ZDM Mathematics Education (2014) 46:389–406
DOI 10.1007/s11858-014-0579-x

ORIGINAL ARTICLE

Student reasoning about the invertible matrix theorem in linear algebra

Megan Wawro

Accepted: 3 April 2014 / Published online: 20 April 2014
© FIZ Karlsruhe 2014

Abstract I report on how a linear algebra classroom community reasoned about the invertible matrix theorem (IMT) over time. The IMT is a core theorem that connects many fundamental concepts through the notion of equivalency. As the semester progressed, the class developed the IMT in an emergent fashion. As such, the various equivalences took form and developed meaning as students came to reason about the ways in which key ideas involved were connected. Microgenetic and ontogenetic analyses (Saxe in J Learn Sci 11(2–3):275–300, 2002) framed the structure of the investigation. The results focus on shifts in the mathematical content of argumentation over time and the centrality of span and linear independence in classroom argumentation.

Keywords Linear algebra · Adjacency matrices · Collective activity

1 Introduction

Consider a discussion from a linear algebra class that had been investigating properties of linear transformations and their associated standard matrices:

Instructor: … If the column vectors of A are linearly independent,[1] then you guys are saying it [the matrix] is invertible[2] because of that. Are you guys able to explain why that should give invertibility? I think you said something, or go ahead, Josiah.

Josiah: When they're linearly independent, there's only one path you can take to get to it, so in order to get back, there can only be one answer to get back. Whereas if they're dependent on each other, then depending on how you got there, would determine how you get back … so you don't have the right information again.

Instructor: Jesse…Yeah?

Jesse: Also, if they're dependent, in the RREF, you'll have a zero row, so it will be like you're losing information when you're trying to go back…If they are dependent, then their RREF will have a zero row.

Instructor: Okay, columns dependent [writes] and so you're saying that the RREF of A has a zero or row of zeros.

Jesse: Right, so then if you try and, if you invert that, you can't, because it's like you're losing that information from that row.

Josiah indicated that if the columns of a matrix were linearly dependent, then one would not have the proper information to "get back," whereas Jesse's statements indicated that the row-reduced echelon form (RREF) of a matrix would indicate whether you could "go back." Both Josiah and Jesse associated invertibility with the notion of "going back," but they did so in different ways. How did these interpretations come to be? How did these notions come to function as if the community shared a common

[1] A set $S=\{\vec{v}_1,\ldots\vec{v}_n\}$ is called *linearly independent* if the only solution to the equation $a_1\vec{v}_1,\ldots a_n\vec{v}_n=\vec{0}$ is $a_i=0$ for scalars $a_i, 1\le i\le n$.

[2] An $n \times n$ matrix A is invertible if there is an $n \times n$ matrix C such that $AC = CA = I$.

M. Wawro (✉)
Department of Mathematics, Virginia Tech, Blacksburg, VA, USA
e-mail: mwawro@vt.edu

Fig. 1 The invertible matrix theorem, as developed in class

The Invertible Matrix Theorem

Let A be an $n \times n$ matrix. The following are equivalent:

a. The columns of A span $\mathbf{R}^n$.
b. The matrix A has n pivots.
c. For every $\mathbf{b}$ in $\mathbf{R}^n$, there is a solution $\mathbf{x}$ to $A\mathbf{x}=\mathbf{b}$.
d. Every $\mathbf{b}$ in $\mathbf{R}^n$ can be written as a linear combination of the columns of A.
e. A is row equivalent to the $n \times n$ identity matrix.
f. The columns of A form a linearly independent set.
g. The only solution to $A\mathbf{x}=\mathbf{0}$ is trivial solution.
h. A is invertible.
i. There exists an n x n matrix C such that $CA = I$.
j. There exists an n x n matrix D such that $AD = I$.
k. The transformation $\mathbf{x} \rightarrow A\mathbf{x}$ is one-to-one.
l. The transformation $\mathbf{x} \rightarrow A\mathbf{x}$ maps $\mathbf{R}^n$ onto $\mathbf{R}^n$.
m. Col $A = \mathbf{R}^n$.
n. Nul $A = \{\mathbf{0}\}$.
o. Det $A \neq 0$.
p. The number 0 is not an eigenvalue of A.

understanding of them? In a broader sense, what other ways of reasoning about the concepts related to the invertible matrix theorem (see Fig. 1) developed for this community?

The following research question brings together and extends the inquiries that this example brings to light: How did the collective classroom community reason about the invertible matrix theorem (IMT) over time? For the community analyzed in this study, as the semester progressed, what became known as the IMT (Fig. 1) was 16 concept statements developed individually and related to one another through the notion of equivalence. Analysis is accomplished through the use of adjacency matrices (discussed further in Sect. 5) to investigate the structure of student reasoning both in isolation and over time.

2 Relevance of the study

Reasoning—in the sense of making connections across ideas, representations, and contexts, as well as in terms of argumentation and justification—is a valuable skill and part of the practice of mathematics. The IMT is a powerful theorem in introductory linear algebra because it provides insights into how key ideas in linear algebra relate to one another. Research has shown, however, that students struggle to understand the concepts involved in the IMT, such as linear independence, span, and linear transformation (e.g., Dreyfus et al. 1998; Harel 1989). Thus, studying the development of the IMT, as well as how students understand and reason about the main ideas of the IMT, contributes to what is known about how students learn linear algebra in particular, and more generally how mathematical connections grow over time.

Furthermore, this study is significant because investigating the ways in which a particular classroom community reasoned about the IMT is compatible with investigating what classroom mathematics practices (Cobb and Yackel 1996) relevant to the IMT developed for that community. The practical constraints placed on a teacher in any given classroom are such that she must operate as if the ways of reasoning about the mathematics at the collective level are shared by all individual members of that community. Thus, reporting on mathematical activity developed for a particular linear algebra class at the collective level may allow other linear algebra instructors to become aware of normative ways of reasoning that may develop in their own classroom.

3 Literature review

I include a review of relevant research on student thinking in linear algebra, as well as various analyses of mathematical development at the collective level. I conclude with detail regarding the theoretical framing of genetic analysis.

3.1 Student thinking in linear algebra

According to Hillel (2000), linear algebra is often the first mathematics course that students see that is a mathematical theory, systematically built and reliant upon definitions, explicit assumptions, justifications, and formal proofs. Indeed, Harel (1989) asserts the importance of linear algebra at the college level because it can be studied as "a mathematical abstraction which rests upon the pivotal ideas of the postulational approach and proof" (p. 139). In many universities, however, a first course in linear algebra occurs prior to an introduction to proof course. Research has shown that proof-related difficulties, such as struggling with quantifier usage, tending to reason inductively, or using various proof techniques inappropriately are not unique to linear algebra (e.g., Harel and Brown 2008; Selden and Selden 1987; Weber 2001). As such, Hillel (2000) focused on what he proposed were student difficulties *specific* to linear algebra: the existence of several modes of description, the problem of representations, and the applicability of the general theory (p. 192).

Hillel stated three different modes of representation exist in linear algebra: abstract, algebraic, and geometric. He detailed difficulties students have within the geometric mode (such as confusion caused by describing vectors as both arrows and points), as well as difficulties students have moving between modes. One particularly difficult aspect for students is to move between abstract and the algebraic when the underlying vector space is $\mathbb{R}^n$ for both. This would occur, for example, in change of basis problems. To think of a string of numbers as a representation of a vector (according to a particular basis) rather than the vector itself is difficult for students to grasp or even see the need for grasping.

In a study regarding students' interactions with transformational geometry in linear algebra, Portnoy et al. (2006) investigated if participants viewed transformations as processes or as objects. The authors found that participants displayed an operational view of transformations, as "processes that map geometric objects onto other geometric objects" (p. 201). The authors conjectured this view—thinking of them as a process that carries out an action on other objects—might have contributed to the difficulty students had writing proofs that relied on thinking about transformations as objects themselves. In another study regarding student understanding of linear transformations, Dreyfus et al. (1998) found that students seemed to equate the term "transformation" with the vector "T(**v**)," as if rather than describing a relation between **v** and T(**v**), "transformation" was an object T(**v**) that depended on **v**.

In addition to this and other research that investigates student difficulties in linear algebra, there also exists descriptions of various creative and productive student work done in linear algebra (Possani et al. 2010; Wawro et al. 2011). For instance, Larson et al. (2008) reported on ways that students conceptualized mathematical objects (such as vectors and matrices) as the class developed algebraic methods to solve for eigenvectors and eigenvalues of a given matrix. The authors found the class productively utilized relationships between geometric interpretations of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, linear independence, and determinants to derive conclusions about eigenvectors and eigenvalues. This is consistent with Sierpinska (2000), who found that geometric, arithmetic, and structural reasoning and the ability to move between them are fundamentally important in learning and understanding the core ideas of linear algebra.

