming computers, cooperative learning, and alternatives to lecturing.

We have not, however, actually begun our work in linear algebra, and therefore the reader should not consider the opinions expressed in this article to be informed by any direct study or experience. What I have to say is based on my experience in working with other, closely related undergraduate courses, and should be taken as speculations that are an essential part of getting ready to do some serious work and an attempt to try out extensions to linear algebra of what I feel has been learned in our previous work.

There is another, rather personal point I would like to make. I will be critical of the LACSG recommendations calling for an emphasis on the coordinate representation of vectors and operations with matrices. This hurts because I love coordinate representations. In the 25 years I spent as a research mathematician in the functional analysis field of nuclear Fréchet spaces, most of my activity was concerned with representations of elements of a vector space as an infinite sequence of coordinates. For me, this point of view was aesthetic, intuitive, rich, and productive. Therefore I begin with much sympathy for the point of view of LACSG, and it is with considerable heaviness of heart that I conclude that, from a pedagogical point of view, their approach is ill-advised.

## THE APPROACHES OF THE LACSG AND OF CARLSON

Let us take these approaches as being "defined" by the two articles [3,4].

It appears to me that there are three major aspects of the LACSG article [4]:

- It is recommended that the course should be recognized as a service course and should relate closely to applications of linear algebra.
- It is recommended that the course should be matrix-oriented.
- A fairly detailed syllabus is proposed.

In [3], Carlson adds three more important points:

- Major topics in elementary linear algebra that give students difficulty are specified.
- Reasons for these difficulties are suggested.
- A set of problems for students is presented to deal with these difficulties.

I will discuss the first two issues raised by Carlson in the next section and the three aspects of the LACSG article as well as Carlson's student exercises in the section after that.

## STUDENT DIFFICULTIES AND SOME REASONS FOR THEM

The Carlson paper mentions the following topics as central to linear algebra and giving students difficulty: subspaces, the spanning set of a subspace, the subspace spanned by a set of vectors, linear dependence and independence, and various aspects of bases and dimension. Examples are also given of student difficulties with row and column spaces of a matrix, rank and nullity of a matrix and their relation to the row and column spaces, and vector spaces of matrices (and of functions).

I would add to this list the geometric interpretation of the action of a linear transformation and I would make the point that applications are not always very easy for students (see Harel [12]). Otherwise, I think that it is a reasonable list and I predict that detailed study would show that these are indeed the main trouble spots for students.

Carlson [3] lists four reasons for the difficulties with subspaces, spans and linear independence/dependence. Here is my interpretation of what he is saying:

1. The course is taught too early and the students are too unsophisticated.
2. The difficulties have to do with concepts and students have little experience with learning ideas as opposed to the less difficult computational algorithms.
3. Students are not experienced with the using — much less the determining — of different algorithms to work with a concept in different settings.
4. Concepts are introduced without substantial connection with students' prior experience.

It is here that I begin to differ with Carlson. It is not that I think that the reasons he offers are wrong, but that they are too general and don't point very clearly to ways in which the difficulties can be overcome. For example, if the students are too unsophisticated and the course should be given later, is it enough just to wait? Are we prepared to assert that another year of college courses will significantly improve the mathematical sophistication of undergraduates? The reports we are receiving from national studies suggest otherwise (see [5] and the references in it.) The effect of another year of coursework is more likely to lead to fewer students taking mathematics and, for those that remain, a stronger conviction that mathematics consists of, in the words of Ed Moise, a repertoire of imitative behavior patterns [15].

The second and fourth reasons given by Carlson tell us that the difficulties listed have to do with conceptual understanding and why it is not an outcome of students taking our courses. But it is far from enough to say that we present these concepts without connecting them to students' prior experience. I think that most of what is done in traditional classrooms has little chance of enhancing student understanding. This is because the standard approach is to put the ideas before the students — by talking about them, or having students read about them, or showing students how to do problems and expecting them to extract general principles by copying what their instructors are doing. I believe that Carlson is correct in his position that such passive activities by students do not contribute much to learning mathematics. The LACSG is correct in calling for more active learning by students. But I do not think that these papers are pointing us in the right direction of how to achieve that.

I am convinced that students develop conceptual understanding as a result of responding to problem situations by making mental constructions of mathematical objects and processes and using them to make sense out of the problem and trying to solve it. I believe there is now enough evidence of this to make it more than a plausible point of view. I also believe that learning concepts in this way can help with Carlson's third point concerning the choice of an appropriate algorithm.

What this implies is that before pedagogical strategies are considered, the particular concepts that give students difficulty in linear algebra need to be analyzed epistemologically. By this I mean that research is needed to determine the specific mental constructions that a student might make in order to understand these concepts. Then, pedagogical strategies need to be developed that can lead to students making these constructions and using them to solve problems.

This is a call for a program of research and development that goes far beyond the suggestion that we get students to work hard in performing matrix manipulations, which is the main explicit pedagogical suggestion in [3]. Thus, my disagreement with Carlson on this point is that his explanation of the difficulties is not rich enough to suggest remedies that, in my opinion, have a good chance of being effective. Later in this article I will try to suggest how the research and development program I am suggesting might work.

## APPLICATIONS AND MATRIX OPERATIONS

The main thrusts of the LACSG recommendations are that a first course in Linear Algebra at the college level should be applicable (and seen as applicable by the students) and should be matrix oriented. It is here that I differ most strongly with their position.

