NOTES

AN ASPIRATIONAL APPROACH TO THE MATHEMATICAL PREPARATION OF TEACHERS

James A. M. Álvarez
Elizabeth G. Arnold
Elizabeth A. Burroughs
Editors

MAA PRESS
NOTES / VOLUME 98

An Aspirational Approach to the Mathematical Preparation of Teachers

Electronic ISBN 978-1-61444-334-6

Published in the United States of America

An Aspirational Approach to the Mathematical Preparation of Teachers

Edited by

James A. M. Álvarez
The University of Texas at Arlington

Elizabeth G. Arnold
Colorado State University

and

Elizabeth A. Burroughs
Montana State University

MAA®

Published and Distributed by
The Mathematical Association of America

The MAA Notes Series, started in 1982, addresses a broad range of topics and themes of interest to all who are involved with undergraduate mathematics. The volumes in this series are readable, informative, and useful, and help the mathematical community keep up with developments of importance to mathematics.

MAA Notes

14. Mathematical Writing, by *Donald E. Knuth, Tracy Larrabee, and Paul M. Roberts.*
16. Using Writing to Teach Mathematics, *Andrew Sterrett,* Editor.
17. Priming the Calculus Pump: Innovations and Resources, Committee on Calculus Reform and the First Two Years, a subcomittee of the Committee on the Undergraduate Program in Mathematics, *Thomas W. Tucker,* Editor.
18. Models for Undergraduate Research in Mathematics, *Lester Senechal,* Editor.
19. Visualization in Teaching and Learning Mathematics, Committee on Computers in Mathematics Education, *Steve Cunningham and Walter S. Zimmermann,* Editors.
20. The Laboratory Approach to Teaching Calculus, *L. Carl Leinbach et al.,* Editors.
21. Perspectives on Contemporary Statistics, *David C. Hoaglin and David S. Moore,* Editors.
22. Heeding the Call for Change: Suggestions for Curricular Action, *Lynn A. Steen,* Editor.
24. Symbolic Computation in Undergraduate Mathematics Education, *Zaven A. Karian,* Editor.
25. The Concept of Function: Aspects of Epistemology and Pedagogy, *Guershon Harel and Ed Dubinsky,* Editors.
26. Statistics for the Twenty-First Century, *Florence and Sheldon Gordon,* Editors.
27. Resources for Calculus Collection, Volume 1: Learning by Discovery: A Lab Manual for Calculus, *Anita E. Solow,* Editor.
28. Resources for Calculus Collection, Volume 2: Calculus Problems for a New Century, *Robert Fraga,* Editor.
29. Resources for Calculus Collection, Volume 3: Applications of Calculus, *Philip Straffin,* Editor.
30. Resources for Calculus Collection, Volume 4: Problems for Student Investigation, *Michael B. Jackson and John R. Ramsay,* Editors.
31. Resources for Calculus Collection, Volume 5: Readings for Calculus, *Underwood Dudley,* Editor.
32. Essays in Humanistic Mathematics, *Alvin White,* Editor.
33. Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results, *James J. Kaput and Ed Dubinsky,* Editors.
34. In Eves' Circles, *Joby Milo Anthony,* Editor.
35. You're the Professor, What Next? Ideas and Resources for Preparing College Teachers, The Committee on Preparation for College Teaching, *Bettye Anne Case,* Editor.
36. Preparing for a New Calculus: Conference Proceedings, *Anita E. Solow,* Editor.
37. A Practical Guide to Cooperative Learning in Collegiate Mathematics, *Nancy L. Hagelgans, Barbara E. Reynolds, SDS, Keith Schwingendorf, Draga Vidakovic, Ed Dubinsky, Mazen Shahin, G. Joseph Wimbish, Jr.*
38. Models That Work: Case Studies in Effective Undergraduate Mathematics Programs, *Alan C. Tucker,* Editor.
39. Calculus: The Dynamics of Change, CUPM Subcommittee on Calculus Reform and the First Two Years, *A. Wayne Roberts,* Editor.
40. Vita Mathematica: Historical Research and Integration with Teaching, *Ronald Calinger,* Editor.
41. Geometry Turned On: Dynamic Software in Learning, Teaching, and Research, *James R. King and Doris Schattschneider,* Editors.

42. Resources for Teaching Linear Algebra, *David Carlson, Charles R. Johnson, David C. Lay, A. Duane Porter, Ann E. Watkins, William Watkins,* Editors.
43. Student Assessment in Calculus: A Report of the NSF Working Group on Assessment in Calculus, *Alan Schoenfeld,* Editor.
44. Readings in Cooperative Learning for Undergraduate Mathematics, *Ed Dubinsky, David Mathews, and Barbara E. Reynolds,* Editors.
45. Confronting the Core Curriculum: Considering Change in the Undergraduate Mathematics Major, *John A. Dossey,* Editor.
46. Women in Mathematics: Scaling the Heights, *Deborah Nolan,* Editor.
47. Exemplary Programs in Introductory College Mathematics: Innovative Programs Using Technology, *Susan Lenker,* Editor.
48. Writing in the Teaching and Learning of Mathematics, *John Meier and Thomas Rishel.*
49. Assessment Practices in Undergraduate Mathematics, *Bonnie Gold,* Editor.
50. Revolutions in Differential Equations: Exploring ODEs with Modern Technology, *Michael J. Kallaher,* Editor.
51. Using History to Teach Mathematics: An International Perspective, *Victor J. Katz,* Editor.
52. Teaching Statistics: Resources for Undergraduate Instructors, *Thomas L. Moore,* Editor.
53. Geometry at Work: Papers in Applied Geometry, *Catherine A. Gorini,* Editor.
54. Teaching First: A Guide for New Mathematicians, *Thomas W. Rishel*.
55. Cooperative Learning in Undergraduate Mathematics: Issues That Matter and Strategies That Work, *Elizabeth C. Rogers, Barbara E. Reynolds, Neil A. Davidson, and Anthony D. Thomas,* Editors.
56. Changing Calculus: A Report on Evaluation Efforts and National Impact from 1988 to 1998, *Susan L. Ganter*.
57. Learning to Teach and Teaching to Learn Mathematics: Resources for Professional Development, *Matthew Delong and Dale Winter.*
58. Fractals, Graphics, and Mathematics Education, Benoit Mandelbrot and Michael Frame, Editors.
59. Linear Algebra Gems: Assets for Undergraduate Mathematics, *David Carlson, Charles R. Johnson, David C. Lay, and A. Duane Porter,* Editors.
60. Innovations in Teaching Abstract Algebra, *Allen C. Hibbard and Ellen J. Maycock,* Editors.
61. Changing Core Mathematics, *Chris Arney and Donald Small,* Editors.
62. Achieving Quantitative Literacy: An Urgent Challenge for Higher Education, *Lynn Arthur Steen.*
64. Leading the Mathematical Sciences Department: A Resource for Chairs, *Tina H. Straley, Marcia P. Sward, and Jon W. Scott,* Editors.
65. Innovations in Teaching Statistics, *Joan B. Garfield,* Editor.
66. Mathematics in Service to the Community: Concepts and models for service-learning in the mathematical sciences, *Charles R. Hadlock,* Editor.
67. Innovative Approaches to Undergraduate Mathematics Courses Beyond Calculus, *Richard J. Maher,* Editor.
68. From Calculus to Computers: Using the last 200 years of mathematics history in the classroom, Amy Shell-Gellasch and Dick Jardine, Editors.
69. A Fresh Start for Collegiate Mathematics: Rethinking the Courses below Calculus, *Nancy Baxter Hastings,* Editor.
70. Current Practices in Quantitative Literacy, *Rick Gillman*, Editor.
71. War Stories from Applied Math: Undergraduate Consultancy Projects, *Robert Fraga*, Editor.
72. Hands On History: A Resource for Teaching Mathematics, *Amy Shell-Gellasch*, Editor.
73. Making the Connection: Research and Teaching in Undergraduate Mathematics Education, *Marilyn P. Carlson and Chris Rasmussen*, Editors.
74. Resources for Teaching Discrete Mathematics: Classroom Projects, History Modules, and Articles, *Brian Hopkins*, Editor.
75. The Moore Method: A Pathway to Learner-Centered Instruction, *Charles A. Coppin, W. Ted Mahavier, E. Lee May, and G. Edgar Parker*.
76. The Beauty of Fractals: Six Different Views, *Denny Gulick and Jon Scott*, Editors.
77. Mathematical Time Capsules: Historical Modules for the Mathematics Classroom, *Dick Jardine and Amy Shell-Gellasch*, Editors.
78. Recent Developments on Introducing a Historical Dimension in Mathematics Education, *Victor J. Katz and Costas Tzanakis*, Editors.
79. Teaching Mathematics with Classroom Voting: With and Without Clickers, *Kelly Cline and Holly Zullo*, Editors.
80. Resources for Preparing Middle School Mathematics Teachers, *Cheryl Beaver, Laurie Burton, Maria Fung, and Klay Kruczek*, Editors.
81. Undergraduate Mathematics for the Life Sciences: Models, Processes, and Directions, *Glenn Ledder, Jenna P. Carpenter, and Timothy D. Comar*, Editors.
82. Applications of Mathematics in Economics, *Warren Page*, Editor.
83. Doing the Scholarship of Teaching and Learning in Mathematics, *Jacqueline M. Dewar and Curtis D. Bennett*, Editors.
84. Insights and Recommendations from the MAA National Study of College Calculus, *David Bressoud, Vilma Mesa, and Chris Rasmussen,* Editors.
85. Beyond Lecture: Resources and Pedagogical Techniques for Enhancing the Teaching of Proof-Writing Across the Curriculum, *Rachel Schwell, Aliza Steurer and Jennifer F. Vasquez,* Editors.

86. Using the Philosophy of Mathematics in Teaching Undergraduate Mathematics, *Bonnie Gold, Carl E. Behrens, and Roger A. Simons,* Editors.
87. The Courses of History: Ideas for Developing a History of Mathematics Course, *Amy Shell-Gellasch and Dick Jardine,* Editors.
88. Shifting Contexts, Stable Core: Advancing Quantitative Literacy in Higher Education, *Luke Tunstall, Gizem Karaali, and Victor Piercey,* Editors.
89. MAA Instructional Practices Guide, *Martha L. Abell, Linda Braddy, Doug Ensley, Lewis Ludwig, Hortensia Soto,* Project Leadership Team.
90. What Could They Possibly Be Thinking!?! Understanding your college math students, *Dave Kung and Natasha Speer*.
91. Mathematical Themes in a First-Year Seminar, *Jennifer Schaefer, Jennifer Bowen, Mark Kozek, and Pamela Pierce,* Editors.
92. Addressing Challenges to the Precalculus to Calculus II Sequence through Case Studies: Report based on the National Science Foundation Funded Project Precalculus through Calculus II, *Estrella Johnson, Naneh Apkarian, Kristen Vroom, Antonio Martinez, Chris Rasmussen, and David Bressoud*, Editors.
93. Engaging Students in Introductory Mathematics Courses through Interdisciplinary Partnerships: The SUMMIT-P Model *Susan L. Ganter, Debra Bourdeau, Victor Piercey, and Afroditi V. Filippas,* Editors.
94. Expanding Undergraduate Research in Mathematics: Making UR more inclusive, *Michael Dorff, Jan Rychtář, Dewey Taylor,* Editors.
95. Sharing and Storing Knowledge about Teaching Undergraduate Mathematics: An Introduction to a Written Genre for Sharing Lesson-specific Instructional Knowledge, *Douglas Lyman Corey and Steven R. Jones*, Editors.
96. Justice Through the Lens of Calculus: Framing New Possibilities for Diversity, Equity, and Inclusion, *Matthew Voigt, Jess Ellis Hagman, Jessica Gehrtz, Brea Ratliff, Nathan Alexander, and Rachel Levy*, Editors.
97. Equitable and Engaging Mathematics Teaching: A Guide to Disrupting Hierarchies in the Classroom, *Daniel Reinholz*
98. An Aspirational Approach to the Mathematical Preparation of Teachers, *James A. M. Álvarez, Elizabeth G. Arnold, and Elizabeth A. Burroughs*, Editors.

MAA Service Center
P.O. Box 91112
Washington, DC 20090-1112
1-800-331-1MAA FAX: 1-301-206-9789

Acknowledgments

The work of the Mathematical Education of Teachers as an Application of Undergraduate Mathematics (META Math) project involved a network of mathematicians, statisticians, mathematics and statistics education researchers, MAA staff members, and research participants. A large part of the research and development of the META Math lessons was supported by the Mathematical Association of America and the National Science Foundation Division of Undergraduate Education (DUE-1726624). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The following people made direct contributions to the META Math project.

Project Team:

James A. M. Álvarez, *The University of Texas at Arlington*

Elizabeth G. Arnold, *Colorado State University*

Elizabeth A. Burroughs, *Montana State University*

Douglas E. Ensley, *Shippensburg University*

Elizabeth W. Fulton, *Montana State University*

Andrew Kercher, *Simon Fraser University*

Nancy Ann Neudauer, *Pacific University*

James Tanton, *Mathematical Association of America*

Rachel Tremaine, *Colorado State University*

Kyle Turner, *The University of Texas at Arlington*

Advisory Board:

Diane Briars, *Mathematics Education Consultant; Past President, National Council of Teachers of Mathematics (NCTM) 2014–2016*

Dave Kung, *Charles A. Dana Center at the University of Texas at Austin*

W. Gary Martin, *Auburn University*

Natasha Speer, *University of Maine*

Alan Tucker, *Stony Brook University*

External Evaluator:

Cynthia Schneider, *Independent Consultant*

We wish to thank the classroom instructors and their undergraduate students who implemented various versions of the META Math lessons. Their willingness to try new things and their feedback provided valuable insight that strengthened the quality of the lessons.