These studies highlight the importance of making connections in linear algebra—between and within modes of representation for a particular concept, across operations with and on a concept, and among multiple concepts—both for content understanding and for facility with proof. Investigating how a classroom investigated the IMT, a theorem that connects concepts through equivalence, adds to what is known in this research area.

3.2 Analysis at the collective level

As compared to a purely individual lens for investigating mathematical development, some mathematics education research investigates the emergence, development, and spread of ideas in a classroom community over time (e.g., Elbers 2003; Sfard 2007; Stephan and Rasmussen 2002). For example, Elbers (2003) focused on how reasoning at the collective level may influence individual students' learning, as well as how individual students contribute to whole class discussion. Defining learning as the changing participation of students in the classroom discourse, he analyzed learning by comparing individual work to whole class discussion to "get a view of the progress of mathematical understanding both in the classroom as a whole and in individual children's minds" (p. 82).

In a compatible view of learning and the inseparable interaction between individual and community, the work of Cobb and his colleagues (Cobb and Yackel 1996; Stephan et al. 2003) takes into consideration the social and situated nature of mathematical activity. Their work analyzes students' mathematical reasoning as acts of participation in the mathematical practice established by the classroom community. At the collective level, they analyze mathematical progress by documenting classroom mathematical practices, which are the "taken-as-shared ways of reasoning, arguing, and symbolizing established while discussing particular mathematical ideas" (Cobb 1999, p. 9). Classroom mathematical practices, then, do not illuminate any information about a particular individual student's mathematical development, nor should it be assumed that the two must overlap. Rather, the classroom community is the unit of analysis, documenting the development of practices and the establishment of meanings at the collective level. One example of such analysis is the work of Stephan and

Rasmussen (2002), which analyzes the emergence of normative ways of reasoning of a university differential equations course. Their methodology (Rasmussen and Stephan 2008) relies on various criteria regarding how mathematical content changes within collective discourse, such as claims that once needed justified within the classroom begin to function as if they are mathematical truths.

The focus of the current study is how the collective classroom community came to reason about the IMT over time. These results contribute directly to this emerging research agenda regarding what is known about mathematical development at the collective level. In addition to the aforementioned work by Cobb and Yackel (1996) and Rasmussen and Stephan (2008), my investigation into the emergence, development, and spread of ideas in this classroom community over time is framed in terms of genetic analysis.

3.3 Theoretical framing: genetic analysis

Researchers from both educational and anthropological fields have investigated the progression of human development through *genetic analysis*. Ercikan and Roth (2006) state that, when dealing with a genetic explanation, "a certain fact is not derived from antecedent conditions and laws (deduction) or observations and antecedents (induction) but rather is shown to be the endpoint of a longer development, the individual stages (phases) of which can be followed" (p. 19). For instance, Saxe and Esmonde (2005) studied how the function of a particular word shifted for a central New Guinea tribe over 20 years. They asked, "How do new collective systems of representation and associated mathematical ideas arise in the social history of a social group?" (p. 172). Within the classroom microculture, Saxe and colleagues investigated how to conduct analyses that would allow them to describe individuals' idea develop in the classroom over time, given that the classroom is also changing over time. They suggested analysis through three different strands—microgenesis, ontogenesis, and sociogenesis. They define microgenesis as the short-term process by which individuals construct meaningful representations in activity; ontogenesis as the shifts in patterns of thinking over development; and sociogenesis as the reproduction and alteration of representational forms that enable communication among participants (Saxe et al. 2009, p. 208).

Tiberghien and Malkoun (2009) investigated methodologies to both distinguish between individual and collective analyses of the classroom as well as relate fine-grain and broader analysis over time. The methods involve three scale sizes: microscopic (key words or utterances), mesoscopic (themes over a few lessons), and macroscopic (longer sequences such as a semester). The authors offer possible coordination between these distinctions. For instance, coordination between the microscopic and macroscopic level from a collective perspective could analyze the number of key words or utterances in relation to the duration of a theme or sequence (density), or the distribution of utterances most reused during a theme or sequence (continuity).

The primary means of genetic analysis within this study were inspired by Saxe (2002), Saxe et al. (2009) and Tiberghien and Malkoun (2009). In Saxe et al. (2009), microgenetic construction occurred when a student turned forms such as the number line into a particular meaning and used it to accomplish a goal in activity. In this study, the constructed representational forms witness their analog in students' explanations regarding the IMT, which function to justify connections within the IMT. In other words, microgenetic analysis is accomplished by considering the content and structure of particular instances of reasoning about the IMT during whole class discussion. Furthermore, the compilation of these microgenetic analyses serves as the data for ontogenetic analysis. Considering how these shift over time, conclusions are made regarding reasoning about the IMT at the collective level. For instance, which concepts from the IMT were used most often overall as students reasoned about novel problems? Were some concepts from the IMT used consistently over the semester, whereas others dropped off? This analysis was inspired by Tiberghien and Malkoun (2009) who consider density and continuity of ideas as analytical frames in genetic analysis.

4 Methods

In this section I detail the setting, participants, and data sources for the study. The results section then begins with an explanation of how adjacency matrices were utilized as a methodological tool to analyze the data set.

Data for this study came from an undergraduate, inquiry-oriented introductory linear algebra course. Students had generally completed three semesters of calculus (at least two were required). Approximately half had also completed a discrete mathematics course, and 75 % were in their second or third year of university. Their major courses of study included Computer Engineering, Computer Science, Mathematics, Statistics, and other science or business fields.

The course was designed based on the instructional design theory of Realistic Mathematics Education (RME) (Freudenthal 1991), which begins with the tenet that mathematics is a human activity. As such, the course was designed to build on student reasoning as the starting point from which more complex and formal reasoning developed. The class engaged in various RME-inspired instructional sequences focused on developing a deep understanding of key concepts such as span and linear independence (Wawro et al. 2012), linear transformations, and change of basis. For instance, the class first reasoned about span by considering "all the places you could get to" with a linear combination of two modes of transportation (i.e., two vectors in $\mathbb{R}^2$). They investigated journeys that began and ended at home and developed the formal definitions of linear independence and dependence out of this situational experience. The class also used a textbook (Lay 2003) as a supplemental resource.

The class met twice a week for 15 weeks, and each class session was video recorded. I was the teacher-researcher during data collection, and I met with the remaining four members of the research team three times a week to debrief after class, discuss impressions of student work and mathematical development, and plan subsequent classes to be mindful of both the course learning goals and data collection opportunities for the research study. Data sources for this study, chosen from this set of video recordings, included video of whole class discussion relevant to the development of and reasoning about the IMT. Portions of classroom discourse were considered relevant if they either implicitly or explicitly involved the class members actively engaging in developing ways of reasoning about two or more concepts from the IMT in conjunction with each other. The video for each relevant portion was transcribed completely. From this transcript, arguments were identified, where argument is defined as "an act of communication intended to lend support to a claim" (Aberdein 2009, p. 1). For each argument, I identified the claim and supporting justification given. Each argument was numbered and compiled sequentially into an argumentation log according to the class day on which it occurred. For example, transcript lines 1–19 in Sect. 5.2 were coded as the sixth argument on Day 20. The creation of argumentation logs followed the methodology developed by Rasmussen and Stephan (2008) for the documentation of collective activity. For the purpose of investigator triangulation (Denzin 1978), a research team member independently created argumentation logs for portions of the first 3 days of data. We discussed the argumentation logs by comparing and defending each coding, and a high level of reliability was reached. I then completed analysis independently, reporting my finding back to the research team.

5 Results

Through grounded analysis, these arguments were broken into separate idea clauses and then coded as either IMT concept statements or interpretations of those statements. This resulted in 100 distinct codes arranged into 15 categories (Fig. 2). The main code for each category (except for "miscellaneous") is a concept statement from the IMT or its negation, and the remaining codes are arranged as subcodes under the appropriate main codes. This information was then organized for adjacency matrix analysis, in which codes are vertices in directed vertex-edge graphs (digraphs) and are connected, when appropriate, with arrows in such a way to match the implication offered by the speaker(s).