There is an implicit assumption in the LACSG proposals that the goal of developing an applicable course is automatically achieved by emphasizing coordinate representations of vectors and linear transformations in the form of tuples and matrices. For example, taken together with some of the details in their proposal, their statement, "...a change of focus from an abstract, inward-looking course to a more practical matrix-oriented course that meets the needs of not only mathematics students but also the students of the various client disciplines..." suggests an identification of abstract with impractical and matrix-orientation with usefulness.

I think this is a mistake that we often make. Our students demand that the material in a course be less abstract and more concrete, or "down-to-earth". I believe that, often, what students are really asking is for us to give them more computational procedures they can imitate. We, the teachers, don't like the idea of doing that, but we do like the idea of having a lot of applications in our courses. So we give them their recipes and call it applications. I am not suggesting that this is what the LACSG intended to call for, but I am worried that their recommendations and some of the solutions that arise from considering them could, unwittingly, cause some teachers to fall into this trap.

Consider, for instance, the examples given in Carlson's paper. I see no connection, direct or implied, between these problems and applications. Every one of them consists, essentially, of a set of matrix manipulations. It is true that the (stated) intention is that these computational procedures be used to help students understand certain important ideas in linear algebra, but there is no indication of how to make that happen. Indeed, in the absence of any other considerations, and based on our experience with mathematics education, a reasonable prediction is that with such a syllabus, the students will, at most, learn to perform the matrix manipulation procedures and little else. I do not see how these activities are going to help make the course more applications oriented.

Actually, it is not certain that these particular problems will even make the course more concrete. They will do so if, as I have suggested, the students relate to them as computational procedures to be memorized and imitated. This is what "concrete" can come to mean. But if an instructor does succeed in getting students to pay attention to the ideas behind these computations then I wonder if the problems remain so concrete. Let me try to express a certain subtlety having to do with the abstract and the concrete in terms of one of Carlson's examples.

There is a problem in which the students are asked to show that the two matrices $\begin{pmatrix} I & 0 \\ A & I \end{pmatrix}$ and $\begin{pmatrix} I & 0 \\ B & I \end{pmatrix}$ commute where $A, B$ are square matrices of numbers. I claim that this is not a "concrete" problem but is rather abstract for many students. That is, it is true that $A$ and $B$ are themselves concrete and will appear so to most students if considered in isolation. But to see an array of numbers as a single object which can be manipulated just like a number requires a level of abstraction of which many students are incapable. Such a difficulty is not less serious than the the difficulty of seeing an example of a set with two binary operations as satisfying a collection of axioms.

This kind of issue is endemic to the point Carlson is getting at in using these problems. A row or a column of a matrix is, for most students, a list of numbers. It is not easily considered to be an object in itself that can be considered in a set along with other such objects. I suggest that thinking of linear independence in terms of row vectors may not make the concept more concrete and accessible to students.

So these problems do not embody applications nor are they certain to make the course more concrete. I have other objections. For me, there is a huge gap — not mathematical, but pedagogical — between the difficulties students have with linear independence, bases, or subspaces and the matrix operations proposed by Carlson. In my opinion there is no way to bridge this pedagogical chasm other than through abstraction. Proposing all these calculations just sweeps the difficulties under the rug. What is needed are some ideas, not just about how to be concrete, but about how to get students to go from the concrete to the abstract. I see nothing of this in the LACSG proposals or in Carlson's paper.

Finally, a very specific sort of objection. It is only my opinion and we need research on the question, but I have long suspected that students' notions of what is accessible and concrete versus what is obscure and abstract may differ from ours. For example, I wonder which makes students more uncomfortable, the notion of a subspace (in terms of the axioms and standard examples) or notation such as the following which appears in [3, p.34].

$$\sum_{j=1}^{k} b_{jk} col_j(A)$$

I, for one, would not bet that in a popularity contest with students the sigma notation would win out over formal definitions and examples of subspaces.

There are also some aspects of the syllabus to which I take exception. The most serious is the almost total lack of geometry. In the syllabus, the only two relevant phrases are a call for "a strong geometric emphasis" and suggesting that ideas be "motivated using geometric examples". These are admirable positions, but very general and not much help in suggesting what should actually be done. Overall, I would say that the LACSG proposals describe a course that is not likely to instill in students an appreciation of the role of geometry in linear algebra.

Finally, in the spirit of meeting the needs of client departments, I suggest that more in the way of specific applications be included in the Core Syllabus and not just the set of Supplementary Topics. In particular, more should be included regarding differential equations. I would argue that solution spaces of systems of linear differential equations are one of the most important areas of applications — especially in engineering schools and in science schools, which are sources for a significant portion of the population of students who take linear algebra courses. Including this topic would require that at least some attention be paid to vector spaces which are not finite dimensional.

Also, what about less traditional, but still important, applications of vector spaces over fields other than the real numbers, such as $Z_p$? There are important uses of such objects in communications and coding, for example. In general, I am concerned about how well thought out is the relation of the LACSG syllabus to the actual needs of client departments. I believe that, as with calculus, we should not rely on conventional wisdom and anecdotal evidence here. Research is needed about how linear algebra really is used in other disciplines and their courses.

## AN ALTERNATIVE

### Work that has been done

Before discussing the approach I am suggesting, let me say something about work of others in linear algebra education that is not necessarily connected with the LACSG. There has been some but not a great deal.

Jerry Porter [17] describes an interesting pedagogical approach in a linear algebra course for non-math majors. He uses a text that is strong on applications but less thorough on the theoretical underpinnings. To supplement the text, he assigned each student in the class the task of writing a chapter explaining the concepts of subspace, basis and dimension. They were to take as their starting point the solution space of a system of linear equations and they were required to relate their material to dimension, lines, planes and hyperplanes. Porter reports that his experiment was successful in terms of student learning and their attitudes about theoretical material.