The following people were generous with their ideas about mathematics and statistics content and teaching approaches that informed the development of the META Math lessons.

Anna Bargagliotti, *Loyola Marymount University*
Ellen Breazel, *Clemson University*
Alex Burstein, *Howard University*
John Caughman, *Portland State University*
Richard Chandler, *University of North Texas at Dallas*
Jessica Deshler, *West Virginia University*
David Duncan, *James Madison University*
Brittney Falahola, *Stephen F. Austin State University*
Jess Ellis Hagman, *Colorado State University*
Natalie Hobson, *Sonoma State University*
Susan Hollingsworth, *Edgewood College*
Klay Kruczek, *Southern Connecticut State University*
Brigitte Lahme, *Sonoma State University*
Yvonne Lai, *University of Nebraska*
Jennifer Lovett, *Middle Tennessee State University*
Yanping Ma, *Loyola Marymount University*
Jim Madden, *Louisiana State University*
Ryota Matsuura, *St. Olaf College*
Victoria Noquez, *Indiana University Bloomington*
Cody Patterson, *Texas State University*
Janice Rech, *University of Nebraska Omaha*
Kathryn Rhoads, *The University of Texas at Arlington*
Dev Sinha, *University of Oregon*
Michael Smith, *Lewis University*
Francis Su, *Harvey Mudd College*
James Tanton, *Mathematical Association of America*
George Tintera, *Texas A&M University-Corpus Christi*
Michaela Vancliff, *The University of Texas at Arlington*
Sunita Vatuk, *City College of New York*
Wendy Weber, *Central College*
Ping Ye, *University of North Georgia*

Suggested Chapter Citation:

Chapter Authors (2024). Chapter Title. In J. A. M. Álvarez, E. G. Arnold, & E. A. Burroughs (Eds.), MAA Notes #98: An Aspirational Approach to the Mathematical Preparation of Teachers (pp. #–#). Mathematical Association of America.

Contents

1

The Why and How of META Math Lessons

Elizabeth A. Burroughs, *Montana State University*
Elizabeth G. Arnold, *Colorado State University*
James A. M. Álvarez, *The University of Texas at Arlington*

1.1 Introduction

Secondary mathematics teacher preparation routinely includes the study of both mathematics content and pedagogical strategies. One challenge for mathematicians who prepare prospective teachers is to structure mathematics courses to engage undergraduates[1] in mathematics content that is part of a mathematics major, while also attending to the particular mathematics content that will be most germane to those studying to be mathematics teachers. And though specific coursework designed to build pedagogical knowledge is often formally assigned as the responsibility of academic departments other than mathematics, the adage "teachers teach how they were taught" indicates that mathematicians cannot discount the role of mathematical learning experiences in the pedagogical preparation of teachers.

Mathematicians and mathematics teacher educators have collaborated for decades in formulating recommendations for mathematics teacher preparation. *The Mathematical Education of Teachers II* (MET II) report (CBMS, 2012), an update of the 2001 *Mathematical Education of Teachers* (CBMS, 2001), recommends that prospective secondary mathematics teachers complete the equivalent of an undergraduate major in mathematics with a focus on examining secondary school mathematics from an advanced perspective. This advanced perspective should focus on identifying and examining connections between the mathematics prospective teachers are learning and the school mathematics they will be teaching. The *Standards for Preparing Teachers of Mathematics* (AMTE, 2017) embraces this recommendation as well, with the imperative that secondary teachers gain "solid and flexible knowledge of relevant mathematical concepts and procedures from the high school curriculum, including connections to material that comes before and after high school mathematics" (p. 122).

At the same time, research studies have found evidence that prospective secondary mathematics teachers complete their undergraduate mathematics degree without having a deep understanding of secondary mathematics content (Moreira & David, 2008; Speer et al., 2015) and perceive their undergraduate mathematical study as disconnected from the practice of teaching secondary mathematics (Winsor et al., 2020; Zazkis & Leikin, 2010). Stylianides and Stylianides (2014) articulate the need for teacher preparation programs to incorporate specific knowledge needed for teaching as well as opportunities for prospective teachers to apply this knowledge in context of their future practice. Grossman et al. (2009) refer to these opportunities as "approximations of practice," or "opportunities to engage in practices that are more or less proximal to the practices of a profession" (p. 2058). And, recent evidence suggests that mathematical

[1] Throughout this volume, we refer to college students as *undergraduates* and use the term *school student* to refer to K–12 students. When commentary or context could apply to either group, we use the more general terms *students* or *learners*.

tasks that link university mathematics courses and school mathematics can counter the disconnect prospective teachers have felt between their university-level coursework and secondary mathematics (Rach, 2022). Yet, it is often the case that prospective secondary mathematics teachers enroll in traditional mathematics major courses that are taught by mathematicians who do not specialize in teacher preparation, creating a context in which this disconnect can persist.

Furthermore, *teaching applications* in materials used in undergraduate mathematics courses are lacking. Courses in the undergraduate mathematics major curriculum often have designated examples and problems addressing specific application areas such as physics, engineering, and business. When examples and problems are chosen from specific applications or careers, the result is a legitimization of the area as a focus of mathematical study and a subtle reminder of career opportunities undergraduates with appropriate mathematical preparation can pursue. When applications specifically addressing teaching are minimal or nonexistent (Lai & Patterson, 2017), mathematics teaching is undervalued as a rich mathematical area and as a career opportunity.

In response to these needs, the META Math project[2] has focused on adding *secondary mathematics teaching* explicitly to the list of application areas of undergraduate mathematics major courses by creating teaching materials to help instructors integrate national recommendations for teacher preparation. Those materials are the focus of this book, with one chapter for each lesson and a final chapter for summarizing the research about the use of these materials. The materials are formatted as lessons with teaching applications and instructor notes that provide detailed recommendations on implementation, accompanied by a set of ready-to-use handouts. The notes were written by members of the META Math project team, along with input from a variety of instructors who were early adopters of these lessons. The notes are intended to help instructors decide how to structure the class and incorporate class discussions by describing specific content and pedagogical emphases. The notes also highlight the ways in which these teaching applications help prepare prospective teachers for interactions with their future students' mathematical work and conceptions.

The design of lessons that feature teaching applications involved in-depth study and practice over the four years of this project. The lessons that we include in this volume are the result of that research (see Burroughs et al., 2023; Álvarez et al., 2020). There are nine lessons, aligning with courses in single variable calculus, introduction to statistics, discrete mathematics or introduction to proof, and abstract algebra (Table 1.1). The lessons have been implemented in a variety of settings nationwide, and we have researched their use in courses whose population of undergraduates consists of prospective secondary mathematics teachers, mathematics majors, and non-mathematics majors. Throughout the project, we have collected undergraduates' written work from the lessons and interviewed both instructors and a subset of their students to understand their experiences teaching or learning from these lessons. These data informed revisions of the lessons, which are what appear in this volume.

With these materials, the META Math project focuses specifically on enhancing prospective secondary teachers' understanding of connections between undergraduate mathematics and school mathematics. Our work complements and extends that of other researchers (e.g., Wasserman, 2018; Heid et al., 2015; Lai & Patterson, 2017; McCrory et al., 2012) who over the last decade have made progress in understanding the benefits of studying advanced mathematics as part of secondary teacher preparation. Just as mathematics applications to physics or chemistry are encountered by all undergraduates in a given course, we also intend for our teaching materials to increase awareness of teaching applications and these connections among all undergraduates in the course, even those not intending to pursue teaching as a career, in a way that deepens their understanding of undergraduate mathematics. We focus on teaching applications in an explicitly human context, so that the prospective teachers who engage with our materials will see that the human context of mathematics is held on par with the mathematics content (Álvarez et al., 2020).

This volume reflects our high regard for the discipline of mathematics and mathematicians and for teaching and teachers. Our lessons aim for a meaningful blend of undergraduate mathematics, school mathematics, and teaching that does not privilege mathematics content over content for teaching or vice versa. We view this volume as an invitation for

[2]The Mathematical Education of Teachers as an Application of Undergraduate Mathematics (META Math) was partially supported by the National Science Foundation (NSF) under DUE-1726624.

Course	Topic of Lesson
Single Variable Calculus	Inverse Functions and Their Derivatives Newton's Method
Introduction to Statistics	Variability: Mean Absolute Deviation and Standard Deviation Using Sampling Distributions to Build Understanding of Margin of Error
Discrete Mathematics or Introduction to Proof	The Binomial Theorem Foundations of Divisibility
Abstract Algebra	Solving Equations in Alternative Number Systems Groups of Transformations Logarithms and Isomorphisms

Table 1.1. META Math teaching materials, summarized by course and lesson topic

scholars to continue their collaborative efforts to improve mathematics teaching. As such, the teaching materials attend to mathematics content suitable for inclusion in courses for mathematics majors, to mathematics content that is foundational to teaching secondary mathematics, and to learning experiences fostered by the teaching materials.

1.2 Connections to Teaching

Many who have considered what prospective secondary teachers should understand about undergraduate mathematics arrive at a notion of "looking back" at the content of school mathematics; in fact, this perspective was advanced by Felix Klein (1932) more than a century ago and has resonated with many scholars since. Our work includes this perspective as one type of connection, and we have identified additional areas with which prospective teachers should gain experience in their study for the profession of teaching. We examined the particular needs of prospective secondary mathematics teachers and defined five types of connections between undergraduate-level mathematics and knowledge for teaching secondary mathematics (Arnold et al., 2020; Table 1.2). All five types of connections to teaching are embedded in the lesson materials in each chapter.

Connection to Teaching	Description
Content Knowledge	Undergraduates use course content in applied teaching contexts or to answer mathematical questions in the course.
Explaining Mathematical Content	Undergraduates justify mathematical procedures or theorems and use of related mathematical concepts.
Looking Back / Looking Forward	Undergraduates explain how mathematics topics are related over a span of K–12 curriculum through undergraduate mathematics.
School Student Thinking	Undergraduates evaluate the mathematics underlying a hypothetical student's work and explain what that student may understand.
Guiding School Students' Understanding	Undergraduates pose or evaluate guiding questions to help a hypothetical student understand a mathematical concept and explain how the questions may guide the student's learning.

Table 1.2. Five types of connections between undergraduate mathematics and teaching secondary mathematics.

What should prospective teachers understand about mathematics content? We propose that a prospective teacher who is well-prepared to teach school mathematics can: 1) solve mathematical problems foundational to teaching school mathematics; 2) explain the reasoning supporting the mathematical concepts; and 3) know what mathematics comes before and after the subject matter being studied. Notice that the three connections *Content Knowledge*, *Explaining Mathematical Content*, and *Looking Back / Looking Forward* are based in mathematical understanding, and yet they all rely on developing an understanding of how human beings interact with mathematical ideas. *Content Knowledge* and *Looking Back / Looking Forward* connections rely on developing an understanding of how school mathematics is organized; *Explaining Mathematical Content* connections rely on explaining mathematical ideas to another human being. The materials in the chapters that follow provide opportunities for undergraduates to encounter these three types of connections to teaching; you'll recognize these connections in activities that prompt undergraduates to explain mathematics concepts that explicitly or implicitly involve ideas from secondary mathematics.

What should prospective teachers understand about fostering mathematical understanding in others? We propose that they should be able to 4) look at student work and articulate what a student may or may not yet understand mathematically; and then 5) formulate meaningful questions aimed at guiding that student's understanding. Notice that in the connections of *School Student Thinking* and *Guiding School Students' Understanding*, human beings are explicitly included. You will recognize these two types of connections to teaching in the chapters that follow when you see characters—human beings—featured in the lesson materials.

We developed these five types of connections to teaching to guide the design of teaching applications, and we recognize that these five types of connections are intertwined. In what follows, we don't sort lesson activities into distinct categories by type of connection. Instead, the lesson materials are intended to provide undergraduates opportunities to encounter these ideas about mathematics and about teaching as they are threaded through the lessons and to provide instructors support in fostering these connections to teaching.

1.3 Principles of Teaching

In addition to featuring the five types of connections that situate undergraduate mathematics topics in the context of teaching secondary mathematics, our materials also incorporate three principles about teaching (Álvarez et al., 2020): *Habit of Respect*, *Active Engagement*, and *Recognition of Mathematics as a Human Activity*.

The *Habit of Respect* principle is our explicit acknowledgment that in teaching, it is instructors' responsibility to take an asset-based view of where students are in their development (rather than a deficit-based view of what students don't know). This principle appears both in how we approach our own teaching and in how we frame lesson materials that engage undergraduates with the human characters in the teaching applications. The materials include situations in which respect for student thinking and reasoning is valued, whether or not a learner's work is correct or incorrect. This principle manifests itself in the teaching applications in which undergraduates engage in addressing hypothetical students' developing notions in an affirming manner that centers and values their thinking.

The *Active Engagement* principle embraces classroom experiences that allow prospective teachers to make, explore, and validate conjectures and make sense of new ideas. Because student-centered approaches to teaching are prevalent in K–12 mathematics instruction, undergraduates who plan to become teachers should have opportunities to experience learning in a manner aligned with the ways that they may be expected to teach and with the ways in which many of them learned school mathematics. With this principle, we advocate for providing opportunities for undergraduates to become co-discoverers of mathematical concepts and for establishing norms for learning mathematics that include encouraging students to take ownership for creating mathematics. We enact this principle by facilitating interactive discussions and scaffolding ideas undergraduates encounter as they develop understanding of key concepts. We look for ways to encourage undergraduates to derive meanings underlying the methods, theorems, or ideas appropriate for a mathematical task. While enacting the principle of active engagement can take various forms, it relies upon and underscores the value of student discourse and social interaction in the learning of mathematics.