For a given direct graph, an adjacency matrix is an $n \times n$ matrix with one row and one column for each of the n vertices in the digraph, and each entry $a_{ij} = k$ in the matrix indicates k edges from the ith vertex to the jth vertex. For example, we say vertex u is *adjacent to* v (and v is *adjacent from* u) if there exists an edge from u to v, and an entry of "1" in row u and column v would correspond to one "$u \rightarrow v$" statement such as "if u then v," or "another way to say u is v." Thus, the 100 codes listed in Fig. 2 serve as the rows and columns for adjacency matrix analysis conducted on the given data set. For example, code F4 references vectors that "lie along the same line," an interpretation of statement F. A statement such as, "because the vectors lie along the same line (code F4), I can't go in all the directions with the vectors (code H3), so the vectors do not span[3] $\mathbb{R}^n$ (code H)," would be shorthanded as "F4 → H3 → H," and a "1" would be placed in the F4 row/H3 column, and in the H3 row/H column of the associated adjacency matrix.

Whole class discussion relevant to reasoning about the IMT occurred on ten of the 31 days of the semester (see Fig. 3). Within these days, I identified 109 arguments that were relevant to the class's reasoning about the IMT. The clauses of separate ideas within the 109 arguments informed the creation of the 100 codes (Fig. 2) used in adjacency matrix analysis, 83 of which were used at least once during whole class discussion. Analyzing how the class structured the arguments revealed 452 connections between codes. For instance, "because the vectors lie along the same line (F4), I can't go in all the directions with the vectors (H3), so the vectors do not span $\mathbb{R}^n$ (H)," coded as "F4 → H3 → H," contributes two connections between ideas because it contributes two adjacencies: (F4, H3) and (H3, H). In digraph language, the codes are vertices and connections between them are directed edges for a digraph T. Thus, T has 83 vertices and 452 edges.

The associated adjacency matrix for digraph T is denoted $A(\mathrm{T})_{\mathrm{tot}}$. The "tot" subscript stands for "total," meaning that all arguments for any $m \times n$ matrices (or n

[3] A set of vectors S in $\mathbb{R}^n$ is said to *span* $\mathbb{R}^n$ if any vector in $\mathbb{R}^n$ can be written as a linear combination of the vectors in S.

Fig. 2 The 100 codes used as the rows and columns for the adjacency matrices

E. Column vectors of *A* are linearly independent
- E1. Only solution to $A\mathbf{x} = \mathbf{0}$ is trivial solution
- E2. There is a unique soln to system/matrix eqn
- E3. Can't get back home / to origin w/ cols. of *A*
- E4. Vectors aren't parallel/on the same line/plane
- E5. No vector is a scalar multiple of another
- E6. No vector is a linear combo of another
- E7. No vector is in the span of the other vectors
- E8. Don't have extra vector needed to return home

F. Column vectors of *A* are linearly dependent
- F1. Is more than one solution to $A\mathbf{x} = \mathbf{0}$.
- F2. No unique/infinitely many solns to system or matrix equation
- F3. Can get back home/ to a point with cols of *A*
- F4. Vectors are parallel/on the same line or plane
- F5. One vector is a scalar multiple of another
- F6. One vector is a linear combo of others
- F7. A vector is in the span of the other vectors
- F8. The matrix *A* has a row or column of zeroes
- F9. Have an extra vector needed to return home

G. Column vectors of *A* span $\mathbf{R}^n$
- G1. Are enough vectors to span the entire space
- G2. Can get to every pt/go everywhere
- G3. Is a linear combo of vectors for all pts in $\mathbf{R}^n$
- G4. Can use each vector to go in a direction
- G5. There is a solution to $A\mathbf{x} = \mathbf{b}$ for every $\mathbf{b}$

H. Column vectors of *A* do not span $\mathbf{R}^n$
- H1. Aren't enough vectors to span the space
- H2. Can't get to every pt/go everywhere
- H3. Cannot go in all directions with the vectors
- H4. The vectors of *A* span a *k*-dim subspace of $\mathbf{R}^n$
- H5. Span is only a point/line/plane
- H6: There is not a solution to $A\mathbf{x} = \mathbf{b}$ for every $\mathbf{b}$

I. Row-reduced echelon form of *A* has *n* pivots
- I1. RREF(*A*) has all ones on the main diagonal
- I2. Can row-reduce / is row equivalent to *I*
- I3. Is a pivot in each row
- I4. Is a pivot in each column
- I5. Each variable is defined in system/matrix eqn
- I6. RREF(*A*) has no rows of zeroes

J. Row-reduced echelon form of *A* has < *n*
- J1. RREF(*A*) does not have all ones on the main diagonal
- J2. Cannot row-reduce /not row equivalent to *I*
- J3. Is not a pivot in each row
- J4. Is not a pivot in each column
- J5. Not every variable is defined in system
- J6. RREF(*A*) has at least one row of zeroes

K. *A* is invertible
- K1. $[A \mid I] \sim [I \mid A^{-1}]$ is possible
- K2. Can calculate, no "divide by 0 errors"
- K3. Can undo/get back from the transformation
- K4. Exists a *C* s.t. *AC*=*I* and/or *CA*=*I*
- K5. Exists seq. of elem row ops that turns *A* into *I*
- K6. Something gets sent to $\mathbf{e}_1, \mathbf{e}_2, \ldots$

L. *A* is not invertible
- L1. $[A \mid I] \sim [I \mid A^{-1}]$ is not possible
- L2. Can't calculate, get "divide by 0 errors"
- L3. Can't undo/get back from the transformation
- L4. Does not exist a *C* s.t. *AC*=*I* and/or *CA*=*I*
- L5. Doesn't exist seq of elem row ops to turn *A* into *I*
- L6. Nothing gets sent to $\mathbf{e}_1, \mathbf{e}_2, \ldots$

M. The transformation defined by *A* is onto
- M1. For every $\mathbf{b}$ there is at least one $\mathbf{x}$ such that $T(\mathbf{x})=\mathbf{b}$
- M2. Range is all of the codomain
- M3. All codomain is mapped to as outputs/images

N. The transformation defined by *A* is not onto
- N1. For every $\mathbf{b}$ there is not at least one $\mathbf{x}$ s.t. $T(\mathbf{x})=\mathbf{b}$
- N2. Range is not all of the codomain
- N3. Not all $\mathbf{b}$ in codomain get used/mapped to as outputs/images
- N4. Transformation collapses everything to a point/line/plane
- N5. The transformation goes up/adds a dimension

O. The transformation defined by *A* is 1-1
- O1. For every $\mathbf{b}$ there is at most one $\mathbf{x}$ s.t. $T(\mathbf{x})=\mathbf{b}$
- O2. Each output has at most one input
- O3. There is only 1 way to "get to" the output/vector

P. The transformation defined by *A* is not 1-1
- P1. For every $\mathbf{b}$ there is more than one $\mathbf{x}$ s.t. $T(\mathbf{x})=\mathbf{b}$
- P2. Each output has more than one input
- P3. There is more than one way to "get to" the output/vector
- P4. At least 2 inputs give the same output
- P5: The transformation goes down/excludes a dimension

Q. det(*A*) ≠ 0: Determinant of *A* is nonzero
- Q1. Unit square/cube has area/volume ≠ 0 after transformation
- Q2. Formula to calculate determinant yields nonzero result

R. det(*A*) = 0: Determinant of *A* is equal to 0
- R1. Unit square/cube has area/volune= 0 after transformation
- R2. Formula to calculate determinant yields zero
- R3. No area/volume to a line/plane

S. Miscellaneous
- S2. The Col *A* is all of $\boldsymbol{R}^n$
- S3. The Col *A* is all not of $\boldsymbol{R}^n$
- S4. The Nul *A* contains only the zero vector
- S5. The Nul *A* contains more than the zero vector
- S6. The number zero is not an eigenvalue of A
- S7. The number zero is an eigenvalue of A
- S8: *A* has more vectors than dimensions
- S9: *A* has less vectors than dimensions
- S10: *A* has the same number of rows/columns
- S11: *A* does not have the same number of rows and columns
- S12: Miscellaneous

vectors in $\mathbb{R}^m$, or a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$) were compiled. This is in contrast to dividing T into sub-digraphs (and corresponding sub-adjacency matrices) for arguments specific to $m < n$, $m = n$, or $m > n$. Only the adjacency matrix $A(\mathrm{T})_{\mathrm{tot}}$ is given in this paper; because of its size, the adjacency matrix was partitioned into four.