Jean-Luc Dorier describes an approach to teaching linear algebra (and other subjects) in which he explicitly teaches what are called "unifier and generalizer concepts". These are concepts built to unify and generalize different methods already operative in various settings. He proposes to use, for a given concept such as the structure of a vector space, a teaching sequence, based on an epistemological analysis, in which he creates an artifical context which motivates the development of the vector space axioms by the students themselves [6].

Guershon Harel has developed and implemented a linear algebra course for high school students. His approach (see [13]) involves a gradual abstraction that develops in three phases. The first phase deals with visual geometric models of vector space notions such as subspace, basis and dimension, in two and three dimensions. In the second phase, coordinate models, first in $\mathcal{R}_1, \mathcal{R}_2$, and $\mathcal{R}_3$ and then in $\mathcal{R}_n$, are used to redefine the concepts considered in Phase 1. The third phase deals with vector spaces whose elements are undefined but their dimensions are one, two and three. In [12] Harel provides data on the effectiveness of his method.

One observation I would make about Harel's approach is that it focuses on the *objects* of linear algebra that is, vectors and their various relationships such as membership in a subspace, geometric and coordinate representations, etc. All of these objects can be visualized geometrically — at least in the lower dimensions. It is quite a different matter to visualize the linear algebra *processes* which transform these objects, that is, linear transformations. According to Piaget, visual perception (which is the main tool used by Harel) is possible for static objects, but not for dynamic processes. To visualize the latter, he argues, it is necessary to perceive a set of static phenomena and to reason about them in making mental constructions of dynamic processes [16]. See [18] for a study of this question.

There is a new textbook by Banchoff and Wermer [2] which supports a curricular approach which has many similarities to Harel's program and also tries to deal with linear transformations.

Finally, I would mention the paper of Hern and Long [14] which is not a discussion of pedagogy so much as a very thorough description of the effect of linear transformations on various geometrical objects in two and three dimensions. These include circles and spheres, rectangles and parallelopipeds, and triangles and tetrahedra. Major concepts in linear algebra such as eigenvectors, invariant subspaces, canonical forms of matrices, and Markov chains are related to these objects. As it appears in the paper, this material would seem very difficult for beginners in linear algebra, but geometric interpretations are an essential tool for anyone who would develop an understanding of linear transformations. I take this as a challenge to curriculum developers — develop a course in which beginning students can deal with some of the ideas in the paper of Hern and Long.

**An interpretation of student difficulties**

I would like to return to some of the difficulties students have with specific topics in linear algebra such as subspaces, bases, linear independence, linear transformations and matrices. My analysis of these difficulties is along somewhat different lines than that of Carlson. I think that most of the trouble comes from three sources.

First, the overall pedagogical approach in most linear algebra courses is, and has been for a long time, that of telling students about mathematics and showing them how it works. The closest we come to students playing an active role is when we have them work on problems. But this does not really work because, as a profession, I believe we succumb to the student demand that we first show them how to do a certain kind of problem and then ask them to solve many instances of this same problem. I have (jointly with others) written several undergraduate textbooks in which there are almost no "worked examples" that are the same as the exercises students are asked to do. For one of these, in Calculus, the publisher has commissioned numerous reviews. The overwhelming majority of reviewers criticise this lack of recipes for imitation.

To summarize, students do not understand these concepts because they never get the chance to construct their own ideas about them. Clearly, this first source of student difficulties is valid for most undergraduate courses in mathematics, whatever the topics.

My second source has to do with students' lack of understanding of background mathematical concepts that are not part of linear algebra, but are essential to learning it. An obvious example is that having a strong function concept is essential for understanding linear transformations. But I think that a much more serious problem (because it is harder to solve) is students' inability to deal with existential and universal quantification. As I think about concepts such as linear independence, bases, and even eigenvalues, I find it hard to imagine how anyone could possibly understand these ideas, except in very superficial ways, without making use of the tool of quantification. This is a very serious problem because of how difficult quantification is for students. In [9] we give a description of the extensive mental growth required to develop a reasonable conception of quantification and in [8] we describe one approach to helping students achieve such growth. Our approach shows indications of being effective, but it takes some time and would have to be in a course other than (and preliminary to) a first course in linear algebra.

The third and final source of difficulty I wish to discuss is a specific version of my first complaint that we do not use pedagogical strategies that give students a chance to construct their own ideas about the important concepts in linear algebra. I am not proposing that students construct their own ideas and leave it at that. The way I think it needs to work is for students to make constructions of mathematical concepts and then to interact with each other and the instructor in the context of the problem situations which the course provides. These interactions will bring out inconsistencies, contradictions and disagreements. The resolution of these conflicts then leads to an understanding on the part of the student that is consistent with the understandings held by the community of mathematicians.

Thus I propose pedagogical strategies that begin with analyses of the specific mental constructions that might be used to understand a certain concept. Then students are presented with problem situations designed to foster their making such constructions. Finally, the interactions among students, instructor and problem involves the students trying to get the solution to the problem and the instructor trying to get the students to revise their constructions so as to make them more effective in solving the problem and more consistent with the instructor's own constructions. Of course we do not rule out the possibility that in some cases it will be the instructor who revises her or his constructions.

This method requires some knowledge of the constructions involved in understanding various topics and teaching strategies that will be effective in getting students to make such constructions. This requires research in how these topics can be learned and I will discuss a theoretical perspective for doing that in the following section.