The *Recognition of Mathematics as a Human Activity* principle affirms that the practice of teaching involves interacting with learners as they engage with mathematical ideas. It positions as equal both knowledge about mathematics and knowledge about teaching and learning mathematics. Thus, it acknowledges the idea that flexible mathematics knowledge is critical when interacting with learners in a manner that affirms their thinking and scaffolds new ideas. Encountering hypothetical learners in a teaching application requires undergraduates to develop an underlying understanding of the mathematics in order to formulate appropriate ways to facilitate mathematical learning or guide student understanding.

1.4 Pedagogical Strategies Employed in Lessons

The chapters in this book reference a variety of pedagogical strategies that we have found effective when teaching these lessons. We intend for the materials to be flexible so that they can be useful to instructors with a wide variety of classroom styles, and the lessons were implemented by instructors with distinct styles. The strategies we reference most frequently are *group work and class discussions*, *selecting and sequencing student work*, and *exit tickets*.

Group Work and Class Discussions. To foster undergraduates' active engagement in the lessons in this volume, we rely on group work and class discussions. The MAA *Instructional Practices Guide* (2018, p. 8–15) gives some examples of effective group work strategies for the undergraduate mathematics classroom. Effective group work requires careful attention to creating a classroom environment that encourages student contributions as well as thinking about how those contributions will be leveraged by the instructor to facilitate learning. For example, instructors should set clear goals and expectations, establish norms for communication and participation, and provide scaffolding for students who may need extra support. Instructors can also encourage productive group dynamics by monitoring group interactions, providing feedback, and modeling good collaboration skills. Additionally, assigning group roles can encourage students to take ownership of their learning and promote accountability. For example, a group may have a leader who ensures everyone has a chance to contribute, a recorder who documents the group's progress, and a presenter who shares the group's findings with the class.

Selecting and Sequencing Student Work. When facilitating group work in mathematics, it is essential for instructors to monitor student interactions with an eye toward selecting and sequencing student work for powerful whole class discussions aimed at promoting student understanding. Smith and Stein's (2018) *5 Practices for Orchestrating Productive Mathematics Discussions* recommends selecting student work that represents a range of strategies and approaches. This helps create opportunities for students to learn from one another and build a deeper understanding of the mathematical concepts. Next, instructors make decisions about sequencing the discussion of the work in a strategic manner, starting with simpler and more accessible strategies or ideas and moving towards more complex ones. This helps ensure that all students have a chance to participate and build confidence before being challenged further. The active listening and interactions of the instructor are a critical component of facilitating active engagement in the lessons.

Exit Tickets. We have found it useful to ask undergraduates to complete an exit ticket at the end of each day of the lesson (see the MAA *Instructional Practices Guide* (2018, p. 7–8) for additional information about using exit tickets in instruction). Exit tickets are used to quickly assess what students learned during a lesson and what aspects of the lesson need to be reinforced in the following class session. We have framed our exit ticket prompt around the connections to teaching that undergraduates are making after working through a lesson. A prompt such as "Describe how you think today's lesson builds mathematical connections to teaching or learning middle school or high school mathematics" has allowed us to examine how undergraduates are building their understanding of the connections to teaching that are described throughout the lessons. When we see discrepancies between what we intend for undergraduates to understand and what they report about their understanding, we take some time at the start of the next class session to re-emphasize the ideas we want them to learn.

1.5 Using the META Math Materials

The lesson materials are formatted as a set of teaching notes—think of those as what you might prepare yourself—along with annotations—think of those as what might be provided by an experienced colleague who has implemented the lesson and has advice for you. (Hiebert and Morris (2012) describe *annotated lesson plans*, and much of our structure in providing these notes derives from our interpretation of their work.) These annotations contain additional information and commentary that is meant to enhance the set of teaching notes and provide additional support to you as you teach a lesson, particularly for the first time. The annotations will help you decide how to structure the class activity and incorporate class discussions, what to emphasize, and what undergraduates might know and still need to learn about different concepts. For example, an annotation indicating challenges undergraduates face during the lesson can help instructors anticipate these challenges in advance and give them an opportunity to prepare how they will address and respond to these challenges during instruction. Other annotations may include pedagogical recommendations regarding group work, for example, that worked well while implementing the lesson. In all of our materials, we included annotations that specifically address connections to teaching and highlight the ways in which the lessons help prepare prospective teachers.

Each META Math lesson contains the following: (1) Overview and Outline of Lesson, which is meant to provide a summary of what you will cover during the lesson; (2) Alignment with College Curriculum and (3) Links to School Mathematics, to situate the content of the lesson within your curriculum and within school mathematics; (4) Lesson Preparation, to indicate prerequisite knowledge, learning objectives, anticipated length, and materials needed; (5) Instructor Notes and Lesson Annotations, including solutions with sample responses, to help you prepare for and implement the lesson. *Solutions* provided in the lessons are author-written, while *sample responses* are based on those undergraduates have given when using these materials, though we have edited many of the original responses in order to improve clarity; (6) References; and (7) Lesson Handouts, including the pre-activities, class activities, homework problems, and assessment problems used in each lesson. These are also available as LaTeX downloads from maa.org/meta-math. The order in which we have presented these chapters does not imply an order in their use, although earlier chapters feature lessons that undergraduates usually encounter earlier in their coursework. Each lesson is independent of the others, and instructors who use multiple lessons for a single course can choose how to order them.

Instructors have found the Instructor Notes and Lesson Annotations to be invaluable to their implementation of the META Math lessons, because of the extra details and the clear indications of when to "Discuss This Connection to Teaching" during the lesson (see Álvarez et al., 2022; Fulton et al., 2022). These types of annotations are intended to remind you of the mathematical ideas to focus on during the lesson and to provide a straightforward way to connect ideas in your undergraduate-level mathematics course to teaching school mathematics.

We encourage you to take a reflective stance while teaching with these materials and to consider the instructor notes and annotations we have included. Consider how you would answer these questions of your own teaching:

- What is "the mathematics needed for teaching secondary mathematics"?
- How do I implement active engagement and reflection in my course, and how does this influence secondary teacher preparation?

The materials in this volume are what resulted when we reflected on these questions for ourselves. We hope that the opportunities to engage in robust teacher preparation that are embedded in the META Math lessons resonate with you. We have wholeheartedly embraced the notion that including mathematics applications for teaching in undergraduate mathematics major courses is valuable for all undergraduates, those who intend to teach mathematics in the secondary grades and those who don't, and we are pleased to share that perspective with you through these materials.

1.6 References

[1] Álvarez, J. A. M, Arnold, E. G., Burroughs, E. A., Fulton, E. W., & Kercher, A. (2020). The design of tasks that address applications to teaching secondary mathematics for use in undergraduate mathematics courses. *The Journal of Mathematical Behavior*, *60*, 100814.

[2] Álvarez, J. A., Kercher, A., Turner, K., Arnold, E. G., Burroughs, E. A., & Fulton, E. W. (2022). Including school mathematics teaching applications in an undergraduate abstract algebra course. *PRIMUS*. *32*(6), 685–703, `doi.org/10.1080/10511970.2021.1912230`.

[3] Arnold, E. G., Burroughs, E. A., Fulton, E. W., & Álvarez, J. A. M. (2020). Applications of teaching secondary mathematics in under- graduate mathematics courses. *TSG33 of the 14th International Congress on Mathematical Education. Shanghai, China: International Mathematics Union.* arXiv:2102.04537.

[4] Association of Mathematics Teacher Educators (AMTE). (2017). *Standards for preparing teachers of mathematics*. Author. Available online at amte.net/standards

[5] Burroughs, E. A., Arnold, E. G., Álvarez, J. A. M., Kercher, A., Tremaine, R., Fulton, E., & Turner, K. (2023). Encountering ideas about teaching and learning mathematics in undergraduate mathematics courses. *ZDM–Mathematics Education*. `doi.org/10.1007/s11858-022-01454-3`.

[6] Conference Board of the Mathematical Sciences (CBMS). (2001). *The mathematical education of teachers* (Vol. 11). American Mathematical Society and Mathematical Association of America.

[7] Conference Board of the Mathematical Sciences (CBMS). (2012). *The mathematical education of teachers II* (Vol. 17). American Mathematical Society and Mathematical Association of America.

[8] Fulton, E. W., Arnold, E. G., Burroughs, E. A., Álvarez, J. A., Kercher, A., & Turner, K. (2022). Including school mathematics teaching applications in an undergraduate discrete mathematics course. *PRIMUS*. *32*(6), 704–720, `doi.org/10.1080/10511970.2021.1905120`.

[9] Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055–2100.

[10] Heid, M. K., Wilson, P. S., & Blume, G. W. (2015). *Mathematical understanding for secondary teaching: A framework and classroom-based situations.* IAP.

[11] Hiebert, J., & Morris, A. K. (2012). Teaching, rather than teachers, as a path toward improving classroom instruction. *Journal of Teacher Education*, *63*(2), 92–102.

[12] Klein, F. (1925). *Elementary mathematics from an advanced standpoint: Arithmetic, algebra, analysis.* Translated from the third German edition by E.R. Hedrick and C. A. Noble. Dover.

[13] Lai, Y., & Patterson, C. (2017). Opportunities presented by mathematics textbooks for prospective teachers to learn to use mathematics in teaching. *Proceedings of the sixth annual mathematics teacher education partnership conference*.

[14] Mathematical Association of America. (2018). *Instructional practices guide*. (Especially Chapter 1, “Classroom Practices”). `https://www.maa.org/programs-and-communities/curriculum%20resources/instructional-practices-guide`

[15] McCrory, R., Floden, R., Ferrini-Mundy, J., Reckase, M. D., & Senk, S. L. (2012). Knowledge of algebra for teaching: A framework of knowledge and practices. *Journal for Research in Mathematics Education*, *43*(5), 584–615.

[16] Moreira, P. C., & David, M. M. (2008). Academic mathematics and mathematical knowledge needed in school teaching practice: Some conflicting elements. *Journal of Mathematics Teacher Education*, *11*, 23–40.

[17] Rach, S. (2022). Tasks for linking school and university mathematics: do such tasks have an impact on the interest of student teachers? In *Professional Science* (pp. 177–189). Springer Spectrum.

[18] Smith, M. S., & Stein, M. K. (2018). *5 practices for orchestrating productive mathematics discussions*. National Council of Teachers of Mathematics; Corwin.

[19] Speer, N. M., King, K. D., & Howell, H. (2015). Definitions of mathematical knowledge for teaching: Using these constructs in research on secondary and college mathematics teachers. *Journal of Mathematics Teacher Education*, *18*, 105–122.

[20] Stylianides, A., & Stylianides, G. (2014). Viewing “mathematics for teaching” as a form of applied mathematics: Implications for the mathematical preparation of teachers. *Notices of the AMS*, *61*(3), 266–276.

[21] Wasserman, N. H. (2018). Exploring advanced mathematics courses and content for secondary mathematics teachers. In N. H. Wasserman (Ed.), *Connecting abstract algebra to secondary mathematics, for secondary mathematics teachers* (pp. 1–15). Springer, Cham.

[22] Winsor, M. S., Barker, D. D., & Kirwan, J. V. (2020). Promoting knowledge integration in teacher education programs. In T. Lehmann (Ed.), *International perspectives on knowledge integration: Theory, research, and good practice in pre-service teacher and higher education* (pp. 349–369). Brill Sense.

[23] Zazkis, R., & Leikin, R. (2010). Advanced mathematical knowledge in teaching practice: Perceptions of secondary mathematics teachers. *Mathematical Thinking and Learning*, *12*(4), 263–281.

2

Inverse Functions and Their Derivatives

Single Variable Calculus

James A. M. Álvarez, *The University of Texas at Arlington*
Elizabeth G. Arnold, *Colorado State University*
Andrew Kercher, *Simon Fraser University*
Elizabeth W. Fulton, *Montana State University*
Elizabeth A. Burroughs, *Montana State University*
Kyle Turner, *The University of Texas at Arlington*

2.1 Overview and Outline of Lesson

High school students and undergraduates encounter inverse functions in various forms in their mathematics courses. Many students memorize and apply the following technique to find the inverse function of a given function: (1) Write $y = f(x)$; (2) Switch the x and y; and (3) Solve for y. This procedure obscures important relationships and concepts that may help bridge student understandings when these topics are revisited in a more advanced setting. The purpose of this lesson is to further develop undergraduates' conceptual understanding of the relationship between a function and its inverse function and apply this understanding to find derivatives of inverse functions, such as using the derivative of $\tan(x)$ to find the derivative of $\arctan(x)$.

1. Launch—Pre-Activity
   Prior to the lesson, undergraduates complete the Pre-Activity, which focuses on having students think about the relationship between a function and its inverse function. By examining graphs of familiar functions, undergraduates identify when a particular function has an inverse function and what the relationship might be between the domain and range for each function and its inverse function. Instructors can launch the lesson by discussing undergraduates' responses to the Pre-Activity.

2. Explore—Class Activity

   - *Problems 1–2*:
     In Problem 1, undergraduates will compare and contrast three students' methods for finding the inverse of a function. One is the commonly taught method from high school of "switch x and y and solve for y," while the others rely on the mathematical principles underlying the relationship between a function and its inverse function. In Problem 2, undergraduates investigate the limitations of the method "switch x and y and solve for y" in the context of two different situations. They consider whether the other two methods from Problem 1 give rise to the same difficulties.
   - *Problems 3–6*:
     In Problems 3 and 4, the properties of inverse functions and the chain rule are used to compute the derivatives

of $\arcsin(x)$ and $\ln(x)$ in terms of x. Undergraduates reflect on the ways in which the properties of inverse functions were used in these computations to produce the desired form of the derivatives. In Problems 5 and 6, undergraduates write a method for finding the derivative of any inverse function when given a function and its inverse function. They then use this method to write an explicit formula for this derivative and explain why it makes sense.