Fig. 3 Summary of the 10 class days relevant to the chronological development of the IMT

Day	**Chronological Development of the IMT**
Days 5	Students explored the concepts of span, linear dependence, and linear independence.
Day 6	The class developed generalizations about linear independence and dependence.
Day 9	"Theorem 4" was developed in class: Suppose A is an m x n matrix. The following are equivalent: a. The columns of A span $\mathbf{R}^m$ b. A has a pivot in each row c. For every **b** in $\mathbf{R}^m$, **b** can be written as a linear combo of the columns of A. d. For every vector **b** in $\mathbf{R}^m$, there exists a solution to the equation $A\mathbf{x} = \mathbf{b}$.
Day 10	Theorem 4 was altered to address square matrices only. New equivalences were added ((a)-(g) from Figure 1). And became known as the "New Theorem."
Day 17	The class developed conjectures about which matrices are invertible.
Day 18	Equivalences (h), (i), and (j) were added to the New Theorem. It was renamed "The Invertible Matrix Theorem" in class (see Figure 1).
Day 19	The class explored the concepts of one-to-one (1-1) and onto transformations.
Day 20	Upon investigating transformations that are both 1-1 and onto, (k)-(l) were added to the IMT. (m)-(n) were added in homework.
Day 24	Students investigated determinants, and (o) was added to the IMT. (p) was added in homework.
Day 31	Students discussed the IMT in small group and in whole class.

Figure 4a is the upper left quadrant of $A(\mathrm{T})_{\mathrm{tot}}$, Fig. 4b is the lower left, Fig. 4c is the upper right, and Fig. 4d is the lower right.

Each numerical entry within a cell in $A(\mathrm{T})_{\mathrm{tot}}$ indicates the frequency that particular adjacency occurred, and the subscript for the numerical value indicates the day on which it occurred. For instance, the cell in the "K" row and "F" column is "$2_{17}1_{24}$." This means that the implication "A is not invertible → the columns of A are linearly dependent," coded "K → F," occurred twice on Day 17 and once on Day 24. Because $A(\mathrm{T})_{\mathrm{tot}}$ compiles connections between vertices for any m and n, care should be given when making conclusions from $A(\mathrm{T})_{\mathrm{tot}}$. For instance, consider "the columns of span $\mathbb{R}^n$ then the columns are linearly independent," coded as "G → E." It is a perfectly valid implication in the $m = n$ type, but it is invalid in the $m < n$ type (for instance, a set of three vectors in $\mathbb{R}^2$ can span $\mathbb{R}^2$ but cannot be linearly independent). My discussion of the results within this paper reflects that. In particular, I present three categories of results: shifts in argumentation over the course of the semester, adjacencies within main codes and within subcodes, and centrality. Each category relates to both microgenetic and ontogenetic analyses of student reasoning about the IMT.

5.1 Shifts in argumentation over the course of the semester

The first category of results illuminates shifts in argumentation over the semester regarding how the classroom reasoned about the IMT. By considering where in the adjacency matrix entries appear, in conjunction with on which days they appeared (Fig. 4a–d), a sense of the "travel of ideas" (Saxe et al. 2009) is revealed. First, the subscripts in $A(\mathrm{T})_{\mathrm{tot}}$ illuminate the movement throughout the semester. Through this large grain size, there is a high concentration of k_5 and k_9 (where "k" is the integer value indicating the frequency of the particular adjacency and "d" is the day on which it occurred) in the upper left of the adjacency matrix; a concentration of k_{10} in the I block; a concentration of k_{19} and k_{20} in the M, N, O, and P blocks; k_{24} in the Q and R blocks, and k_{31} scattered throughout. We consider k_5, k_9, k_{19} and k_{20} in more detail.

Given the chronology of concepts developed in class and what tasks were posed for discussion each day, the aforementioned concentrations are sensible. First, the concentration of k_5 and k_9 in the upper left corner of $A(\mathrm{T})_{\mathrm{tot}}$ indicates frequent explanation within the F category early in the semester. For instance, on Day 9, a student, Gil (all names are pseudonyms), claimed you could always create a linearly dependent set. When asked why, he stated the following (the adjacency matrix codes are given in bold, as explained in Sect. 5):

> If you have anything from 1 or greater vectors, if 1 of those vectors happens to be 0, it's dependent (**F8** → **F**). And then if you have 2 or more vectors, if 2 of those vectors are the same … you can use 1 vector to go out and 1 vector to come back, so linear dependent that way (**F5** → **F3** → **F**).

In lines 1–2, Gil discusses how having the zero vector in a set (code F8, Fig. 2) gives linear dependence (code F), but he does not provide data for that claim. In lines 2–4, he states a set that has at least two that are "the same vector"

(a)

	E	E1	E2	E3	E4	E5	E6	E7	E8	F	F1	F2	F3	F4	F5	F6	F7	F8	F9	G	G1	G2	G3	G4	G5	H	H1	H2	H3	H4	H5	H6	I	I1	I2	I3	I4	I5	I6	J	J1	J2	J3	J4	J5	J6	K	K1	K2	K3
E		$1_{9}1_{17}1_{31}$	1_{10}																	$1_{5}2_{17}2_{20}1_{24}1_{31}$		1_{17}		1_{5}							1_{5}		2_{31}		$2_{17}1_{18}$	3_{19}											$3_{18}3_{24}$			
E1	$1_{17}1_{31}$																																																	
E2	$2_{10}1_{31}$																																																	
E3	1_{5}																			1_{10}																														
E4	1_{10}																		1_{5}	1_{5}		2_{5}																												
E5	1_{5}																			2																														
E6																																																		
E7	1_{5}																																																	
E8																																																		
F										1_{5}	1_{20}															$2_{17}1_{20}$					$1_{17}1_{24}$														1_{19}		1_{24}			
F1										1_{9}																																								
F2										1_{9}																																								
F3										1_{9}					1_{5}					1_{9}																														
F4					1_{5}					$2_{5}1_{9}$			1_{5}															1_{9}																						
F5										3_{5}			1_{9}	1_{9}													1_{5}																		1_{17}	1_{19}				
F6																																																		
F7													1_{5}																																					
F8										$1_{9}1_{24}$																2_{19}		1_{17}															1_{17}							
F9										1_{5}				1_{5}																																				
G	$2_{5}4_{10}1_{20}1_{31}$			1_{5}						$1_{9}2_{17}$		2_{10}							3_{5}			$1_{9}1_{20}1_{31}$	1_{31}		1_{31}											1_{19}		1_{10}							1_{5}		1_{24}			
G1																			3_{5}	1_{9}																														
G2									1_{5}										2_{5}	1_{31}																														
G3				1_{10}															1_{5}	1_{31}		2_{31}			1_{31}																									
G4														1_{5}						$1_{5}3_{9}$	1_{5}									1_{9}																				
G5																				1_{31}		1_{31}																												
H																																																		
H1																										$1_{9}1_{20}$		1_{9}		1_{9}																				
H2																																																		
H3																																																		
H4																																																		
H5																												1_{20}																						
H6																																																		
I																				$2_{31}2_{19}1_{5}$				1_{5}																										
I1			2_{10}																	1_{10}																														
I2																																															1_{18}			
I3																				$1_{10}1_{19}4_{20}2_{5}$				2_{5}									1_{10}	1_{10}			1_{10}	2_{10}												
I4				1_{10}																																														
I5																																							2_{10}								1_{18}			
I6																																																		
J																													1_{5}																					
J1																																												1_{5}						
J2																																																		
J3										1_{10}															1_{5}				3_{5}																					
J4										1_{5}																																			3_{5}					
J5										1_{5}		$1_{20}1_{5}$							1_{5}	1_{5}																										1_{17}				
J6										1_{24}																																								
K	$3_{17}2_{18}1_{24}$																			4_{17}															2_{17}				1_{24}									1_{24}		
K1																																																		
K2																																																		
K3																																																		

◀**Fig. 4** **a** Adjacency matrix $A(\mathrm{T})_{\mathrm{tot}}$: upper left quadrant, **b** adjacency matrix $A(\mathrm{T})_{\mathrm{tot}}$: lower left quadrant, **c** adjacency matrix $A(\mathrm{T})_{\mathrm{tot}}$: upper right quadrant, **d** adjacency matrix $A(\mathrm{T})_{\mathrm{tot}}$: lower right quadrant

(code F5) is linearly dependent (code F) because you can travel out on one vector and back on the other (F3), which is geometric interpretation of linear dependence.