**A theoretical perspective for research in learning linear algebra**

Both the conduct and evaluation of research in any field takes place, explicitly or implicitly, in the context of a particular research paradigm. It is better to be explicit, and so I will begin by describing the paradigm for the research discussed here. For a more detaied discussion with several examples and references, see [1] by the *RUMEC* group.

The paradigm is neither the mathematical analysis cum laboratory experiment cum confirmation/rejection method of the physical sciences, nor is it the treatment/control cum statistical analysis approach favored in psychology and the social sciences. Needless to say, however, it is not devoid of elements taken from both of these traditions.

There are three parts to this paradigm for research: theoretical analysis, design and implementation of instructional treatments based on this analysis, and observations followed by evaluations of the instruction. The theoretical analysis produces assertions about mental constructions that can be made in order to learn a particular mathematical topic; the instruction tries to create situations which can foster making these constructions; and the evaluation tries to determine if the constructions appear to have been made. The evaluation also tries to determine the extent to which the students actually learned and it is used formatively to reconsider both the theory and the instruction.

In this section I will consider the first part of the paradigm — theoretical or epistemological analysis, with consequent assertions about mental constructions. This presentation must then be followed, in future work, with a more complete analysis of the concepts in linear algebra in terms of this theory. In the next section, I present some first thoughts about the second part, design of instruction. The empirical studies in the third part are, for linear algebra, just coming onto the drawing board.

The theoretical perspective can be summed up in a single word: constructivism. This word means many things to many people and so it is necessary to say something about the particular brand of constructivism used in this approach.

*A particular interpretation of constructivism.*

Piaget is the father of constructivism. He wrote about it at length, in many papers, in many forms and over a long period of time during which his ideas were developing. Many people have written about his ideas, sometimes repeating them, sometimes revising and expanding them, and sometimes misunderstanding them. Of these things, I have tried in my writings to do a great deal of the first, some of the second and, hopefully, not too much of the third. Following is a general statement which I consider to be a succinct expression of my interpretation of Piaget's constructivism. It forms the basis of my theoretical perspective.

> Mathematical knowledge is an individual's tendency to respond, in a social context, to a perceived problem situation by constructing, re-constructing and organizing, in her or his mind, mathematical actions, processes, objects and schemas with which to deal with the situation.

The first thing which must be done with this very general statement is to make it more specific to mathematics and to particular topics in mathematics, e.g., linear algebra. This is done by describing a mechanism for constructing mathematical knowledge and then applying that mechanism to describe specific constructions that can be made in order to understand a particular topic in mathematics. I call such a description for a particular concept a *genetic decomposition* for that concept. Initially, a genetic decomposition is derived from a theoretical analysis. The remaining steps of the research paradigm result in continued evaluation (as an effective description of how a concept may be learned) and revision of genetic decompositions.

*A Mechanism for Constructing Mathematical Knowledge.*

We begin with another idea of Piaget which he stated as a sort of motto for his constructivism.

> To know a thing is to transform it.

It is interesting that this simple phrase can refer to such a wide variety of learning experiences, from a child investigating an object by picking it up and throwing it, to a mathematician studying a geometric object by applying a linear transformation to it.

As we try to construct our understanding of Piaget's motto, we realize that it implies two kinds of components of knowledge: transformations and that which is transformed. I will refer to that which is transformed as *objects*. Transformations are a little more complicated in that it is useful to include consideration of the relationship between an individual and a transformation which he or she is able to make. For this purpose I will consider two kinds of transformation, *action* and *process*.

Before explaining what I mean by these terms, I would like to mention two things that make it difficult to distinguish between them.

First of all, I believe that the nature of transformations is a continuum running from (before) actions to (after) processes. Thus, although I will consider various characteristics as suggesting one or the other, it should be realized that, in observing individuals, what are found, in reality, are mixtures of these characteristics.

The second difficulty has to do with the fact that the property of being an action or a process is not found completely in either the transformation or the individual, but in the relation between the two. We shall analyze this relationship in terms of the extent to which the subject is part of, or is controlled by, the transformation (action) or the extent to which the transformation is part of, or is controlled by, the subject (process).

*Action.*

A transformation is considered to be an action when it is a reaction to stimuli which the subject perceives as external. These stimuli, and the reaction, can be physical or mental. An action can be either a single-step response or a multi-step sequence of responses. In either case, the effect is to transform, in a physical or mental way, one or more objects.

Thus, a physical reflex (which the subject may not control) or a response that the derivative of $\sin x$ is $\cos x$ can be examples of single-step actions. In such cases, the action is a direct response to a stimulus such as the question, "What is the derivative of $\sin x$?" Recalling from memory can be yet another example of a single-step action.

A transformation can be constructed by coordinating several steps. I will consider an action to be a multi-step action if the "next step", although it comes from the individual, must be triggered by the steps which have been performed up to that point, rather than from an overall conscious control by the individual. Consider, for example, the problem situation to determine (if possible) a representation of a given vector $\mathbf{v}$ as a linear combination of two given vectors $\mathbf{e}$, $\mathbf{f}$, in $\mathcal{R}_2$. Suppose an individual knows (from memory or by following a written algorithm) that the first step is to set up a vector equation of the form

$$\mathbf{v} = x\mathbf{e} + y\mathbf{f}$$

where $x, y$ are numbers to be determined. It can happen that it is only after setting up this equation that the individual knows that the next step is to express the given vectors in coordinate form and only then realizes that it is possible to set up a system of two linear equations in two unknowns. Finally, the individual is then able to think about methods for solving such a system and may remember one or have to look it up.