3. Closure—Wrap-Up
   Conclude the lesson by discussing how the properties of inverse functions and the chain rule play an important role in understanding how a function f, its derivative f', and its inverse function f^{-1} can be used to find a formula for the derivative of f^{-1}. Make connections to secondary mathematics by emphasizing how focusing on the properties of inverse functions, in particular the composition of a function and its inverse function (and vice versa), lays the groundwork for establishing the formulas that use these relationships, such as $\frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1-x^2}}$ for $-1 < x < 1$.

2.2 Alignment with College Curriculum

Inverse functions are encountered in single variable calculus courses in the following ways: (1) defining an inverse function using concepts introduced in precalculus and (2) using inverse functions and inverse relationships to derive the derivatives of inverse functions, in particular finding the derivatives of inverse trigonometric functions. This lesson addresses both of these by (1) reviewing methods to find inverse functions with a focus on the domain and range of the inverse function and the fact that the composition of a function and its inverse function is the identity function; and (2) finding derivatives of inverse functions with a focus on using the chain rule on the composition of a function and its inverse.

2.3 Links to School Mathematics

Finding inverse functions is addressed in high school algebra and precalculus courses, as indicated in the Common Core State Standards for Mathematics (CCSSM, 2010). Frequently, inverse functions are taught as an algebraic manipulation that students memorize, which does not provide natural places to develop automaticity in using composition of functions. This lesson problematizes the common method of "switch x and y and solve for y" to find the inverse of a function by providing sample student work for undergraduates to analyze and discuss.

This lesson highlights:

- Connections between calculus concepts and ideas about functions, including how inverse functions can be leveraged to find certain derivatives;
- The importance of unambiguous mathematical notation when deriving the inverse of a function.

This lesson addresses several mathematical knowledge and practice expectations in common high school standards documents (e.g., CCSSM). To work with functions, high school students are expected to understand the concept of function and use function notation. In particular, most state standards for high school mathematics include detailed expectations related to the definition of function (c.f. CCSS.MATH.CONTENT.HSF.IF.A.1). Additionally, high school students are expected to build functions that model relationships between two quantities. To build appropriate functions, they also learn about function composition and how to find the inverse of a function (c.f. CCSS.MATH.CONTENT.BF.A.1.C, CCSS.MATH.CONTENT.HSF.BF.B.4). This lesson also provides opportunities for prospective teachers to think about the role of the context of a problem, the reasoning of others, and construct sound mathematical arguments.

2.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know how to:

- Find the inverse function of a given function algebraically and graphically;
- Use the chain rule.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Describe the mathematical principles underlying different methods used to compute the inverse of a function;
- Describe the limitations of the method of "switch x and y and solve for y" to find the inverse of a function;
- Compute derivatives of inverse functions using compositions and the chain rule;
- Analyze hypothetical student work and evaluate three different methods used to find the inverse of a function;
- Pose guiding questions to help a hypothetical student understand when certain methods for finding the inverse of a function have mathematical limitations.

Anticipated Length

One 75- to 80-minute class session.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity
- Computer (for instructor to display a dynamic sketch during the activity)
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

2.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to be completed in preparation for this lesson. We recommend that you collect this Pre-Activity the day before the lesson so that you can review undergraduates' responses before you begin the Class Activity. This will help you determine if you need to spend additional time reviewing the solutions to the Pre-Activity with your undergraduates.

Pre-Activity Review (10 minutes)

Discuss undergraduates' responses to the Pre-Activity, and introduce the lesson by discussing the following connection to teaching. If you wish, provide specific examples of situations in which the existence of an inverse function, in particular, is important.

Discuss This Connection to Teaching

Functions are an important component of school mathematics. A robust understanding of functions provides students with foundational tools for success in calculus and other courses that rely on quantitative thinking and determining relationships between quantities. Despite the importance of building a strong conceptual understanding of functions in grades 8–12, teachers encounter many challenges teaching functions. For example, students "tend to limit the concept of functions to equations or orderly rules" and "frequently overlook many-to-one correspondences or irregular functions that could be very useful in describing or representing real-world phenomena" (Cooney et al., 2010, p. 1). These challenges make it essential for prospective teachers to cultivate their own powerful understanding of functions.

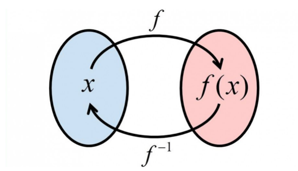
Pre-Activity Problem 1

1. Given a function f, its ***inverse function*** (if it exists) is the function f^{-1} such that $y = f(x)$ if and only if $f^{-1}(y) = x$.

   (a) If we know that $f(♫) = ❤$, what is $f^{-1}(❤)$? What about $f(f^{-1}(❤))$?

   Solution:
   $f^{-1}(❤) = ♫$
   $f(f^{-1}(❤)) = f(♫) = ❤$

   Commentary:
   When discussing 1(a), consider using a function diagram (as seen below) rather than a graph to represent a function and its inverse. Emphasize that the inverse function "undoes" the original function. This behavior is exemplified by their composition.



   (b) If we know a function has an inverse function, what do we know about the properties, behavior, or graph of its inverse function? Create a list of these attributes.

   Sample Responses:

   - It's what you get when you switch x and y and then solve for y. [most common response]
   - All the x and y coordinates are flipped $\rightarrow$ opposite function.
   - Reflects about $y = x$.
   - The inverse "undoes" f [the original function].
   - Opposite solving steps.
   - It's a function that "reverses" another function.
   - Domain and range flipped.
   - It's whatever function returns the input variable when composed with f [the original function].
   - Function notation: $f^{-1}(x)$.
   - Logarithm and exponential growth graphs are inverses of each other.
   - It's the opposite of f [the original function].
   - It's the reciprocal of f [the original function].

   Commentary:
   For 1(b), ask undergraduates to share their answers. Some common undergraduate responses that are correct, incorrect, or incomplete are provided above. Record their responses on the board or document camera. If the following do not arise in their responses, be sure to add them to the list of attributes since they will be used in discussions that will occur in the rest of the lesson:

   - The graph of the inverse function appears as the reflection over the line $y = x$ of the graph of the original function.
   - An inverse function *"undoes"* the mapping or process induced by its associated function and vice-versa. That is, the composition of a function and its inverse function is the identity, no matter the order of the composition.

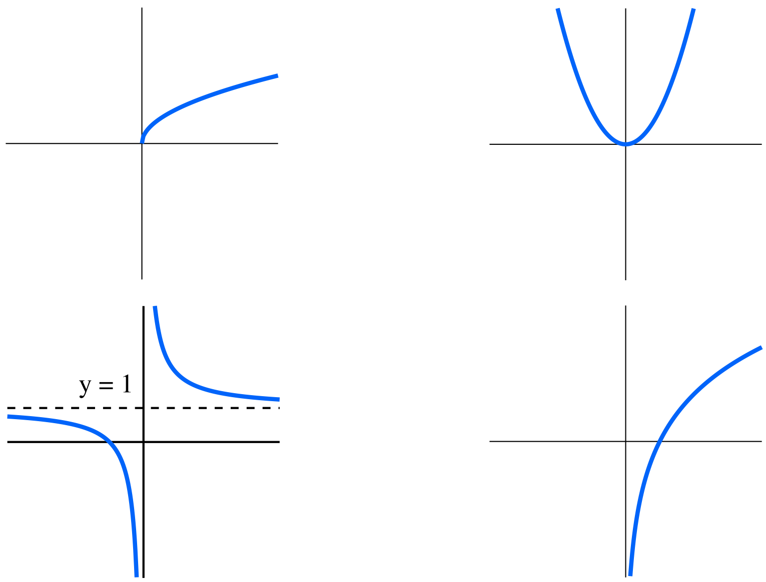
After discussing Problem 1, make sure undergraduates are using correct notation to indicate the inverse of a function and discuss the following connection to teaching, which highlights a common misapplication of the notation used to indicate an inverse function.

Discuss This Connection to Teaching

The fact that $f^{-1}(x) \neq \frac{1}{f(x)}$ runs counter to students' experiences with inverses of real numbers. As a result, students might use their experience with the superscript (-1) in the context of real numbers to incorrectly conclude, for example, that $\sin^{-1}(x) = \frac{1}{\sin(x)}$. Prospective teachers should be able to compare the use of negative exponents when writing 3^{-1} versus $f^{-1}(x)$ and explain why mathematicians have chosen this notation, which in different contexts, yields seemingly "different" actions. See Zazkis & Kontorovich (2016) for further insight.

Pre-Activity Problem 2

2. Consider each of the functions below.



(a) Which of the these functions has an inverse function? Explain how you can know without being given the defining expression.

Solution:
Only the "parabola-looking" function does not have an inverse function. Undergraduate explanations may vary, and include:

- The graph that looks like a parabola is not one-to-one.
- The graph that looks like a parabola does not pass the "horizontal line test."
- If you reflect the graph of the parabola over the line $y = x$, the resulting graph does not represent a function because it fails the vertical line test.

(b) What is the domain and range of each function? What is the domain and range of its inverse function (if it exists)?

Solutions:
Function graph that resembles a square root function:

- $D[f] = (0, \infty)$, $R[f] = (0, \infty)$
- $D[f^{-1}] = (0, \infty)$, $R[f^{-1}] = (0, \infty)$

Function graph that resembles a quadratic function:

- $D[f] = [0, \infty)$, $R[f] = (-\infty, \infty)$
- Does not have an inverse function.

Function graph that resembles a rational function with vertical asymptote at $x = 0$ and horizontal asymptote at $y = 1$:

- $D[f] = (-\infty, 0) \cup (0, \infty)$, $R[f] = (-\infty, 1) \cup (1, \infty)$
- $D[f^{-1}] = (-\infty, 1) \cup (1, \infty)$, $R[f^{-1}] = (-\infty, 0) \cup (0, \infty)$

Function graph that resembles a logarithmic function:

- $D[f] = (0, \infty)$, $R[f] = (-\infty, \infty)$
- $D[f^{-1}] = (-\infty, \infty)$, $R[f^{-1}] = (0, \infty)$

Commentary:

After discussing 2(b), ask the class to compare the domain and range of each function with the domain and range of its inverse function. Do they notice a relationship between the domain and range of a function and the domain and range of its inverse function? Highlight the following relationship and make sure it is added to the list of attributes about functions and their inverse functions on Problem 1b of the Pre-Activity (if it was not there already):

- The domain of f is the range of f^{-1} and the range of f is the domain of f^{-1}.

Wrap up the Pre-Activity by letting undergraduates know that during the Class Activity they will explore how the properties of functions and their inverse functions can be used to determine the derivative of an inverse function by using only the derivative of the original function.

Class Activity: Problems 1–2 (30 minutes)

Distribute the Class Activity and discuss the following connection to teaching.

Discuss This Connection to Teaching

Problem 1 gives undergraduates opportunities to analyze other students' thinking in order to develop their skills in understanding school student thinking. All undergraduates (especially prospective teachers) should examine how others use, reason with, and communicate mathematics. In addition, Problems 1 and 2 of the Class Activity lay the groundwork for problematizing the method of "switch x and y and solve for y" [without attending to the corresponding swap of domain and range] for finding an inverse of a function represented by $y = f(x)$. Raising awareness regarding how this method may not advance understanding or connect to conceptual ideas in calculus for prospective teachers, in particular, may influence their own teaching of inverse functions in the future.

Give undergraduates a few minutes to study Problem 1 before directing them to work in groups on this problem. (See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion.) In this problem, undergraduates will examine the work of three hypothetical students, Alex, Jordan, and Kelly, who used different methods to find the inverse of a function. If undergraduates focus on whether the algebra in each student's work is correct, let them know that the computations are correct and that they should focus on the methods each student is using to find the inverse function.

Alex's method in Problem 1 is not always appropriate, and even when it is appropriate mathematically, its use may obscure the meaning behind the variables and can lead to confusion in contextualized problems. Undergraduates will see specific examples of this shortcoming in Problem 2. Jordan's and Kelly's method makes use of the definition of inverse functions. By recognizing that the inverse function of $y = f(x)$ is $x = f^{-1}(y)$, learners can make sense of inverse functions in multiple mathematical contexts including real world data analysis and modeling (Wilson et al., 2016). By discussing the relationship between different input values and output values of a function and its inverse function, undergraduates can focus on the concept that inverse functions undo the mapping produced by the original function. In addition, analyzing three different (although incomplete) sample student approaches to finding the inverse function puts undergraduates in a teaching situation by having them analyze *student work. This analysis may sharpen conceptual understanding of inverse functions (e.g., understanding why and how different methods work and how they are grounded in mathematical operations) as well as develop skills for examining student work for mathematical understanding (Cooney et al., 2010).*

Class Activity Problem 1

1. Consider how Alex, Jordan, and Kelly found the inverse function of $f(x) = \frac{2}{3}x + 1$.

Alex's Work	Jordan's Work	Kelly's Work
$y = 2/3\,x + 1$ $x = 2/3\,y + 1$ $x - 1 = 2/3\,y$ $\frac{x-1}{2/3} = y$ $\boxed{\frac{3}{2}(x-1) = y}$	$f \circ f^{-1}(y) = y$ $f(x) = \frac{2}{3}x + 1$ So: $f(f^{-1}(y)) = \frac{2}{3}f^{-1}(y) + 1 = y$ $\frac{2}{3}f^{-1}(y) = y - 1$ $\frac{3}{2} \cdot \frac{2}{3} f^{-1}(y) = \frac{3}{2}(y-1)$ $\boxed{f^{-1}(y) = \frac{3}{2}(y-1)}$	$y = \frac{2x}{3} + 1$ $y - 1 = \frac{2x}{3}$ $3(y-1) = 2x$ $\frac{3(y-1)}{2} = x$ But $f(x) = y \Rightarrow$ $f^{-1}(f(x)) = f^{-1}(y) \Rightarrow$ $x = f^{-1}(y)$ So $\boxed{f^{-1}(y) = \frac{3(y-1)}{2}}$

Compare and contrast the key mathematical ideas used by Alex, Jordan, and Kelly to find the inverse function of $f(x) = \frac{2}{3}x + 1$. Make sure to identify which properties of inverse functions each student uses, if any.