This type of argument was common early in the semester, when the class was developing ways to reason about linear dependence; however, this reliance on explanations within the F category became less prominent over time. The various F codes moved away from being adjacent to or adjacent from other F codes towards making claims about (i.e., being adjacent to) other concepts in the IMT. Through an ontogenetic analysis lens, the community's ways of reasoning about linear dependence became normative over time as they did not need to unpack what "linear dependence" means each time it was discussed. This coincides with Criterion 1 (Rasmussen and Stephan 2008) for documenting normative ways of reasoning: when justifications for claims are initially present but then drop off.

Second, consider the concentration of k_{19} and k_{20} in the M, N, O, and P blocks; these correspond to discussing one-to-one (1-1) or onto[4] transformations. On these days, there was a high concentration in the rows for the G and H categories and in the columns for the M and N categories. This implies a high frequency of adjacencies from the interpretations of column vectors spanning or not spanning all of $\mathbb{R}^n$ leading to conclusions about if transformations are or are not onto $\mathbb{R}^n$. There are far fewer adjacencies from rows for the M and N categories to columns for the G and H categories (6 and 2, respectively). This means that "the transformation defined by A is not onto" was the claim more often than data. Considered ontogenetically, this is compatible with a normative way of reasoning of using the span of vectors to draw conclusions about onto transformations.

5.2 Adjacencies within main codes and within subcodes

The concepts involved in the IMT were negotiated at various times during the semester, and the early-developed concepts were often integral to the development of latter ones. Furthermore, the structure of argumentation was not uniform over time. For instance, one main structure that surfaced was that when an argument served towards developing a way of reasoning about a new concept or implication between two concepts, it involved multiple uses of the "interpretation" subcodes; the class's reliance on subcodes to explain the implication subsequently dropped off over time.

For instance, on Day 20, the class was parsing out the relationship between 1-1 and onto (which are properties of linear transformations) and linear independence and span (which are properties of sets of vectors). Previously, Abraham had volunteered, "If it's 1-1, the columns of A have to be linear independent" (coded O → E). The instructor asked another student, David, how his small group discussed why that implication might be true:

David: We just went over his explanation, we're having a hard time why that works, why that makes it a multiple solution, it's not 1-1 (**F2** → **P**).

Instructor: So he's saying it makes sense but he's having a hard time explaining it…Brad, can you tell me what your table talked about for this one?

Brad: When you reduce matrices linear dependent, you're going to have a free variable (**F** → **J5**). When you have that free variable, there has to be more than one input to get the same output (→**P4**).

Instructor: I think that's a great start. Does that make sense? So remember those, when you took a matrix and you row reduced it to the row-reduced echelon form, if you got a free variable, then that gave you some leniency for how you could put down a solution for what was in the span of the vectors (**J5** → **F2**). Well, that gives you some variability as to how you're going to answer. That variability is giving you more than one way to get to that certain output (→**P3**), not a unique solution (→**P1**).

In his response, David voiced that his group struggled with why "if the columns of A are linearly dependent" then "the transformation defined by A is not 1-1" (lines 5–7), so the instructor asked Brad to contribute to the explanation as well (lines 11–23). The implication "F → P" was still emerging in this community at this time, evidenced through the variety of interpretations of these concepts utilized to justify the proposed implication. Each of the six adjacencies in lines 7, 12, 14, 20, 22, and 23 involved interpretations of F or P, as well as "there exists a free variable" (J5), which served as an intermediary vertex between interpretations of both F and P. Thus, this argumentation served towards developing a way of reasoning about a new connection between concepts (namely linear dependence and not being 1-1) that involved multiple uses of interpretation subcodes.

[4] $T:\mathbb{R}^n \to \mathbb{R}^m$ is *one-to-one* if every $x \in \mathbb{R}^n$ has at most one $b \in \mathbb{R}^m$ such that $T(x) = b$ and *onto* if every $x \in \mathbb{R}^n$ has at least one $b \in \mathbb{R}^m$ such that $T(x) = b$.

(b)

	E	E1	E2	E3	E4	E5	E6	E7	E8	F	F1	F2	F3	F4	F5	F6	F7	F8	F9	G	G1	G2	G3	G4	G5	H	H1	H2	H3	H4	H5	H6	I	I1	I2	I3	I4	I5	I6	J	J1	J2	J3	J4	J5	J6	K	K1	K2	K3
K4																																															$2_{18}1_{31}$			
K5																																			1_{18}															
K6																																																	
L										$2_{17}1_{24}$																																								
L1																																														1_{24}				
L2																																																	
L3																																																	
L4																																																	
L5																																																	
L6																												1_{17}																						
M	1_{20}																			$4_{20}1_{19}$																											1_{20}			
M1																																																	
M2																				1_{19}																														
M3																				1_{31}																														
N																										1_{19}																								
N1																																																	
N2																																																	
N3																										1_{19}																								
N4																																																	
N5																																																	
O	3_{20}																			1_{20}																											1_{20}			
O1	1_{20}																																																
O2																																																	
O3																																																	
P																																																	
P1																																																	
P2																																																	
P3																																																	
P4																																																	
P5																																																	
Q	2_{24}																																														1_{24}			
Q1																																																	
Q2																																																	
R	1_{24}									6_{24}																																								
R1										$1_{24}1_{31}$																																								
R																																										1_{24}								
R3																																																	
S																																																	
S1																		1_{17}																																
S2																				1_{31}		1_{31}	1_{31}		1_{31}																									
S3																																																	
S4																																																	
S5																																																	
S6																																																	
S7																																																	
S8										$1_{5}4_{20}$			2_{5}											1_{5}																										
S9																										1_{9}	$1_{9}1_{20}$			1_{20}												1_{9}								
S10			1_{10}							3_{9}									1_{5}	$2_{10}1_{9}$																									1_{9}						
S11																																																	
S12																																																	

Fig. 4 continued

(c)

	K4	K5	K6	L	L1	L2	L3	L4	L5	L6	M	M1	M2	M3	N	N1	N2	N3	N4	N5	O	O1	O2	O3	P	P1	P2	P3	P4	P5	Q	Q1	Q2	R	R1	R	R3	S	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
E				$1_{24}1_{17}$							1_{20}						1_{20}		1_{24}		6_{20}			1_{20}							1_{24}																			
E1																																																		
E2																																																		
E3																																																		
E4																																																		
E5																																																		
E6																																																		
E7																																																		
E8																																																		
F				$8_{17}4_{24}1_{31}$														1_{20}	1_{24}						3_{20}									6_{24}		1_{24}			1_{17}											
F1																													1_{20}																					
F2																									2_{20}			1_{20}																						
F3																																																		
F4																																																		
F5				1_{17}														1_{17}							1_{20}																									
F6																			1_{24}																															
F7																																																		
F8				2_{17}			1_{17}	1_{17}							1_{19}			4_{19}																					2_{17}											
F9				1_{17}																																										1_{5}				
G											$7_{20}4_{19}$		1_{19}								1_{20}																			2_{31}										
G1																																																		
G2											1_{20}			1_{19}																																				
G3																																								1_{31}										
G4																																																		
G5														1_{31}																										1_{31}										
H				1_{17}											$2_{20}3_{19}$			1_{19}																																
H1																																																		
H2															1_{20}			1_{20}																																
H3																																																		
H4																		1_{20}																																
H5										2_{17}					1_{19}			1_{20}	$1_{17}2_{19}$																1_{24}															
H6																																																		
I																																																		
I1																																																		
I2		1_{18}									1_{20}										1_{20}																													
I3											$1_{20}1_{19}$										1_{20}																													
I4																																																		
I5																																																		
I6																																																		
J																																																		
J1																																																		
J2				1_{24}																																														
J3																		1																																
J4																																																		
J5																													1_{20}																					
J6				11_{24}																					1_{20}																									
K	$2_{17}1_{18}1_{31}$																														1_{24}																	1_{17}		
K1																																																		
K2																																																		
K3																																																		

Fig. 4 continued

(d)