In this case I would say that this transformation (in the sense of converting three given vectors to two scalars), at this time, is an action for this person. Here, the results of the individual steps are at least partially controlling the person (by triggering memory, for example), who is, in this sense, a part of the transformation.

As an individual experiences a number of the same (or very similar) instances of a particular action, it becomes repeatable for her or him. At this point, it is referred to as an *action scheme*. When an individual reflects on an action scheme, it begins to become a process.

*Process.*

When an individual reflects on an action scheme, it can become perceived as a part of the individual and he or she can establish control over it. Thus, an individual might think of expressing a given vector in terms of two given vectors as something like: *write the first vector as a linear combination of the two other vectors with unknown coefficients, use the coordinate form to express this as a system of equations, and then solve the system for the coefficients.* He or she can think of the action without specific vectors or even without specifying the number of coordinates.

We refer to such a construction of a process from an action as an *interiorization*. Once an individual has constructed a process, several things are possible. For example, two or more processes can be coordinated to get a new process. Thus in the above representation example, instead of starting from scratch to construct all steps of the action, an individual might make use of previously constructed processes for setting up a vector equation, converting an equation from vector to coordinate form, and solving a system of simultaneous linear equations. These can be coordinated so as to consruct the representation process directly.

Perhaps the most important point is that when it is a process, an individual can think about a transformation by imagining it, without actually performing it. It is only at this point that a question like: *which vectors in* $\mathcal{R}_2$ *can be represented as a linear combination of* $\mathbf{e}$ *and* $\mathbf{f}$? becomes meaningful. To answer it, the individual must imagine going through the representation process for many vectors and having enough control over performing the process to be able to reflect on when it might work and when it might fail. Gradually in the mind of the individual, the action of checking a particular vector is replaced by thinking about checking all vectors and deriving conditions for success or failure.

An individual can reverse, in her or his mind, a process. Consider, for example, the process of transforming a vector by expressing it in coordinate form and multiplying it by a matrix. A part of understanding matrix inverses is to mentally reverse this process by asking what vector to use in order to get a prespecified answer.

All of these are operations on processes. Reflecting on them leads the individual to construct objects.

*Object.*

When an individual reflects on operations applied to a particular process, becomes aware of the process as a totality, realizes that transformations (whether they be actions or processes) can act on it, and is able to actually construct such transformations, then he or she is thinking of this process as an object. We say that the subject has *encapsulated* the process to construct an object.

One of the best examples of encapsulation arises in the formation of the dual space $V^*$ of a vector space $V$. A linear functional is a process which transforms a vector into a number. Encapsulating this process into an object allows one to construct a new vector and ultimately a new vector space. The dual space is thus completely different from the original vector space and the nature of this difference, which arises through encapsulation, is a powerful tool in straghtening out the confusion that often arises in the finite dimensional case from the fact that there are very simple isomorphisms between the vector space $\mathcal{R}_n$ and its dual space.

Very often, in mathematics, it is important to *de-encapsulate* an object back to the process from which it came, and even to go back and forth. Thus, in thinking about which linear functionals annihilate a subspace, one is thinking about the set of linear functionals as objects. But in trying to derive some information about the situation, it is necessary to de-encapsulate these objects and study them as processes to see what it is that determines whether the result of applying the functional is 0.

*Summary of theoretical perspective.*

The main ingredients of this theoretical perspective are objects, actions, and processes. The construction of objects begins with the child's construction of the concept of *permanent object* in the first year of life and develops into the complex mental objects of advanced mathematics. Actions are transformations of objects that take place in response to stimuli which the individual may perceive as external. An action may be interiorized into a process in which the transformation is consciously controlled by the individual. It can be performed entirely in thought, or reflected upon without actually performing it. It can also be reversed and/or coordinated with other processes. Transforming processes leads to their encapsulation into objects so that higher order actions can be constructed. In thinking about an object, a subject can de-encapsulate it back into a process. Mathematical activity often involves going back and forth between a process and an object conception.

Included in the basic assumptions of this theoretical perspective are the following assertions: actions are constructed by repeated responses to stimuli; processes are constructed either by interiorizing actions or by transforming existing processes; objects are constructed by encapsulating processes; and, in de-encapsulating an object, the only processes an individual can obtain are processes which were encapsulated to construct this object.

All of this is illustrated in Figure 1.

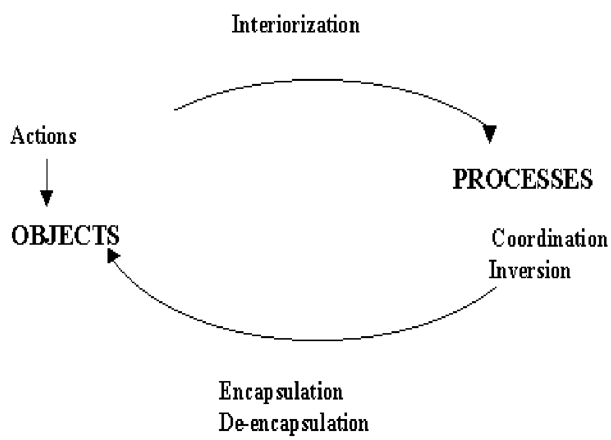


**Figure 1: Mental Constructions**

Finally I will add the point that the action/process/object development is not a one-shot affair. Objects, once constructed, can be transformed to make higher level actions and then processes, and so on. This can continue indefinitely. Moreover, any action, process, or object can be reconstructed, as a result of experiencing new problem situations on a higher plane, interiorizing more sophisticated actions and encapsulating richer processes. The lower level construction is not lost, but remains as a part of the enriched conception. Thus Figure 1 might really look more like a spiral rising out of the paper.