Sample Responses:

- Alex switches x and y and solves for y.
- Jordan uses the fact that the composition of a function and its inverse function is the identity to write an equation involving $f^{-1}(y)$, then solves for that as if it were a variable.
- Kelly solves for x, then uses that fact that $f^{-1}(f(x)) = f^{-1}(y) = x$ to make a substitution.
- Both Jordan and Kelly end up with a function of y, while Alex's method produces a function of x.
- Both Jordan and Kelly use the definition of an inverse function.

Commentary:
From our experience, undergraduates have expressed their familiarity with using Alex's method (i.e., "switch x and y and solve for y"), and some have argued that Jordan or Kelly are "wrong" in some way. An important feature of this problem is to have undergraduates engage in providing a mathematical explanation for *why* they feel this way. This affords them an opportunity to enhance their capacity to make sense of and justify unfamiliar methods. If undergraduates are unfamiliar with Jordan's and Kelly's method, the following questions may help them to make sense of Jordan's and Kelly's work.

- How is Jordan/Kelly starting the problem?
- Why do you think Jordan/Kelly is using $f^{-1}(y)$? What property of inverse functions does this highlight?
- Which methods make use of the definition of the inverse function that we saw in the Pre-Activity?
- What understanding of inverse functions does Jordan/Kelly have that Alex may not?
- Why does Jordan's/Kelly's/Alex's method work?
- Why do you think Jordan's/Kelly's/Alex's method does not work?

After undergraduates have had some time to work through Problem 1, gather the class back together for a discussion. First ask undergraduates to report out their responses. As you facilitate the discussion, call attention to explanations that involve the properties of inverse functions that they noted while working to answer the problem. Consider compiling a master list of these "mathematical ideas" as undergraduates volunteer them. Undergraduates will refer back to these ideas during Problem 4 of the Class Activity and being able to quickly remember the highlights of the discussion will help streamline the process.

Consider asking the following questions, which can prompt discussion about the applicability of each method:

- Which method would you most likely use? Why?
- Which method makes the most sense to you? Why?
- Is Alex's solution "the same" as the other two? (i.e., Does this difference "matter"?) Why/Why not?

Instruct undergraduates to work in groups on Problems 2(a) and 2(b) and emphasize that the computations in the students' work are correct which redirects students' attention to the conceptual ideas instead of their focusing on the computations. From our experience, we have found that undergraduates did not immediately know the best way to answer these questions. Allow sufficient time to work in groups so that undergraduates have time to recoup if they initially misinterpreted the problem.

Class Activity Problem 2 : Parts a & b

2. Now consider two problems where a high school student used Alex's method of switching the variables and solving for the dependent variable to find the inverse function.

Find the inverse function of $T(C) = \frac{9}{5}C + 32$ where C is the temperature in Celsius and $F = T(C)$ gives the temperature in Fahrenheit.	Find the inverse function of $f(x) = \frac{2x+1}{x-1}$ for $x \neq 1$.
Let $F = \frac{9}{5}C + 32$ $C = \frac{9}{5}F + 32$ $C - 32 = \frac{9}{5}F$ $\frac{5}{9}(C-32) = F$	$y = \frac{2x+1}{x-1}$, $x \neq 1$ $x = \frac{2y+1}{y-1}$ $(y-1)x = 2y+1$ $yx - x = 2y+1$ $yx - 2y = 1+x$ $y(x-2) = x+1$ $y = \frac{x+1}{x-2}$, $x \neq 1$

(a) Describe why the student's work for the temperature function is problematic.

Sample Responses:

- No inverse notation is used.
- We should not switch F and C because they represent different units.
- We wanted a function where we plug in Fahrenheit and get the temperature in Celsius; this would "undo" the original function.
- We appear to have two different functions for changing Celsius to Fahrenheit which behave differently.

Commentary:

Below are some questions that you might ask undergraduates to facilitate their group conversations:

- What does the inverse function mean in this context?
- Is the property that an inverse function "undoes" the original function preserved? How can we check?
- What are the units attached to F and C? Are they preserved when we switch F and C?
- How might we fix this work so that these issues are resolved?

(b) Describe why the student's work for the rational function is problematic.

Sample Responses:

- $x \neq 1$ is an appropriate restriction on the first function, but does not make sense for its inverse function.
- The domain and the range of the function and its inverse were not "switched"—the last line has an unnecessary restriction on x, but without the necessary restriction of $x \neq 2$.

Commentary:

Below are some questions that you might ask undergraduates to facilitate their group conversations on Problem 2(b):

- Where are the horizontal and vertical asymptotes of the original function? Of the inverse function?
- Look back at the rational looking function from Problem 2 of the Pre-Activity. How did its domain and range relate to the domain and range of its inverse function?
- How might we fix this work so that these issues are resolved?

Before moving on to Problem 2(c), call the class back together and allow groups to share and discuss their answers to Problems 2(a) and 2(b).

Class Activity Problem 2 : Part c

(c) What are the limitations of using Alex's method of switching the variables and solving for the dependent variable to find an inverse function? Would Jordan and Kelly have the same problem(s)? Explain.

Sample Responses:

- Variables might have an associated unit which we ignore when we switch them.
- Switching the variables is misleading if we also do not account for the change in domain/range when considering what x and y now represent in the context of the inverse function.
- Jordan and Kelly would not have the same problem because their notation would preserve the identity of the variables and would switch the domain and range.
- When using Jordan's method, we use the fact that the composition of a function and its inverse function is the identity, which implicitly requires that the range of the inner function be the domain of the outer function. Jordan's method also does not switch the units attached to each variable.
- When using Kelly's method, the substitution $x = f^{-1}(y)$ reminds us that the inverse function "undoes" the original function by using the range of f as its domain. Kelly's method also does not mix up the units which may be attached to each variable.

Commentary:

Note that there is some overlap between Problem 2(c) and Problems 2(a) and 2(b). The following prompts can be used to facilitate a class discussion.

- "Is the y at the end the same as the y at the beginning? Why or why not?"
- For many functions, the meaning of the variables are important. Switching them confuses their intended meaning.
- Solving by switching x and y hides these properties of inverse functions:
  - The inverse function *undoes* whatever the function does. If the original function maps x values to y values, the inverse function should send y values back to x values.
  - The domain of the function is the range of its inverse function and vice versa.
- Jordan and Kelly avoid these problems by indicating that the inverse function is a function of y, not x.

After discussing Problem 2(c), emphasize to your class that Alex's method is not, strictly speaking, incorrect; but that it does come with certain drawbacks. Discuss the following connection to teaching to highlight this point.

Discuss This Connection to Teaching

Finding the inverse of a function is taught in second year school algebra (i.e., Algebra II in most states in the United States), and it is often taught with a procedural emphasis. Students memorize a procedure (e.g., "switch x and y and solve for y") and apply it to find the inverse of a given function. Although this procedure "works" when students are working with linear functions, students tend to struggle when working with other functions, such as transcendental functions (Teuscher et al., 2018). In addition, procedural approaches often are not grounded in the mathematical operations associated with the relationship between a function and its inverse function, may hinder a learner's understanding of the derivatives of inverse functions as they relate to the original function, and may inadvertently limit students' contextual practice in composition of functions (Wilson et al., 2016).

Class Activity: Problems 3–6 (30 minutes)

Introduce the next part of the Class Activity by explaining to undergraduates that they will use the previous explorations on inverse functions in order to learn to find derivatives of inverse functions. Explain that to find such a derivative, they will need to already know the derivative of the original function and will use function compositions and the relationship between a function and its inverse function.

First, tell your class to only work on Problem 3(a) in groups. Before undergraduates begin, ask them what they know about the inverse function of $y = \arcsin(x)$.

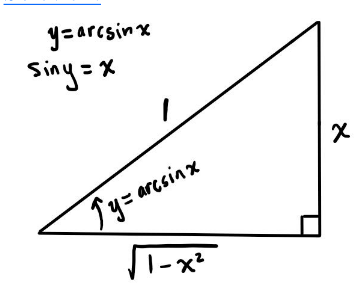
- Ask undergraduates what x and y mean in the context of the problem. You may need to remind them that $\arcsin(x) = \sin^{-1}(x)$ and that $\arcsin(x)$ is the angle whose sine value is x.
- It may be helpful to refer back to the definition of an inverse function to prompt undergraduates to conclude that $\sin(y) = x$.

Class Activity Problem 3 : Part a

3. We can use the properties of inverse functions to find their derivatives with respect to x.

   (a) Draw a right triangle that illustrates the relationship $y = \arcsin(x)$ for $0 < x < 1$.

   Solution:



   Commentary:
   From our experience, an undergraduate who uses Jordan's method will likely feel more comfortable working with $\sin(\arcsin(x)) = x$ or other mathematical statements involving the composition of functions. Jordan's method also requires the manipulation of function notation as a mathematical object, a skill which many undergraduates struggle to develop.

After most groups have a reasonable sketch of the triangle (such as the one presented in the solution above), share an undergraduate's correct drawing and discuss the following connection to teaching. You may wish to emphasize that this triangle model only considers values of y within the interval $(0, \frac{\pi}{2})$, but can be extended to account for $y \in (-\frac{\pi}{2}, 0]$. Alternatively, consider asking your class to consider why we might need to make this restriction.

Discuss This Connection to Teaching

High school students typically create these types of right triangles to illustrate how the values of trigonometric functions of an acute angle are ratios of the lengths of the sides of a right triangle. These images can be used to help students compute all six trigonometric functions of a given angle and examine other characteristics of trigonometric functions.

Now allow your class to complete Problem 3(b) by finding the derivative of $y = \arcsin(x)$ in terms of x. You may want to hint that they should take the derivative of both sides of the given equation. Remind them to use the picture to make substitutions when necessary to write their final answer as a function of x.

Class Activity Problem 3 : Part b

(b) Use the fact that $\sin(\arcsin(x)) = x$ and the chain rule to compute $\frac{d}{dx}\arcsin(x)$ in terms of x.

Solution:

$$\begin{aligned} \sin(\arcsin(x)) &= x \\ \tfrac{d}{dx}\sin(\arcsin(x)) &= \tfrac{d}{dx}x \\ \cos(\arcsin(x)) \cdot \tfrac{d}{dx}\arcsin(x) &= 1 \\ \tfrac{d}{dx}\arcsin(x) &= \tfrac{1}{\cos(\arcsin(x))} \\ \tfrac{d}{dx}\arcsin(x) &= \tfrac{1}{\cos(y)} \\ \tfrac{d}{dx}\arcsin(x) &= \tfrac{1}{\sqrt{1-x^2}} \end{aligned}$$

Commentary:
The following prompts can be used to facilitate a class discussion.

- How does the chain rule come into play in this problem?
- How did you use your picture while you worked? In what ways was your picture helpful as you attempt to write the derivative of $\arcsin(x)$ in terms of x (that is, "by eliminating or not involving a trigonometric function")?

Instruct undergraduates to work on Problem 3(c). Encourage them to use their work on Problem 3(b) as a guide, telling them that the way the chain rule is involved in each part is largely the same.

Class Activity Problem 3 : Part c

(c) Use the fact that $e^{\ln(x)} = x$ and the chain rule to show that $\frac{d}{dx}\ln(x) = \frac{1}{x}$.

Solution:

$$\begin{aligned} \tfrac{d}{dx}e^{\ln(x)} &= \tfrac{d}{dx}x \\ e^{\ln(x)} \cdot \tfrac{d}{dx}\ln(x) &= 1 \\ \tfrac{d}{dx}\ln(x) &= \tfrac{1}{e^{\ln(x)}} \\ \tfrac{d}{dx}\ln(x) &= \tfrac{1}{x} \end{aligned}$$

Commentary:
You may or may not want to re-engage the whole class for a discussion depending on the proficiency they demonstrate as they solve Problem 3(c) in their smaller groups. Some discussion questions you might use include:

- The composition of two functions in Problem 3(c) is slightly less obvious. Which is the interior function and which is the exterior?
- For what values of x is this formula for the derivative valid? Why?

Transition to the final three problems in the Class Activity by appealing to the need for a generalized formula; that is, while it is nice to be able to find the derivatives for particular functions, as in Problem 3, it would certainly be useful if we could find a "nice" relationship between the derivative of any function to the derivative of its inverse.

Class Activity Problem 4

4. What key mathematical idea(s) from Problem 1 did we use to find the derivatives in Problem 3?

Solution:
We used the idea from Jordan and Kelly that a function composed with its inverse produces the identity.

Commentary:
If you chose to write a master list of "mathematical ideas" on the board during Problem 1, you might choose not to break out into smaller groups before the class discussion on this problem. No matter how the discussion begins, make sure that undergraduates recognize that in both derivations from Problem 3, they had to begin with the idea that the composition of a function with its inverse function is the identity. This allowed them to apply chain rule and solve for the required derivative.

Instruct undergraduates to work on Problem 5(a) and 5(b) and verify that they are using chain rule correctly. In Problem 5(a), undergraduates need to write a few sentences *describing* the process they would use to compute $\frac{d}{dx}g(x)$ for any function f with inverse function g, and in Problem 5(b), they will apply that process.

Class Activity Problem 5 : Parts a & b

5. Let f be a function with inverse function g.

   (a) In the form of written sentences, describe how you would use the fact that $f(g(x)) = x$ and the chain rule to compute $\frac{d}{dx}g(x)$ for any function f with inverse function g.