	K4	K5	K6	L	L1	L2	L3	L4	L5	L6	M	M1	M2	M3	N	N1	N2	N3	N4	N5	O	O1	O2	O3	P	P1	P2	P3	P4	P5	Q	Q1	Q2	R	R1	R	R3	S	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
K4																																																1_{17}		
K5	1_{18}																																																	
K6																																																		
L																																		1_{24}																
L1																																																		
L2																																																		
L3																																																		
L4				1_{17}																																														
L5																																																		
L6				2_{17}			1_{17}																																											
M												2_{20}	2_{20}								5_{20}				2_{20}																									
M1														1_{20}																																				
M2											1_{20}										1_{20}																													
M3																																																		
N																					2_{20}				1_{20}																									
N1															1_{19}																																			
N2																																																		
N3											1_{20}				$1_{19}2_{20}$		1_{19}																																	
N4										2_{17}																									2_{24}		3_{24}													
N5																																																		1_{20}
O											6_{20}				2_{20}							1_{20}																												
O1																					1_{20}																													
O2																					1_{20}																													
O3																																																		
P											2_{20}				1_{20}																																			
P1																																																		
P2																																																		
P3																										1_{20}																								
P4																									1_{19}																									
P5																									1_{20}																									
Q																																																		
Q1																																																		
Q2																																																		
R				2_{24}		1_{24}																													1_{31}															
R1																																		1_{24}																
R																																		1_{24}																
R3																																			1_{24}															
S																																																		
S1				2_{17}			1_{17}	3_{17}																																										
S2																																																		
S3																																																		
S4																																																		
S5																																																		
S6																																																		
S7																																																		
S8																									2_{20}																									
S9															1_{20}																																			
S10													1_{20}																																					
S11																																																		
S12																		1_{20}																																

Fig. 4 continued

Furthermore, consider the implication "F → P," which only occurred three times during the semester, each on Day 20 (see the "3_{20}" in row F/column P of $A(\mathrm{T})_{\mathrm{tot}}$ in Fig. 4c). Each occurred after the aforementioned argument, and none contained any justification of that implication. Thus, there is an indication that "if the columns of *A* are linearly dependent, then the associated transformation is not 1-1" eventually functioned as if shared within this classroom.

A second prevalent structure of argumentation that surfaced via adjacency matrix analysis was that when the argumentation made use of relatively well-established concepts or connections between concepts, the argument involved mainly the "concept statement" main codes, rather than substantive use of the interpretive subcodes. This is somewhat illustrated in the previous example of "F → P." As an additional example, consider a portion of discourse from Day 20.

Instructor: So Lawson was saying something about being linear independent, is that the same as being able to say something about being onto?

Abraham: It is, if the matrix is square (**E** → **M**, **M** → **E**)

Instructor: Yeah, I think that's something that we need to get at. So the one that we had from before was if the columns of *A* span $\mathbb{R}^n$ (**G**), so the transformation T is onto $\mathbb{R}^m$ (→**M**). So now we're saying, can we say anything about connecting onto to linear independence? And Abraham's talking about we can if they're square. So I think I agree, can you say a little bit more?

Abraham: I just remember if it's square, we had the $n \times n$ Theorem way back when. And if a square matrix is linear independent (**E**), it also spans (→**G**). And if it spans (**G**), it's also linear independent (→**E**). And so that means that if it's 1-1 (→**O**), it has to be onto (→**M**); if it's onto (**M**), it has to be 1-1 (→**O**). Do you know what I mean, like connecting the ideas?

Abraham stated that Lawson's claim was true if the matrix is square (line 28). Lawson's claim of "the same as" was interpreted as equivalence of the concepts, thus Abraham's agreement in line 27 was coded both E → M and M → E. Abraham's justification of this claim relied on two more equivalences for square matrices: between span and linear independence ((G, E), and (E, G), lines 35–37), and between onto and one-to-one ((O, M) and (M, O), lines 37–39). In addition, this is a different structure of reasoning than in the example given in lines 5–23. In lines 24–39, although the claim Lawson suggested and Abraham justified was novel (that onto and linear independence were equivalent for square matrices), the justification Abraham provided involved previously established ways of reasoning. For example, Abraham referenced the "$n \times n$ Theorem [from] way back when" to justify his claim that linear independence implies span for column vectors of a square matrix. Abraham's justification in lines 34–39 utilized previously established connections between main codes rather than subcodes, which are often used to "unpack" or "interpret" the main codes, as in lines 5–23.

5.3 Centrality of concepts over time

The final results presented investigate what IMT concepts statements (main codes in Fig. 2) or interpretations of those concept statements (subcodes in Fig. 2) were prominent in argumentation over time. This is synchronously both ontogenetic and microgenetic analysis because considering the behavior of a code over time is ontogenetic but then investigating individual instances of that code, which contributed to its behavior over time, is microgenetic. This analysis is also consistent with what Tiberghien and Malkoun (2009) refer to as continuity—the distribution of utterances most reused during a theme or sequence. To accomplish this analysis, I use a measure of *centrality*.

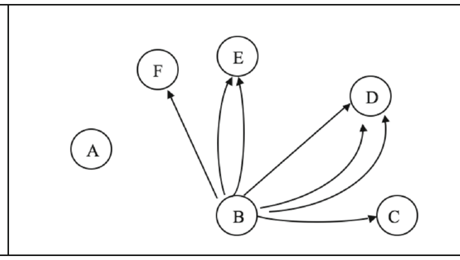
The construct of centrality is used to indicate "a node's degree of participation in the structure of the graph by measuring the relative connectivity of a node within a graph" (Strom et al. 2001, p. 752). It measures how central a vertex is—does it serve as a sort of "hub," connected to multiple vertices, or is it connected to a low number of vertices but with high edge frequency? For example, consider two different digraphs with six vertices and seven edges (Fig. 5). Vertex B is adjacent to four vertices in Fig. 5a but only two vertices in Fig. 5b. Centrality is one method of parsing how the participation of B differs in the two examples.

The *out-degree* of vertex v (od v) is the number of edges emanating from vertex v, whereas the *out-connection* of v (oc v) is the number of distinct vertices to which these edges emanate. The *in-degree* of v (id v) is the number of edges to v, whereas the *in-connection* of v (ic v) is the number of distinct vertices from which these edges emanate. Finally, I measure the *centrality* of a vertex v by $C(v) = (\mathrm{ic}\ v + \mathrm{oc}\ v)/2r$, where r is the total number of vertices in the graph. For example, $C(\mathrm{B}) = (0 + 4)/2(7) = 4/14$ in Fig. 5a, but $C(\mathrm{B}) = (0 + 1)/2(7) = 1/14$ in Fig. 5b. Note that $0 \le C(v) \le 1$; $C(v) = 0$ if vertex v is not connected to the rest of the graph, and $C(v) = 1$ if vertex v is adjacent to and adjacent from every vertex (including itself) in the graph; thus, the closer the value of $C(v)$ is to zero, the less central the vertex v is in the graph.

For the data regarding the community's reasoning about the IMT, Fig. 6 provides the centrality measure for each main code, subcode, and category (a code grouped with its subcodes) for the adjacency matrix in Fig. 4a–d.

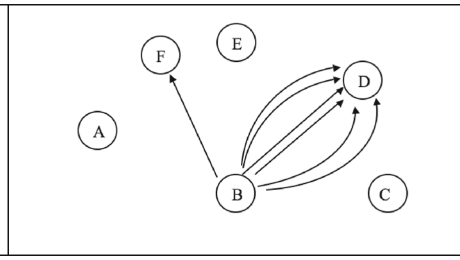
Fig. 5 Two examples of a digraph (and their associated adjacency matrices) such that the out-degree of vertex B is 7

(a)



	A	B	C	D	E	F
A						
B			1	3	2	1
C						
D						
E						
F						

(b)



	A	B	C	D	E	F
A						
B				6		1
C						
D						
E						
F						

The rightmost column in Fig. 6 shows that categories F and G had the highest and second highest centrality, respectively, in $A(\mathrm{T})_{\mathrm{tot}}$. Furthermore, Fig. 6 can be read as a comparison between categories and their negations. For instance, category E involves "the columns of A are linearly independent" and interpretations of that, whereas category F relates to linear dependence. First, comparing categories and their negations reveals that three categories as stated in the IMT were more central than their negations in whole class discussion: "The columns of A span $\mathbb{R}^n$" (category G); "The row-reduced echelon form of A has n pivots" (category I); and "The transformation defined by A is 1-1" (category O). Second, four categories from the IMT were less central in whole class discussion than their negations: "The columns of A are linearly dependent" (category F); "The matrix A is not invertible" (category L); "The transformation defined by A is not onto $\mathbb{R}^n$" (category N); and "The determinant of A is zero" (category R).

Thus, reasoning about the IMT also involved reasoning about negations of the statements involved in the theorem. This could have occurred as the justifications for an implication were given in the form of a contrapositive. Additionally, some concepts may lend to reasoning about their negations, such as determinants. The centrality of "the determinant of A is zero" was drastically higher than "the determinant of A is nonzero." A deeper examination of the involved arguments reveals that the class reasoned about how determinants connected to other concepts by focusing more on matrices with zero determinants rather than nonzero.