### A beginning of curriculum development in linear algebra — some speculations on instructional strategies

In the previous section, I gave some examples of how topics in linear algebra could be analyzed in terms of our action/process/object theory. This would have to be done for all of the concepts in the course. The next step would be to design instruction that would get students to construct the appropriate actions, processes and objects and use them in solving problems.

Before turning to a discussion of such instruction let me try to remove a possible misapprehension about our paradigm that may have arisen from this discussion of an approach to linear algebra that is at a very early level of development — at least as far as linear algebra is concerned. It may appear as though our ideas of how students can learn about the concepts of linear algebra come entirely from our theory and our own knowledge of the mathematics. This is at most true in the first phase. The second phase is to design and implement instruction based on the results of the first phase. Then in the third phase the observations of students experiencing this instruction are analyzed in terms of the theoretical results and the latter are adjusted to be consistent with the data. The result is a new set of genetic decompositions for concepts in linear algebra and these are then used in a second round of design and implementation of instruction. That is followed by another round of observations and analysis resulting in more adjustments to the genetic decompositions.

All of this is repeated as long as researchers and instructors feel that the course can be improved. The overall result is expected to include a set of genetic decompositions of the concepts in linear algebra that is derived from a synthesis of our theory, our own knowledge of the mathematics, and our analyses of observations of students trying to construct understandings of these concepts. The results are also expected to include a course that is effective for helping students construct their understandings.

In practice it works a little differently. A complete theoretical analysis of all topics in a course before instruction takes place is not practical, and probably not desireable. The paradigm begins by analyzing a small number of the most important concepts and then proceeds to instruction and observation. The revisions in each round then include not only adjustments to the genetic decompositions that were made for the previous round, but also the gradual addition of new topics.

*Pedagogical strategies.*

The overall approach that our project uses has three components: alternatives to lecturing, cooperative learning and computers. Although the first two are extremely important, we have found that in their general formulations, it has not been necessary to change very much in moving from one course to another. These formulations are discussed in [1] and we will not consider them here. Of course the task of working out the specifics for linear algebra will take some effort and this must be done as part of the actual development of the course once that begins.

*Use of computers.*

This universality of our general formulations also holds for the way in which we use computers and we have described what we do in many places. Since our methods are very different from most educational uses of computers, it may be well to review them here. Our approach emphasizes the idea of having students construct implementations of mathematical concepts on computers, essentially by writing programs. We recognize the value of computer demonstrations and the value of giving students computers tools they can use to solve mathematical problems, but we believe that it is much more effective to have students make and use their own tools.

Of course this is not always practical. Sometimes the programs are too difficult for students to write, or it may just take too much time, or involve too many distractions that have to do with programming issues and not mathematics. In such cases, we may just give the tools to the students, or we might ask them to modify a given program, or to explain how it works (emphasizing the mathematical aspects of its operation.)

But our experience has been that there is a rich set of concepts in many mathematical subjects in which the task of writing programs is within the scope of our students' abilities, is free of distracting programming issues, and is closely related to some of the psychological tasks of constructing the processes and objects which the theoretical analyses propose as leading to understanding the mathematics.

We will try to explain how computer constructions can relate to these mental constructions and then finish by giving some examples.

*Computer constructions and mental constructions.*

In this general discussion, we will be concerned with the two most important mental constructions — processes and objects. For a fuller account of the history and nature of our use of computers, see [10].

Understanding a process is difficult in part because, as I indicated earlier, it is hard to visualize a process. You can draw a picture of a plane in $\mathcal{R}_3$ but you can't really draw a picture of the process of a linear transformation which rotates that plane by a certain angle. All you can do with pictures is to draw the beginning and ending figures (and possibly some intermediate ones) with perhaps some arrows to indicate the transformation. The study of the effectiveness of the traditional linear algebra course that I discussed at the very beginning of this article would show, I predict, that such "visualization" is not very helpful in getting students to construct an understanding of linear transformation.

With today's technology, there are intriguing possibilities of visualization through videos and computer animation. I believe that these should be part of the curriculum which helps students construct mental processes that will help them learn linear algebra.

But visualization is not the only powerful tool for thinking about mathematics. There are also analytic strategies. Consider linear functionals, for example, and the mental process associated with a particular linear functional $\mathbf{v}$ in which a vector in $\mathcal{R}_3$ is transformed into a scalar, say by adding its first two components and subtracting twice its third component. We suggest that writing the following computer program (in the language ISETL [7]) will help students construct this process.

```
v := func(x);       $ x is a vector in R3, represented as a tuple.
        return x(1) + x(2) - 2*x(3);
     end;
```

It turns out that very many processes in mathematics can be implemented on the computer in such a simple way.

Constructing objects is a very serious problem for students and many authors have expressed pessimism about how many students can go very far in doing so. For example, once a student has constructed a process for the above functional **v**, we would like her or him to think about it (and other functionals) as objects on which certain operations can be performed. There is very little in our traditional pedagogical repertoire that can help with this. We feel that our approach offers a unique contribution in this area.