   Sample Response:
   You would take the derivative of both sides of $f(g(x)) = x$ with respect to x, using the chain rule on the left hand side. The derivative of both sides of the equation, $f(g(x)) = x$ is $f'(g(x)) \cdot g'(x) = 1$. We are then able to divide both sides of the equation by $f'(g(x))$, provided it is not equal to zero, in order to arrive at an expression for the derivative of a general inverse function.

   (b) Use the procedure you wrote above to compute $\frac{d}{dx}g(x)$ for any function f with inverse function g.

   Solution:

$$\begin{aligned} f(g(x)) &= x \\ \frac{d}{dx}[f(g(x))] &= \frac{d}{dx}x \\ f'(g(x)) \cdot g'(x) &= 1 \\ g'(x) &= \frac{1}{f'(g(x))} \text{ where } f'(g(x)) \neq 0 \end{aligned}$$

Ask undergraduates to complete Problem 5(c), and have this dynamic sketch (`https://www.desmos.com/calculator/6fegvrt5xv`) in a place where it is visible to the whole class in order to help stimulate their discussion. When groups begin working on Problem 5(c), listen to the group discussions and find an appropriate time to bring the whole class together to demonstrate the features of the dynamic sketch. At this point, discuss the connection to teaching below.

Discuss This Connection to Teaching

Using multiple representations of functions (in this case, an algebraic representation and a graph) may help students connect mathematical ideas and strengthen understanding. Emphasizing qualitative ways to make sense of algebraic and graphical representations provides students with more experiences making sense of function behavior, similar to the sense-making needed for understanding rates of change and accumulation in calculus.

Class Activity Problem 5 : Part c

(c) Why does the result in Problem 5(b) make sense given what we know about visualizing derivatives and the graphs of inverse functions?

Sample Response:
The derivative at a point represents the slope of the tangent line at that point. If f has an inverse function, the graph of f^{-1} is the graph of f reflected over the line $y = x$. Finally, we can also easily see that the reflection of a linear function reflected over the line $y = x$ has the reciprocal slope of the original line. Putting all these facts together, it makes sense that the derivative of the inverse function is some kind of reciprocal, since it is the slope of the reflection of a tangent line.

Commentary:
Consider posing the following questions as undergraduates continue their work in groups:

- How might we visualize the derivative of a function on its graph?
- How does the graph of a function relate to the graph of its inverse function?
- Given an arbitrary linear function, $f(x) = mx + b$, what is its inverse function? What is the slope of this inverse function? How do the slope of the function and the slope of its inverse function relate?

Finally, for Problem 6, allow groups only a few minutes to try their procedure from Problems 5(a) and 5(b) on the composition $g(f(x))$ to see if it is still possible to solve for $g'(x)$. Facilitate a short discussion in which undergraduates establish that the function we "want the derivative of" should be the "inner" or "inside" function. Compare this to our method of solving for the derivative of arcsine or natural logarithm.

Class Activity Problem 6

6. What happens when we reverse the order of the composition in the previous problem before we differentiate? Can we still find a formula for $\frac{d}{dx}g(x)$? Why or why not?

Solution:
Taking the derivative of both sides of the equation, $g(f(x)) = x$, with respect to x gives us $g'(f(x)) \cdot f'(x) = 1$. This doesn't give us an easy way to solve for the derivative of g since the chain rule "pulled out" the derivative of f instead.

Wrap-Up (5 minutes)

Conclude the lesson by revisiting some of the ideas that were discussed during the lesson, emphasizing how they connect undergraduate calculus to secondary mathematics concepts. Some points to consider include:

- Inverse functions are more than just the result of an algorithm. They have important properties, such as:
  - The composition of a function with its inverse function is the identity and vice-versa.
  - The domain of the original function is the range of the inverse function and vice-versa.
  - The inverse function "undoes" the original function (and vice-versa!).
- In particular, we can use these ideas as a tool to help us find formulas for derivatives of functions, such as the inverse trigonometric functions, for which direct use of the definition of the derivative yields expressions that may require tools beyond first-semester calculus to resolve.
- The method of "switch x and y and solve for y" comes with certain drawbacks:
  - It disregards units which may be attached to the original variables.
  - It glosses over the fact that the domain and range must be appropriately adjusted.
  - It does not provide an opportunity for students to revisit composition of functions in this setting as well as develop computational fluency with this operation.

  Prospective teachers who understand the importance of the role of function composition in working fluently with inverse functions in later courses are better equipped to convey this understanding to their students, who in turn benefit from increased awareness of the power of inverse functions.

At the end of the lesson, you can collect exit tickets if you choose. See Chapter 1 for guidance on using exit tickets in instruction.

Homework Problems

At the end of the lesson, assign the following homework problems.

In Problem 1, undergraduates must attend to the fact that the domain and range of the original function are the range and domain of the inverse function, respectively. Asking leading questions to a hypothetical student requires undergraduates to identify the underlying contradictions that arise from failing to recognize this fact.

Homework Problem 1

1. A theater concludes that their total revenue for the week is a function of the number of tickets they sell. They use the equation $R(t) = 15t - 100$ to represent this relationship, where t is the number of tickets sold. Akira uses the method of "switch x and y and solve for y" to find the inverse function of the relationship described above.

$$\text{Let } r = 15t - 100$$
$$t = 15r - 100$$
$$t + 100 = 15r$$
$$r = \frac{1}{15}(t + 100)$$

What two questions would you ask Akira to help them see the limitations of their work? Why would your questions be helpful?

Sample Response:

- If you are asked to calculate the revenue from 20 tickets sold, which equation would you use?
- The equation $r = \frac{1}{15}(t + 100)$ implies that for $t = 0, r > 0$. Does this make sense in the context of the situation?

My two questions are helpful because they would show Akira that the notation used in their "solution" is misleading. By switching the variable names, they are no longer representing the situation as intended.

Solving Problem 2 requires an understanding of the fact that inverse functions "undo" the process of the original function. It also helps address the possible misconception that $(f \circ g)^{-1} = f^{-1} \circ g^{-1}$.

Homework Problem 2

2. Using the properties of inverse functions, find the inverse function of the composite function $h(x) = f(g(x))$, where both f and g are known to have inverse functions.

Solution:

$$
\begin{aligned}
h(x) &= f(g(x)) \\
f^{-1}(h(x)) &= f^{-1}(f(g(x))) \\
f^{-1}(h(x)) &= g(x) \\
g^{-1}(f^{-1}(h(x))) &= g^{-1}(g(x)) \\
g^{-1}(f^{-1}(h(x))) &= x
\end{aligned}
$$

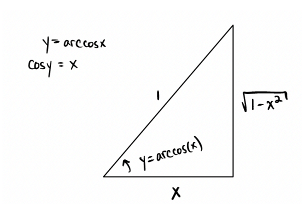
So, if $j(x) = g^{-1}(f^{-1}(x))$, then $j(h(x)) = x$. We can also show that $h(j(x)) = x$. By definition, we then have that $h^{-1}(x) = j(x) = g^{-1}(f^{-1}(x))$

Problem 3 gives undergraduates the opportunity to practice the techniques introduced in the lesson and look for patterns in their results. In particular, they will see how the derivatives of inverse trigonometric functions relate to the derivatives of their inverse cofunctions.

Homework Problem 3

3. Find a formula for $\frac{d}{dx}\arccos(x)$ in terms of x. Compare this formula to $\frac{d}{dx}\arcsin(x)$ (from Problem 3(b) in the Class Activity). What do you notice about the two derivatives?

Solution:



$$
\begin{aligned}
\cos(\arccos(x)) &= x \\
\tfrac{d}{dx}\cos(\arccos(x)) &= \tfrac{d}{dx}x \\
-\sin(\arccos(x)) \cdot \tfrac{d}{dx}\arccos(x) &= 1 \\
\tfrac{d}{dx}\arccos(x) &= \tfrac{-1}{\sin(\arccos(x))} \\
\tfrac{d}{dx}\arccos(x) &= \tfrac{-1}{\sin(y)} \\
\tfrac{d}{dx}\arccos(x) &= \tfrac{-1}{\sqrt{1-x^2}}
\end{aligned}
$$

The derivatives only differ by their sign; that is, $\frac{d}{dx}\arccos(x) = -\left(\frac{d}{dx}\arcsin(x)\right)$.

Problem 4 offers an interesting extension that examines functions with fixed points. Graphical representations of inverse functions, using equal-sized scales on each axis, exhibit symmetry about the line $y = x$. Fixed points also involve this identity line, and comparing the two concepts in Problem 4 helps undergraduates develop well-connected representations of both ideas.

Homework Problem 4

4. We call a point a in the domain of a function f a *fixed point* if $f(a) = a$.

   (a) Give an example of a continuous function with no fixed points.

   Sample Responses:
   - $f(x) = \ln(x)$
   - Any line parallel to $y = x$, such as $f(x) = x + 1$.
   - Any parabola which has been shifted up so that its graph no longer intersects the line $y = x$, such as $f(x) = x^2 + 100$
   - Any function f for which its graph lies strictly above the line $y = x$ or strictly below the line $y = x$.

   (b) Give an example of a continuous function with precisely one fixed point.

   Sample Responses:
   - $f(x) = \ln(x) + 1$ has exactly 1 fixed point at $(1, 1)$.
   - Any line that is not parallel to $y = x$, such as $f(x) = 2x$, will have one fixed point.
   - $f(x) = \frac{1}{2}(x + \frac{1}{2})^2$ has a fixed point at $\left(\frac{1}{2}, \frac{1}{2}\right)$.

   (c) For some continuous function f, assume its inverse function exists and has a fixed point (i.e., $f^{-1}(b) = b$). Does f have a fixed point? If so, for what value of x?

   Solution:
   If $f^{-1}(b) = b$, then $b = f(b)$ by the definition of an inverse function. This means any fixed point for the inverse function is also a fixed point for the original. Graphing the function and its inverse makes this clear, since they are reflections of each other over the line $y = x$.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Assessment Problems 1 & 2

1. A classmate is finding the formula of an inverse function and does not yet understand why the method of "switch x and y and solve for y" is problematic. Provide an example of a situation in which this approach is problematic and explain why.

   Sample Responses:
   - An equation like $y = x^2 + 2^x$ that mixes "types" of functions (e.g., quadratic and exponential) defy our solution strategies. We can neither "isolate the x^2 and take square roots" nor "isolate the 2^x and use logarithms." So we can find examples where the expression in terms of x makes it difficult or impossible to solve for y after a variable switch.
   - In the case of a function that relates two real-world quantities together, operations that ignore the units associated with the variables often leads to inconsistencies and incorrect conclusions (for example, the linear relationship between Fahrenheit and Celsius as seen in Problem 2(a) from the Class Activity). Switching the variables also interchanges the units, which could lead to confusion.

- The domain and the range of a function and its inverse are switched. If we switch x and y and solve for y, this obscures the fact that the x and y are "different" in a fundamental way than the variables we began with. For example, $y = \frac{1}{x+1}$ has a vertical asymptote at $x = -1$; if we write its inverse function (after "switching") as $y = \frac{1}{x} - 1$, we might also think $x \neq 0$, which isn't true in the case of the original function.

2. Show the steps used to find a formula for $\frac{d}{dx}\arctan(x)$ in terms of x. Express $\frac{d}{dx}\arctan(x)$ as a rational function.

   Solution:
   This is finding a formula for $\left(f^{-1}\right)'(x)$ where $f^{-1}(x) = \arctan(x)$ (and $f(x) = \tan(x)$). We also know that $\tan(\arctan(x)) = x$. We then take the derivative of both sides of this equation.

$$\begin{aligned} \tan(\arctan(x)) &= x \\ \frac{d}{dx}\tan(\arctan(x)) &= \frac{d}{dx}x \\ \sec^2(\arctan(x))\left(\frac{d}{dx}\arctan(x)\right) &= 1 \\ \frac{d}{dx}\arctan(x) &= \frac{1}{\sec^2(\arctan(x))} \end{aligned}$$

   The trig identity $\sec^2(\theta) = 1+\tan^2(\theta)$ allows us to rewrite the above as $\frac{d}{dx}\arctan(x) = \frac{1}{1+\tan^2(\arctan(x))}$. We use $\tan(\arctan(x)) = x$ to rewrite this as $\frac{d}{dx}\arctan(x) = \frac{1}{1+x^2}$.

2.6 References

[1] Cooney, T. J., Beckmann, S., & Lloyd, G. M. (2010). *Developing essential understanding of functions for teaching mathematics in grades 9–12*. National Council of Teachers of Mathematics.

[2] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from `http://www.corestandards.org/`

[3] Teuscher, D., Palsky, K., & Palfreyman, C. Y. (2018). Inverse functions: Why switch the variable? *The Mathematics Teacher*, *111*(5), 374–381.

[4] Wilson, F., Adamson, S., Cox, T., & O'Bryan, A. (2016, November 28). *Inverse functions: We're teaching it all wrong!*. AMS Blogs on Teaching & Learning Math. Retrieved from `https://blogs.ams.org/matheducation/2016/11/28/inverse-functions-were-teaching-it-all-wrong/`

[5] Zazkis, R. & Kontorovich, I. (2016). A curious case of superscript (-1): Prospective secondary mathematics teachers explain. *The Journal of Mathematical Behavior*, *43*, 98–110.