Figure 6 also summarizes the centrality of each individual vertex for the adjacency matrix $A(\mathrm{T})_{\mathrm{tot}}$. It reveals that vertex G was the most central single vertex in class discourse related to the IMT, followed by vertex F. As a category (a main code and its subcodes), however, category F was more central than category G. Thus, although concept statement G ("the columns of A span $\mathbb{R}^n$") was most central throughout the semester, the concept statement and subcodes for "the columns of A are linearly dependent" were more central than that of span and its associated interpretations. One explanation for this could be that the notion of linear dependence had a wider variety of powerful interpretations for this particular classroom community than did span.

Figure 6 provides information about a vertex's connectivity in the various sub-digraph cases ($m < n$, $m = n$, and $m > n$). First, gray font indicates that a particular vertex was adjacent to or adjacent from other vertices in only one type of sub-digraph. For instance, all Q and R entries (both OD/OC and ID/IC) are in gray font because they appear in the $m = n$ sub-digraph but not the others. It is sensible that Q and R codes (determinant of A equals zero and does not equal zero, respectively) were used only in $m = n$ argumentation because determinant is only defined for square matrices. Second, if the adjacency (u, v) occurred in more than one sub-digraph, then the cell "od(u)" and the cell "in(v)" would be shaded gray; in other words, (u, v) was not mutually exclusive across sub-digraphs. The gray shaded cells in Fig. 6 indicate that for both F and G, neither in-connections nor out-connections were mutually exclusive across m, n types. Not only were their centrality measures the highest, indicating their participation in the structure of argumentation was central by being adjacent to or adjacent from the highest number of distinct nodes, codes F and G were adjacent to or adjacent from some of the same vertices in more than one m, n type. Thus, across two measures, F ("the columns of A are linearly dependent") and G ("the columns of A span $\mathbb{R}^n$") were central ways of reasoning for the collective during the semester.

vertex	total		total		centrality	category centrality
	OD	OC	ID	IC		
E	41	17	31	13	0.1807	0.3012
E1	2	1	3	1	0.0120	
E2	3	1	4	3	0.0241	
E3	2	2	3	3	0.0301	
E4	5	4	1	1	0.0301	
E5	3	2			0.0120	
E6						
E7	1	1			0.0060	
E8			1	1	0.0060	
F	35	13	39	18	0.1867	0.5181
F1	2	2	1	1	0.0181	
F2	4	3	4	2	0.0301	
F3	3	3	5	4	0.0422	
F4	6	4	3	3	0.0422	
F5	11	9	1	1	0.0602	
F6	1	1			0.0060	
F7	1	1			0.0060	
F8	17	10	1	1	0.0663	
F9	4	4	11	6	0.0602	
G	41	16	51	21	0.2229	0.4639
G1	4	2	2	2	0.0241	
G2	6	5	10	6	0.0663	
G3	7	6	3	3	0.0542	
G4	7	4	5	4	0.0482	
G5	5	5	3	3	0.0482	
H	7	3	11	7	0.0602	0.2169
H1	4	3	3	2	0.0301	
H2	2	2	5	5	0.0422	
H3						
H4	1	1	6	4	0.0301	
H5	9	6	4	3	0.0542	
H6						
I	6	2	3	2	0.0241	0.2349
I1	3	2	1	1	0.0181	
I2	4	4	6	6	0.0602	
I3	18	8	4	4	0.0723	
I4	1	1	1	1	0.0120	
I5	3	2	3	3	0.0301	
I6			3	3	0.0181	
J	1	1			0.0060	0.1807
J1	1	1			0.0060	
J2	1	1	1	1	0.0120	
J3	6	4	2	2	0.0361	
J4	4	2	2	2	0.0241	
J5	7	6	6	4	0.0602	
J6	4	3	3	3	0.0361	
K	20	8	16	9	0.1024	0.1506
K1			1	1	0.0060	
K2						
K3						

vertex	total		total		centrality	category centrality
	OD	OC	ID	IC		
K4	4	2	5	2	0.0241	
K5	2	2	1	1	0.0181	
K6						
L	4	2	30	12	0.0843	0.1627
L1	1	1			0.0060	
L2			1	1	0.0060	
L3			3	3	0.0181	
L4	1	1	4	2	0.0181	
L5						
L6	4	3	4	2	0.0301	
M	18	7	26	9	0.0964	0.1747
M1	1	1	2	2	0.0181	
M2	3	3	4	3	0.0361	
M3	1	1	3	3	0.0241	
N	4	3	16	9	0.0723	0.2169
N1	1	1			0.0060	
N2			2	2	0.0120	
N3	6	4	12	9	0.0783	
N4	7	3	6	4	0.0422	
N5	1	1			0.0060	
O	14	6	19	9	0.0904	0.1205
O1	2	2	1	1	0.0181	
O2	1	1			0.0060	
O3			1	1	0.0060	
P	3	2	14	9	0.0663	0.1084
P1			1	1	0.0060	
P2						
P3	1	1	1	1	0.0120	
P4	1	1	2	2	0.0181	
P5	1	1			0.0060	
Q	3	2	2	2	0.0241	0.0241
Q1						
Q2						
R	11	5	9	4	0.0542	0.1205
R1	3	2	5	4	0.0361	
R	2	2	1	1	0.0181	
R3	1	1	3	1	0.0120	
S						0.1024
S1	7	4	3	2	0.0361	
S2	4	4	4	3	0.0422	
S3						
S4						
S5						
S6						
S7						
S8	10	4	1	1	0.0301	
S9	6	5			0.0301	
S10	10	6	2	2	0.0482	
S11						
S12	1	1	1	1	0.0120	

Fig. 6 Summary information for adjacency matrix $A(\mathrm{T})_{tot}$

One vertex whose out-connection was mutually exclusive, however, was G2 "Can use all the vectors to get to every point/to get everywhere." This vertex was adjacent to five different vertices, and for any particular v_i, (G2, v_i) occurred only within one sub-digraph: (G2, F9), (G2, E8), (G2, G), (G2, M), and (G2, M3). The language in both F9 and E8 coincide with the travel language of G2, all of which grew out of the class's experience with the Magic Carpet Ride Problem (Wawro et al. 2012). Thus, it is reasonable that students would, early in the semester, reason that a set of two vectors in $\mathbb{R}^2$ are linearly independent because they can get everywhere with the two vectors (G2) but do not have an extra vector needed to get back home (→E8). It is also sensible that reasoning about span and linear independence over time became less dependent on these situational ways of reasoning, moving more into general and formal ways of reasoning. This can be seen through the drop-off of the E8, F9, G2, and H2 codes and using the "parents codes" E, F, G, and H to make claims about new concepts such as invertibility or 1-1 transformations (see Fig. 4). Finally, the travel language of G2 did surface in argumentation concerning onto transformations ((G2, M) and (G2, M3)). As such, there is indication that interpreting span as "getting everywhere" was salient throughout the semester for this particular classroom community.

6 Discussion and conclusions

This study investigated how a linear algebra community reasoned about the invertible matrix theorem, and much of the theorem's "coming to be" is reflected in the ways in which the community constructed arguments about how the different concept statements were related. The results show the classroom community first developed rich ways of reasoning about what linear dependence means, and the reliance on unpacking that meaning decreased over time (Sect. 5.1). Analysis also showed that conclusions regarding "the transformation is (not) onto" most often resulted from claims regarding the span of vectors (Sect. 5.1). The collective also first developed a relationship between linear dependence and transformations that are not 1-1 by unpacking both concepts through various interpretations, until a stable relationship between the two was established (Sect. 5.2). Finally, the notion of centrality in Sect. 5.3 revealed that span and linear dependence were the most densely connected to the highest number of other concept statement throughout the semester, and that the "travel" interpretation for span was central throughout. Taken together, these results illuminate aspects of the research question regarding the collective's way of reasoning about the IMT.

Methodologically, this study adapts and extends the work of Selinski et al. (2013) by not only using adjacency matrices on data that spanned the entire semester (rather than at one moment in time) but also by analyzing at the collective level (rather than the individual). As such, it opens a door towards gaining new insights into the mathematical ways of reasoning of a classroom community, both within specific arguments and over time. The method, however, does involve possible limitations. First, although the codes are grounded in data, there is an inferential risk when coding discourse, and the "if–then" structure of adjacency matrix analysis could be restrictive when arguments do not follow that cleanly. Second, the codes are only about concept statements in the IMT or interpretations of those statements. Thus, utterances about other notions, such as self-reflection (e.g., "I prefer to reason about negations") or mathematics beyond the IMT (e.g., translating between matrix equation and vector equation notation), are not captured in the present study.