The key to coming to think about a process as an object is to have problem situations in which actions need to be applied to the processes. For example, we could set the problem of writing a computer program that will take two linear functionals and return the linear functional which is their (pointwise) sum. Here is a solution.

```
fa := func(v,w);          $ v, w are linear functionals
         return func(x);
                   return v(x) + w(x);
                end;
      end;
```

Once this program has been written and assigned to the variable `fa` then if `v`, `w` are any linear functionals and `x` is a vector, the expression

```
(v .fa w)(x);
```

will be interpreted by the computer and the appropriate scalar will be returned.

We have found that writing and using such programs leads students to begin to encapulate processes into objects.

Before leaving this topic we want to raise a practical issue that relates to our theory. In discussing the problem of getting students to think about functions as objects in the context of real valued functions of real variables, for example, some people have felt that the problem is solved by using graphs. Surely the graph of a function is an object and students will treat it as such. Similarly, the linear functional **v** described above may be considered as an object — the triple $[1, 1, -2]$. Again there is little difficulty in getting students to think about such triples as objects.

I think that both examples represent important pedagogical errors and a misinterpretation of our theory. It is not just thinking about objects, but the relationship between an object and the process which is encapsulated to construct the object. This is because it is so often necessary to go back and forth between a process and object interpretation of the same concept. According to our theory, this can only happen if the object comes from an appropriate process. This explains why so many students can work with graphs and perform various operations with them, but cannot interpret them or relate them to the function's process. Similarly, students who think about our **v** as the object $[1, 1, -2]$ may have difficulty in understanding its relationship to the process involved in this function.

*Some examples.*

I would like to consider three kinds of examples. First, I will discuss finite dimensional vector spaces over finite fields at a fairly detailed level. Second, I would like to set a challenge for synthesizing the ideas behind our approach and the material in the paper of Hern and Long [14]. Finally, I would describe somewhat vaguely an approach to helping students understand vector spaces of functions.

We already mentioned the importance of vector spaces over fields such as $Z_n$. Let us consider the vector space $V$ of all triples of elements of $Z_3$. This space has 27 elements in it and the fastest desk computers (e.g., MAC II and MS-DOS 486 machines) can easily handle binary and ternary operations on a set of this size, and can even go to $Z_4$ or $Z_5$. We will assume in our discussion that the programming language ISETL has been equipped with a user-friendly front end for entering and dsiplaying vectors and matrices. The very basic acivities are to construct the vector space itself with computer operations such as the followng.

```
Z3 := {0,1,2};
V := {[a,b,c] :  a,b,c in Z3};

va := func(x,y);
        return [(x(i) + y(i)) mod 3 :  i in [1,2,3]];
      end;

sm := func(s,x);
        return [(s*x(i)) mod 3 :  i in [1,2,3]];
      end;
```

It is now quite easy for students to write programs that check the vector space axioms for these two sets and two operations. For example, here is code that checks that scalar multiplication distributes over addition of vectors.

```
forall s in Z3, x,y in V | s .sm (x .va y) = (s .sm x) .va (s .sm y);
```

(This is one of the longest calculations and it takes about 4 minutes on a Power Book 170. )

A more important example is the following code which checks if a given vector can be written as a linear combination of two given vectors.

```
exists s,t in Z3 | w = (s .sm u) .va (t .sm v);
```

One can easily imagine the investigations of linear independence, spanning sets, and bases that could take place using such tools. Our point is that because the students will have constructed these tools, their use will relate to ideas in the students' minds and this will immensely enrich the experience of investigation.

Don Muench at St John Fisher College has been trying out some of these suggestions in a linear algebra course and he reports (private communication) that the preliminary indications of results are encouraging.

The second kind of example is really more of a challenge to synthesize an analytic approach and a visual approach. A language such as ISETL is good for the kinds of constructions we have just described and can also be used to produce graphs in two dimensions. One can envision giving students a $2 \times 2$ matrix and having them write code that expresses its action on a vector and also having them use graphics tools to produce pictures of the effect of this linear transformation on a square and on a circle. This can lead to investigation and construction of understanding of the difference between a matrix of rank 1 and a matrix of rank 2 in terms of transformations of a square, invariant subspaces, eigenvectors and eigenvalues. These investigations can also be applied to transition matrices of Markov chains. (See [14] for a discussion of the mathematics that would be very helpful in developing such instruction.)

Any development of instruction along these lines would have to be preceeded, of course, by a theoretical analysis that would guide the development of the instruction and would be refined as the results of that instruction were analyzed.

These methods can be extended to three dimensions once the software is enhanced to permit three dimensional graphics.

Finally, I would like to propose the use of the pedagogical methods described here to discuss the linear algebra aspects of the solution space of a linear differential equation of order $n$. Such an equation can always be transformed into a system of $n$ first order linear equations. It then has the form,

$$y' = Ay$$

where $y$ is a vector of $n$ functions of one variable, $y'$ is the vector of the derivatives of those functions, and $A$ is an $n \times n$ matrix of functions of one variable. Assume that the functions that make up $A$ are infinitely differentiable on $\mathcal{R}$.

According to the theory of differential equations, for each $n$-dimensional vector $\mathbf{c}$ of numbers, there is a unique solution $y$ of this equation with the initial value $y(0) = \mathbf{c}$. Moreover, the set of solutions of this equation is an $n$-dimensional vector space of infinitely differentiable functions on $\mathcal{R}$ and a basis for this vector space is obtained by taking those solutions whose initial values form a basis for $\mathcal{R}_n$.

It is not hard to write a program that will apply Euler's method to approximate the solution of our differential equation for a given initial value. We have our students do it for one dimension in our second semester calculus class. Extending this program to $n$ dimensions requires little more than replacing addition and multiplication by vector addition and scalar multiplication. This seems an appropriate exercise for a linear algebra course.