2.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

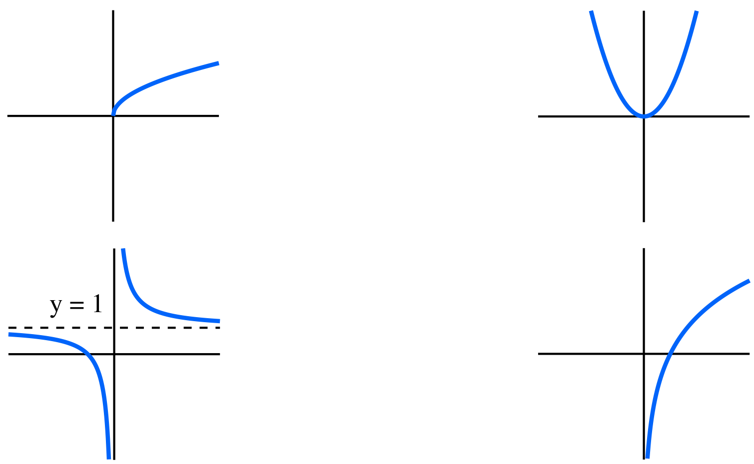
NAME: **PRE-ACTIVITY:** INVERSE FUNCTIONS AND THEIR DERIVATIVES (page 1 of 1)

1. Given a function f, its ***inverse function*** (if it exists) is the function f^{-1} such that $y = f(x)$ if and only if $f^{-1}(y) = x$.

   (a) If we know that $f(♫) = ❤$, what is $f^{-1}(❤)$? What about $f(f^{-1}(❤))$?

   (b) If we know a function has an inverse function, what do we know about the properties, behavior, or graph of its inverse function? Create a list of these attributes.

2. Consider each of the functions below.



   (a) Which of the these functions has an inverse function? Explain how you can know without being given the defining expression.

   (b) What is the domain and range of each function? What is the domain and range of its inverse function (if it exists)?

NAME: **CLASS ACTIVITY:** INVERSE FUNCTIONS AND THEIR DERIVATIVES (page 1 of 4)

1. Consider how Alex, Jordan, and Kelly found the inverse function of $f(x) = \frac{2}{3}x + 1$.

Alex's Work	Jordan's Work	Kelly's Work
$y = 2/3\, x + 1$ $x = 2/3\, y + 1$ $x - 1 = 2/3\, y$ $\frac{x-1}{2/3} = y$ $\boxed{\frac{3}{2}(x-1) = y}$	$f \circ f^{-1}(y) = y$ $f(x) = \frac{2}{3}x + 1$ So: $f(f^{-1}(y)) = \frac{2}{3}f^{-1}(y) + 1 = y$ $\frac{2}{3}f^{-1}(y) = y - 1$ $\frac{3}{2} \cdot \frac{2}{3} f^{-1}(y) = \frac{3}{2}(y-1)$ $\boxed{f^{-1}(y) = \frac{3}{2}(y-1)}$	$y = \frac{2x}{3} + 1$ $y - 1 = \frac{2x}{3}$ $3(y-1) = 2x$ $\frac{3(y-1)}{2} = x$ But $f(x) = y \Rightarrow$ $f^{-1}(f(x)) = f^{-1}(y) \Rightarrow$ $x = f^{-1}(y)$ So $\boxed{f^{-1}(y) = \frac{3(y-1)}{2}}$

Compare and contrast the key mathematical ideas used by Alex, Jordan, and Kelly to find the inverse function of $f(x) = \frac{2}{3}x + 1$. Make sure to identify which properties of inverse functions each student uses, if any.

2. Now consider two problems where a high school student used Alex's method of switching the variables and solving for the dependent variable to find the inverse function.

Find the inverse function of $T(C) = \frac{9}{5}C + 32$ where C is the temperature in Celsius and $F = T(C)$ gives the temperature in Fahrenheit.	Find the inverse function of $f(x) = \frac{2x+1}{x-1}$ for $x \neq 1$.
Let $F = \frac{9}{5}C + 32$ $C = \frac{9}{5}F + 32$ $C - 32 = \frac{9}{5}F$ $\frac{5}{9}(C-32) = F$	$y = \frac{2x+1}{x-1}, \quad x \neq 1$ $x = \frac{2y+1}{y-1}$ $(y-1)x = 2y+1$ $yx - x = 2y + 1$ $yx - 2y = 1 + x$ $y(x-2) = x+1$ $y = \frac{x+1}{x-2}, \quad x \neq 1$

(a) Describe why the student's work for the temperature function is problematic.

(b) Describe why the student's work for the rational function is problematic.

(c) What are the limitations of using Alex's method of switching the variables and solving for the dependent variable to find an inverse function? Would Jordan and Kelly have the same problem(s)? Explain.

3. We can use the properties of inverse functions to find their derivatives with respect to x.

   (a) Draw a right triangle that illustrates the relationship $y = \arcsin(x)$ for $0 < x < 1$.

   (b) Use the fact that $\sin(\arcsin(x)) = x$ and the chain rule to compute $\frac{d}{dx}\arcsin(x)$ in terms of x.

   (c) Use the fact that $e^{\ln(x)} = x$ and the chain rule to show that $\frac{d}{dx}.\ln(x) = \frac{1}{x}$

4. What key mathematical idea(s) from Problem 1 did we use to find the derivatives in Problem 3?

5. Let f be a function with inverse function g.

   (a) In the form of written sentences, describe how you would use the fact that $f(g(x)) = x$ and the chain rule to compute $\frac{d}{dx}g(x)$ for any function f with inverse function g.

   (b) Use the procedure you wrote above to compute $\frac{d}{dx}g(x)$ for any function f with inverse function g.

   (c) Why does the result in Problem 5(b) make sense given what we know about visualizing derivatives and the graphs of inverse functions?

6. What happens when we reverse the order of the composition in the previous problem before we differentiate? Can we still find a formula for $\frac{d}{dx}g(x)$? Why or why not?

NAME: **HOMEWORK PROBLEMS:** INVERSE FUNCTIONS AND THEIR DERIVATIVES (page 1 of 1)

1. A theater concludes that their total revenue for the week is a function of the number of tickets they sell. They use the equation $R(t) = 15t - 100$ to represent this relationship, where t is the number of tickets sold. Akira uses the method of "switch x and y and solve for y" to find the inverse function of the relationship described above.

   Let $r = 15t - 100$

   $t = 15r - 100$

   $t + 100 = 15r$

   $r = \frac{1}{15}(t + 100)$

   What two questions would you ask Akira to help them see the limitations of their work? Why would your questions be helpful?

2. Using the properties of inverse functions, find the inverse function of the composite function $h(x) = f(g(x))$, where both f and g are known to have inverse functions.

3. Find a formula for $\frac{d}{dx}\arccos(x)$ in terms of x. Compare this formula to $\frac{d}{dx}\arcsin(x)$ (from Problem 3(b) in the Class Activity). What do you notice about the two derivatives?

4. We call a point a in the domain of a function f a *fixed point* if $f(a) = a$.

   (a) Give an example of a continuous function with no fixed points.

   (b) Give an example of a continuous function with precisely one fixed point.

   (c) For some continuous function f, assume its inverse function exists and has a fixed point (i.e., $f^{-1}(b) = b$). Does f have a fixed point? If so, for what value of x?

NAME: **ASSESSMENT PROBLEMS:** INVERSE FUNCTIONS AND THEIR DERIVATIVES (page 1 of 1)

1. A classmate is finding the formula of an inverse function and does not yet understand why the method of "switch x and y and solve for y" is problematic. Provide an example of a situation in which this approach is problematic and explain why.

2. Show the steps used to find a formula for $\frac{d}{dx}\arctan(x)$ in terms of x. Express $\frac{d}{dx}\arctan(x)$ as a rational function.

3

Newton's Method

Single Variable Calculus

James A. M. Álvarez, *The University of Texas at Arlington*
Elizabeth W. Fulton, *Montana State University*
Andrew Kercher, *Simon Fraser University*
Elizabeth G. Arnold, *Colorado State University*
Elizabeth A. Burroughs, *Montana State University*
Kyle Turner, *The University of Texas at Arlington*

3.1 Overview and Outline of Lesson

In high school and undergraduate mathematics classes, students often solve equations of the form $f(x) = 0$, where $f(x)$ is a polynomial function. When $f(x)$ is a quadratic function, finding the zeroes of f is relatively straightforward, because students can use the quadratic formula. When $f(x)$ is a polynomial function of higher degree, or any other nonlinear function, a way to estimate or explicitly find the zeroes (if possible) of the function is to use a linear approximation. Linearizing a nonlinear function is an often used technique across mathematics. This lesson emphasizes linear approximation as a useful technique for finding zeroes of a function, either by hand, or with technology, focusing on Newton's method as an example of an algorithmic way to numerically find the zeroes of a function. Undergraduates explore the geometry and calculus used to develop Newton's method, derive and apply the Newton's method procedure, and analyze hypothetical student work as an application to teaching. Because the study of linear functions is a core component of high school algebra, studying Newton's method provides prospective teachers an opportunity to develop a deeper understanding of and an increased fluency with linear functions.

1. Launch—Pre-Activity
   Undergraduates complete the Pre-Activity prior to the lesson. The purpose of the Pre-Activity is to review writing equations of tangent lines and algebraic methods for finding zeroes of a function. Undergraduates will also consider a function, which will be used in the Class Activity, whose zeroes cannot be found algebraically.

2. Explore—Class Activity

   - *Problems 1–4*:
     Undergraduates follow three hypothetical students' reasoning to algebraically and graphically create tangent lines to determine estimates for a zero of a function by finding *x*-intercepts of the tangent lines. They apply this reasoning three times to eventually sketch the first three iterations of Newton's method. Then they discuss which iterations produce a better estimate of the zero and consider how linearization is at the core of this process. If you are teaching this lesson over two class sessions, this may be an appropriate place to end Day 1.

- *Problem 5*:
  Undergraduates generalize the graphical and algebraic processes that they used in Problems 1–4 to describe Newton's method. Undergraduates then discuss when they can stop the iterative process.
- *Discussion—Newton's Method*:
  After Problems 1–5, the instructor formally defines Newton's method according to the instructional material of their course; then, undergraduates determine how to use technology to more quickly compute several iterations of the algorithm.

3. Apply and Extend—Class Activity
   - *Problem 6*:
     Undergraduates first practice applying the Newton's method algorithm to approximate the zeroes of a function. Based on this experience, they describe how to tell if a choice of x_0 will be a "good" initial guess for finding a given zero of a function.
4. Closure—Wrap-up
   The instructor concludes the lesson by describing how Newton's method provides a mathematical algorithm for estimating zeroes of functions that can be used by calculation devices in producing estimates of the zeroes of a given function. Taking time to discuss that linear approximations can be used by a calculation device to produce solutions to an equation provides future teachers perspectives for helping their future students see that many of the "answers" to these types of problems are approximations or estimates of a possible exact solution. As calculation devices are used quite extensively in secondary school mathematics courses, this helps strengthen understandings that procedures similar to Newton's method provide the mathematical foundation and algorithms used by the calculation devices to produce answers within a given error tolerance.

3.2 Alignment with College Curriculum

Newton's method is a topic that naturally fits into an "Applications of Differentiation" section in single variable calculus courses. Undergraduates are asked to consider methods for solving for zeroes that they commonly use (e.g., factoring) and are engaged in thinking about what to do when those algebraic methods are not applicable. Exploring and explaining why Newton's method works offers undergraduates the opportunity to learn about how Newton's method is an application of linear approximation and that linearization in general is a commonly used technique in mathematics for analyzing nonlinear functions.

3.3 Links to School Mathematics

By studying Newton's method, prospective teachers will develop deeper understanding of how an algorithmic method can be used to approximate zeroes of functions when the algebraic methods they teach fail to work. The *Mathematical Education of Teachers II* (CBMS, 2012) recommends that prospective teachers both derive results that may have been taken for granted in high school and that prospective teachers become familiar with technology. (Note that most graphing calculators do not use Newton's method in their algorithm; they likely use a version of a QR-Algorithm. See, for example, `http://sections.maa.org/okar/papers/2010/lloyd.pdf`).

This lesson highlights the following:

- Connections between calculus concepts and mathematical algorithms for calculating zeroes of a function;
- The use of linear functions as a tool for analyzing non-linear functions.

This lesson addresses several mathematical knowledge and practice expectations included in high school standards documents, such as the Common Core State Standards for Mathematics (CCSSM, 2010). The lesson emphasizes an application in which high school students use their knowledge of linear functions in a flexible manner. They look for key features (e.g., x-intercepts) of the graph of a function and use symbolic expressions coupled with graphical representations and the use of technology to reach resolutions to the tasks (c.f. CCSS.MATH.CONTENT.HSF.IF.C.7). The lesson relies on looking for and expressing regularity in repeated reasoning to uncover Newton's method and provides opportunities to consider the reasoning of others as well as construct sound arguments to support conclusions (c.f. CCSS.MATH.CONTENT.HSF.LE.B5, CCSS.MATH.PRACTICE.MP3).

3.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know how to:

- Find the zeroes of a polynomial function both algebraically (when possible) and graphically;
- Compute derivatives of polynomial functions;
- Write an equation of a tangent line.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Apply Newton's method, both graphically and algebraically, to approximate zeroes of a function;
- Flexibly use the derivative of a function to determine the slope of the tangent line to the graph of a function at a given point;
- Explain how to use Newton's method to compute the zeroes of a function;
- Analyze hypothetical student work in order to derive the procedure for Newton's method;
- Evaluate questions one might ask a hypothetical student to guide their understanding of Newton's method.

Anticipated Length

One or two 50-minute class sessions, depending on the instructor's pacing choices.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity (print out Problems 1–4 and 5–6 to pass out separately)
- Computer (for instructor to display a dynamic sketch during the Class Activity)
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

3.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to be completed in preparation for this lesson. We recommend that you collect this Pre-Activity the day before the lesson so that you can review undergraduates' responses before you begin the Class Activity. This will help you determine if you need to spend additional time reviewing the responses to the Pre-Activity with your undergraduates.