These results are compatible with those found via Toulmin's model of argumentation (Toulmin 1969) for documenting collective activity (Cole et al. 2012; Rasmussen and Stephan 2008). Recall that vertices F and G had high centrality measures, indicating their integral participation in the argumentation structure by being adjacent to or from many other vertices. Furthermore, vertices F and G were adjacent to or from other vertices in each of the m, n sub-digraphs, indicating their importance in the classroom argumentation in a variety of situations. Given that the adjacency (u, v) could be read as "if u then v," a vertex with a high out-connection means that the concept statement served as data for a variety of claims. This aligns with Criterion 3 for documenting normative ways of reasoning (Cole et al. 2012). High diversity within a particular cell (such as the (G, E) cell), rather than over a particular row, indicates that the given implication (rather than a given concept) served a role in multiple arguments. While the specifics of the argument that these adjacent pairs belong to is not provided in the adjacency matrix, knowing on which days they occurred provides information about what concepts the pairs are used to reason about. This allows an analysis compatible with Criterion 2 (Rasmussen and Stephan 2008).

In closing, I offer two implications for teaching. First, the class had two main ways of explaining why concept statements in the IMT were equivalent: they were indistinguishable and/or defined in terms of each other, or they are connected through a sequence of if–then deductions. For instance, consider "the columns of A are linearly independent" and "the only solution to $A\mathbf{x} = \mathbf{0}$ is the trivial solution." These are equivalent by definition. That is different than, say, "the columns of A span $\mathbb{R}^m$" and "the number zero is not an eigenvalue of A," which requires a more substantial effort to demonstrate their equivalence. Indeed, research on student understanding of proof highlights difficulty with identifying assumptions and creating logical arguments (e.g., Hoyles 1997), with attention specifically paid to "how to best enculturate students into a proving culture while … [taking] account of student views and ideas" (p. 7) in the classroom. The notion of equivalence in general is most likely new for many students, and, given their limited experience in proof, students would benefit from conversations that explicitly bring this distinction to light. Explicitly studying the role of the instructor in the development of mathematical meaning for this community is beyond the scope of the present study and remains a direction for future work.

Second, the concept statements in the IMT are matrix-, vector-, or transformation-oriented. This orientation shift is not unlike switching between various modes of representation (Hillel 2000). Teachers are often not aware when they ask students to move between modes, and, as Hillel stated, this is one aspect that makes linear algebra difficult to learn and to teach. Making explicit that the IMT is a collection of equivalent statements that all "say the same thing but in different ways" may help students see the power of the IMT, as well as how to leverage these subtleties in proof activity.

References

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science, 14*(1–2), 1–8.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical data analysis. *Mathematical Thinking and Learning, 1*(1), 5–43.

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*, 175–190.

Cole, R., Becker, N., Towns, M., Sweeney, G., Wawro, M., & Rasmussen, C. (2012). Adapting a methodology from mathematics education research to chemistry education research: Documenting collective activity. *International Journal of Science and Mathematics Education, 10*, 193–211.

Denzin, N. K. (1978). *The research act: A theoretical introduction to research methods*. New York: McGraw-Hill.

Dreyfus, T., Hillel, J., & Sierpinska, A. (1998). Cabri-based linear algebra: Transformations. In I. Schwank (Ed.), *Proceedings of the first conference of the European society for research in mathematics education* (Vol. I + II, pp. 209–221). Osnabrueck: Forschungsinstitut fuer Mathematikdidaktik.

Elbers, E. (2003). Classroom interaction as reflection: Learning and teaching mathematics in a community of inquiry. *Educational Studies in Mathematics, 54*(1), 77–99.

Ercikan, K., & Roth, W. M. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher, 35*(5), 14–23.

Freudenthal, H. (1991). *Revisiting mathematics education*. Dordrecht: Kluwer.

Harel, G. (1989). Learning and teaching linear algebra: Difficulties and an alternative approach to visualizing concepts and processes. *Focus on Learning Problems in Mathematics, 11*(1), 139–148.

Harel, G., & Brown, S. (2008). Mathematical induction: Cognitive and instructional considerations. In M. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 111–123). Washington: Mathematical Association of America.

Hillel, J. (2000). Modes of description and the problem of representation in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 191–207). Dordrecht: Kluwer.

Hoyles, C. (1997). The curricular shaping of students' approaches to proof. *For the Learning of Mathematics, 17*(1), 7–16.

Larson, C., Zandieh, M., & Rasmussen, C. (2008). A trip through eigen-land: Where most roads lead to the direction associated with the largest eigenvalue. In *Proceedings of the eleventh special interest group of the mathematical association of America on research in undergraduate mathematics education*, San Diego, CA.

Lay, D. C. (2003). *Linear algebra and its applications* (3rd ed.). Reading: Addison-Wesley.

Portnoy, N., Grundmeier, T. A., & Graham, K. J. (2006). Students' understanding of mathematical objects in the context of transformational geometry: Implications for constructing and understanding proofs. *Journal of Mathematical Behavior, 25*(3), 196–207.

Possani, E., Trigueros, M., Preciado, J. G., & Lozano, M. D. (2010). Use of models in the teaching of linear algebra. *Linear Algebra and its Applications, 432*(8), 2125–2140.

Rasmussen, C., & Stephan, M. (2008). A methodology for documenting collective activity. In A. E. Kelly, R. A. Lesh, & J. Y. Baek (Eds.), *Handbook of innovative design research in science, technology, engineering, mathematics (STEM) education* (pp. 195–215). New York: Taylor and Francis.

Saxe, G. B. (2002). Children's developing mathematics in collective practices: A framework for analysis. *Journal of the Learning Sciences, 11*(2–3), 275–300.

Saxe, G. B., & Esmonde, I. (2005). Studying cognition in flux: A historical treatment of fu in the shifting structure of Oksapmin mathematics. *Mind, Culture, and Activity, 12*(3–4), 171–225.

Saxe, G., Gearhart, M., Shaughnessy, M., Earnest, D., Cremer, S., Sitabkhan, Y., et al. (2009). A methodological framework and empirical techniques for studying the travel of ideas in classroom communities. In B. Schwarz, T. Dreyfus, & R. Hershkowitz (Eds.), *Transformation of knowledge through classroom interaction* (pp. 203–222). New York: Routledge.

Selden, A., & Selden, J. (1987). Errors and misconceptions in college level theorem proving. In J. D. Novake (Ed.), *Proceedings of the second international seminar on misconceptions and educational strategies in science and mathematics* (Vol. III, pp. 457–470). New York: Cornell University.

Selinski, N., Rasmussen, C., Wawro, M., & Zandieh, M. (2013). *A methodology for using adjacency matrices to analyze the connections students make between concepts in linear algebra*. Manuscript submitted for publication.

Sfard, A. (2007). When the rules of discourse change but nobody tells you: making sense of mathematics learning from a commognitive standpoint. *The Journal of the Learning Sciences, 16*(4), 565–613.

Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 209–246). Dordrecht: Kluwer.

Stephan, M., Cobb, C., & Gravemeijer, K. (2003). Coordinating social and individual analyses: Learning as participation in mathematical practices. *Journal for Research in Mathematics Education. Monograph, 12*, 67–102.

Stephan, M., & Rasmussen, C. (2002). Classroom mathematical practices in differential equations. *The Journal of Mathematical Behavior, 21*(4), 459–490.

Strom, D., Kemeny, V., Lehrer, R., & Forman, E. (2001). Visualizing the emergent structure of children's mathematical argument. *Cognitive Science, 25*(5), 733–773.

Tiberghien, A., & Malkoun, L. (2009). The construction of physics knowledge in the classroom from different perspectives: The classroom as a community and the students as individuals. In B. Schwarz, T. Dreyfus, & R. Hershkowitz (Eds.), *Transformation of knowledge through classroom interaction* (pp. 42–55). New York: Routledge.

Toulmin, S. (1969). *The uses of argument*. Cambridge: Cambridge University Press.

Wawro, M., Rasmussen, C., Zandieh, M., Sweeney, G., & Larson, C. (2012). An inquiry-oriented approach to span and linear independence: The case of the magic carpet ride sequence. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies, 22*(8), 577–599.

Wawro, M., Sweeney, G., & Rabin, J. M. (2011). Subspace in linear algebra: Investigating students' concept images and interactions with the formal definition. *Educational Studies in Mathematics, 78*(1), 1–19.

Weber, K. (2001). Student difficulty in constructing proofs: the need for strategic knowledge. *Educational Studies in Mathematics, 48*(1), 101–119.