Let's consider the case of $n = 3$. One can use the graphics facility to plot one function on a chosen interval or even three functions on the same graph. The kind of exercise I am thinking of would give students three sets of initial values. Their task would be to first apply their Euler's method program to obtain the three solutions and convert them to functions of one variable that solve the original third order equation. Then, using plots of these functions on a given interval, to look for two numbers that will express one of the functions as a linear combination of the others. One could advise the students to use a very large mesh size in their Euler approximations at first so that examples would be generated quickly and then when they were close to a solution, to use smaller mesh sizes to get better approximations.

An interesting side issue concerns the fact that what happens at 0, or on a small interval determines the value of a solution function at every point of $\mathcal{R}$.

It seems reasonable to expect that many students would discover the very beautiful theorem to the effect that expressing one initial condition as a linear combination of the other two (which can be done quickly and mechanically) gives the right coefficients for expressing the entire solution function as a linear combination of the other two solution functions, valid on the entire domain. Students will actually see the effect of this linear combination on their graphs. Even if they do not discover the theorem, their search for it will enhance their understanding of it when one of their colleagues explains it to them, or they read about it or even if the instructor explains it.

I believe that a great deal of mathematics can be learned by students using these methods. We have only to develop a little more software, make our theoretical analyses, design the instruction and then go ahead and do it!

## REFERENCES


1. Mark Asiala, Anne Brown, David J. DeVries, Ed Dubinsky, David Mathews, and Karen Thomas, A Paradigm for Research and Development in Undergraduate Mathematics Education, Research in Collegiate Mathematics Education II, in press.

2. Thomas Banchof and John Wermer, *Linear algebra through geometry* 2nd Edi-

tion, New York: Springer-Verlag (1992).

3. David Carlson, Teaching linear algebra: must the fog always roll in? *College Mathematics Journal* **24** 1, (1993) 29-40.

4. David Carlson, Charles Johnson, David Lay, and A Duane Porter, The Linear Algebra Curriculum Study Group recommendations for the first course in linear algebra, *College Mathematics Journal* **24** 1, (1993) 41-46.

5. Committee on the Mathematical Sciences in the Year 2000, *Moving beyond myths: Revitalizing undergraduate mathematics*, Washington, D.C.: National Research Council (1991).

6. Jean-Luc Dorier, Sur l'enseignement des concepts élémentaire d'algèbre linèaire à úniversité, *Reserches end Didactique des Mathématiques*, **11** 2, (1991) 325-364.

7. Jennie Dautermann, *ISETL: A language for learning mathematics*, St. Paul: West Educational Publishing, (1992).

8. Ed Dubinsky, On learning quantification, Journal of Computers in Mathematics and Science Teaching, in press.

9. Ed Dubinsky, Flora Elterman & Cathy Gong, The student's construction of quantification, *For the Learning of Mathematics*, 1989.

10. Ed Dubinsky, ISETL: A Programming Language for Learning Mathematics, *Comm. in Pure and Appl. Math.*, 48, 1995, pp. 1-25.

11. Nancy Hagelgans, Barbara Reynolds, Keith Schwingendorf, Ed Dubinsky, Draga Vidakovic, Mazen Shahin, and Joe Wimbish, *A Practical Guide to Cooperative Learning in Collegiate Matheatics* ,MAA Notes 37, MAA:Washington, DC 1995

12. Guershon Harel, Learning and teaching linear algebra: Difficulties and an alternative approach to visualizing concepts and processes, *Focus on Learning Problems in Mathematics*, **11** 2, (1989) 139-148.

13. Guershon Harel, Using geometric models and vector arithmetic to teach high-school students basic notions in linear algebra, *International Journal for Education in Science and Technology*, **21** 3, (1990) 387-392.

14. Thomas Hern and Cliff Long, Viewing some concepts and application in linear algebra, in Walter Zimmerman and Steve Cunningham (ed.), *Visualization in teaching and learning mathematics*, MAA Notes **19** (1991) 173-190.

15. Ed Moise, Mathematics, computation, and psychic intelligence, in V.P. Hansen & M.J. Zweng (Eds.), *Computers in mathematics education*, 1984 Yearbook of the National Council of Teachers of Mathematics, pp. 35-42. Reston, VA : NCTM (1984).

16. Jean Piaget, *Mental imagery in the child: a study of imaginal representation.* P. A. Clinton, translator. London: Routledge & Kegan Paul and New York: Basic Books, **19** (1966).

17. Gerald J. Porter, Writing about linear algebra: Report on an experiment, *UME Trends*, **3** 3, (1991), 1-3.

18. Rina Zazkis, Ed Dubinsky, and Jennie Dautermann, *Using visual and analytic strategies: A study of students' understanding of permutation and symmetry groups*, Journal of Research in Mathematics Education, in press.

I am just going to start linear algebra for my undergraduate course. And which is the best book available for linear algebra in terms of rigor and a book similar to calculus by Tom M Apostol (I like this book because of it's rigorous and clear ideas, the qualities which I want). Note: I have browsed through other similar topics here and didn't found anything helpful. linear-algebra reference-request.Â  Lax: This book is notable for its speed (duals and quotients are in the first 20 pages). The great part about Lax is he covers a bunch of things other linear algebra books don't, like matrix calculus. (For instance, if you have a differentiable function $A(t)$ taking as its values invertible matrices, what is the derivative of $A^{-1}(t)$?) This is really a great book, but I would never recommend it as a first book.