Pre-Activity Review (10 minutes)

Discuss undergraduates' responses to the Pre-Activity as needed. The point of this Pre-Activity is to refresh undergraduates' memory of (1) how to write equations of tangent lines in point-slope form; (2) how to find zeroes of a function algebraically; and (3) the idea that you can't algebraically find zeroes of some functions. The last problem of the Pre-Activity contains the situation that undergraduates will consider in the Class Activity.

Pre-Activity Problems 1, 2 & 3

1. Write an equation of **a line** with slope 3 that passes through the point $(2, 1)$ in point-slope form. Then, write an equation of this line in slope-intercept form.

   Solution:
   The equation of the line in point-slope form is $y - 1 = 3(x - 2)$. Converted to slope-intercept form, this is $y = 3x - 5$.

2. Write an equation of the **tangent line** to the graph of $f(x) = x^2 + 2$ at the point $(1, 3)$ in point-slope form. Then, write an equation of this tangent line in slope-intercept form.

   Solution:
   Since $f'(x) = 2x$, $f'(1) = 2(1) = 2$ is the slope of the tangent line. Then, the equation of the line in point-slope form $y - 3 = 2(x - 1)$. Converted to slope-intercept form, this is $y = 2x + 1$.

3. More with tangent lines.

   (a) For a given function f, describe how to find an equation of the tangent line to the graph of f at $x = a$.

   Solution:
   To find the equation of a line, we need a point and slope. Then, we can substitute these values into the point-slope form of a line. To find the slope, we need to take the derivative of the function at $x = a$ (i.e., we need to find $f'(a)$). To find a point on the line, we can use $(a, f(a))$. When we substitute these values into the point-slope form of a line, we get $y - f(a) = f'(a)(x - a)$.

   (b) Now, write an equation of the tangent line to the graph of f at $x = a$.

   Solution:
   Calling this function g, we have that $g(x) = f'(a)(x - a) + f(a)$ or $y = f'(a)(x - a) + f(a)$.

Commentary:
When reviewing undergraduates' work for Problems 1–3, consider the following points:

- Verify that undergraduates are writing equations in point-slope form for Problems 1–3.
- Although Problems 2 and 3 may serve as review of concepts addressed earlier in your course, ensure that undergraduates effectively use the derivative of a function at a point to determine the slope of the corresponding tangent line on the graph of the function.
- Undergraduate responses to 3(a) should be in the form of written sentences that describe how and why their process will correctly produce an equation for the tangent line.

To motivate the Class Activity, facilitate a class discussion on Problems 4 and 5. Questions to ask about Problem 4 to motivate discussion include:

- How did you find the zeroes of each of these functions?
- Do these techniques always work? When is one technique preferred over another?

Pre-Activity Problem 4

4. Find the zeroes of the following functions.

   (a) $f(x) = x^2 - 4$

   Solution:
   Undergraduates will likely use one of two approaches:

Difference of Squares:

$$
\begin{aligned}
x^2 - 4 &= 0 \\
(x-2)(x+2) &= 0 \\
x \pm 2 &= 0 \\
x &= \pm 2
\end{aligned}
$$

Solving for x:

$$
\begin{aligned}
x^2 - 4 &= 0 \\
x^2 &= 4 \\
x &= \pm 2
\end{aligned}
$$

(b) $g(x) = 3x^2 + 7x - 2$

Solution:
Using the quadratic formula on $3x^2 + 7x - 2 = 0$:

$x = \frac{-7 \pm \sqrt{49 - 4(3)(-2)}}{2(3)} = \frac{-7 \pm \sqrt{73}}{6}$

(c) $h(x) = x^3 + x^2 - 2x$

Solution:
By factoring:

$$
\begin{aligned}
x^3 + x^2 - 2x &= 0 \\
x(x^2 + x - 2) &= 0 \\
x(x-1)(x+2) &= 0 \\
x &= 0, 1, -2
\end{aligned}
$$

Commentary:
We have found that most undergraduates will factor or use the quadratic formula to find the zeroes of these functions. Make sure that undergraduates recognize that these techniques only apply to some functions. As appropriate for your class, you may also want to discuss *why* each method works. For instance,

- When we factor, we use the fact that a product of real numbers can only be zero if at least one of the factors is zero.
- The quadratic formula is derived from completing the square.

Finally, ask undergraduates how finding zeroes of a function has arisen in our study of calculus. Discuss when finding zeroes is used in calculus (e.g., for identifying critical points and points of inflection) and discuss the following connection to teaching:

Discuss This Connection to Teaching

High school algebra courses focus on different techniques for finding zeroes of a function; these commonly include algebraic techniques (e.g., factoring) and graphing strategies (e.g., using graphing calculators, Desmos). While learning to solve for zeroes of polynomial functions remains the focus of many lessons, undergraduates may not have had opportunities to think about the reasons for focusing on solving for zeroes or situations in which these skills and techniques provide insight when solving problems. Contextualize the focus in high school algebra on solving for zeroes of functions by pointing out to undergraduates that solving for zeroes arises in the context of calculus, for example, with optimization problems.

Before reviewing Problem 5, let undergraduates know that the function they are considering in this problem will be used throughout the Class Activity.

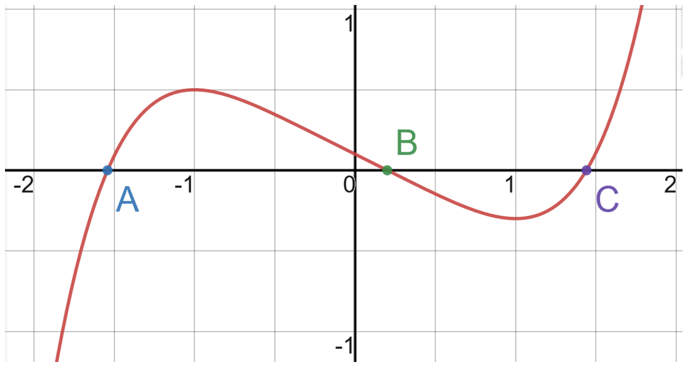
Pre-Activity Problem 5

5. Consider the function, $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$

   (a) Nnamdi has excellent algebra skills, yet he tries to find the zeroes algebraically and gets stumped. Explain why he is having trouble.

   Solution:
   $f(x)$ is a fifth degree polynomial. The usual algebraic methods of (e.g., solving explicitly for x, factoring, the quadratic formula) do not work for most quintic polynomials. In fact, there is no algebraic method that is guaranteed to solve quintic polynomials.

   (b) Nnamdi decides to graph f to find the zeroes. The zeroes are indicated on the graph as A, B, and C. Estimate the value of C.



   Solution:
   The answer rounded to the nearest thousandth is 1.441. From our experience, undergraduates will either make their best guess by looking at the graph or use a graphing utility to find a more exact answer.

   Commentary:
   The first part of the Class Activity focuses on the zero at C. Later, undergraduates will also examine the zeroes at A and B.

Ask undergraduates how a calculator or graphing utility might find the zero at C. Tell undergraduates that there is no general formula for algebraically finding zeroes of quintic polynomials explicitly and that, in general, for most non-linear functions the calculator is only able to estimate a zero very closely.

Discuss This Connection to Teaching

When they were in high school, undergraduates may have used calculators and computers to find the zeroes of a function. These tools use estimation methods that are hidden from the user. One goal of this lesson is to show undergraduates that these devices operate on programmed algorithms that use or extend mathematical concepts introduced in calculus. Furthermore, using calculus and their understanding of linear functions, they can develop an algorithm by hand for finding zeroes of a function and gain some insight into how their calculation device might be generating the zeroes. In addition to examining how calculation devices might be generating estimates, this lesson also provides undergraduates with a practical example of how linear functions are a useful tool when analyzing non-linear functions.

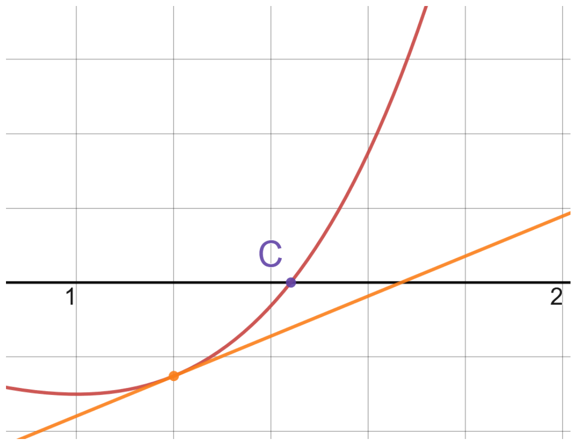
Class Activity: Problems 1–4 (30 minutes)

Pass out **Problems 1–4** of the Class Activity. Instruct undergraduates to work in small groups on Problem 1. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion.

Class Activity Problem 1

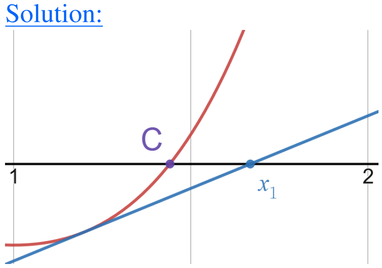
Recall the function from Problem 5 on the Pre-Activity, $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$, for which Nnamdi wanted to find the zeroes of the function. Nnamdi initially thinks that $x = 1.2$ is a good estimate of the zero, C, but when he zooms in on the graph he realizes that C is further to the right. He starts to experiment with linear functions to try to find a better estimate for C.

1. Nnamdi zooms in on the graph and sketches the tangent line at $x_0 = 1.2$ (see graph below).



(a) Label the x-intercept of Nnamdi's tangent line as x_1.

Solution:



(b) Write an equation of Nnamdi's tangent line in point-slope form and find the value of x_1.

Solution:
Since $f'(x) = \frac{1}{2}x^4 - \frac{1}{2}$, we can compute that $f'(1.2) \approx 0.537$. Furthermore, $f(1.2) \approx -0.251$. Taken together, we can write the equation of the tangent line as $y + 0.251 = 0.537(x - 1.2)$. Then, we can calculate the value of x_1 by finding the x-intercept of this line:

$$\begin{aligned} 0 + 0.251 &= 0.537(x_1 - 1.2) \\ 0.251 &= 0.537x_1 - 0.644 \\ 0.895 &= 0.537x_1 \\ x_1 &= 1.667 \end{aligned}$$

Commentary:
Circulate the room while undergraduates work. As you do so,

- Make sure undergraduates are using the algebraic methods from Problems 2 and 3 in the Pre-Activity.
- Watch for undergraduates who try to visually estimate the slope of (or a point on) the tangent line. Remind these undergraduates that they can find the slope by taking the derivative of the function at $x = 1.2$ and they can find a point on the tangent line by plugging $x = 1.2$ into the original equation.
- You may want to encourage undergraduates to round to the nearest thousandth because later questions will ask for accuracy to the thousandths place.

When you have noticed that most groups are done, bring the class back together for a whole group discussion. Briefly discuss undergraduates' responses to Problem 1(b) before instructing undergraduates to work in small groups on Problems 2–4.

Class Activity Problem 2

2. Taking inspiration from Nnamdi's idea, Mari decides to sketch another tangent line to the graph of $f(x)$ at the point $(x_1, f(x_1))$. She claims that the x-intercept of her tangent line will be closer to the zero C than x_1.

(a) Do you agree with Mari's claim? Explain why or why not.

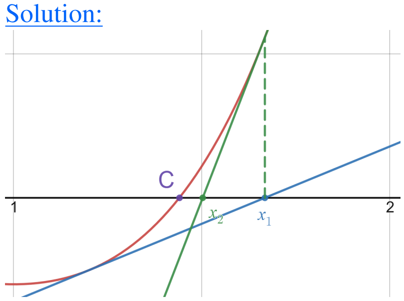
Sample Responses:

- Yes. It looks like the tangent line that Mari will draw is a better linear approximation of the function f, which will give a better approximation of the zero.
- No, since x_1 was farther away from the zero than Nnamdi's x_0.

Commentary:
We have found that most undergraduates will agree with Mari's claim. You might encourage undergraduates to consider whether Mari's claim would be true in general (i.e., for any choice of x_0 and any arbitrary function). Mari's claim is not always true. Undergraduates do not need to recognize this fact yet in their answer, however; throughout the lesson, they will see applications of Newton's method whose successive approximations do not always get closer to a zero.

(b) Sketch in Mari's tangent line. Label the x-intercept of her tangent line as x_2.

Solution:



Commentary:
As you circulate the room identify groups that could present their responses to the class (ideally, use a document camera to showcase their work). Alternatively, you can use undergraduates' work to reconstruct the tangent lines and zeroes using a computer graphing tool (e.g., Desmos) to share with the class.

(c) Write the equation of Mari's tangent line in point-slope form and find the value of x_2.

Solution:
Since $f'(x) = \frac{1}{2}x^4 - \frac{1}{2}$, we can compute that $f'(1.667) \approx 3.361$. Furthermore, $f(1.667) \approx 0.554$. Taken together, we can write the equation of the tangent line as $y - 0.554 = 3.361(x - 1.667)$. Then, we can calculate the value of x_2 by finding the x-intercept of this line:

$$\begin{aligned} 0 - 0.554 &= 3.361(x_2 - 1.667) \\ -0.554 &= 3.361x_2 - 5.603 \\ 5.049 &= 3.361x_2 \\ x_2 &= 1.502 \end{aligned}$$

Class Activity Problem 3

3. Amy uses both Mari's and Nnamdi's ideas to find a point, x_3, even closer to the zero C.

 (a) What do you think she did? Explain.

 Solution:
 I think that Amy drew a tangent line to the graph of $f(x)$ at the point $(x_2, f(x_2))$. She labeled the x-intercept of this tangent line as x_3. Then, she found the equation of this tangent line and solved for its x-intercept, which is x_3.

 Commentary:
 If undergraduates struggle with this, ask them to describe Nnamdi's and Mari's procedures to you.

 (b) Find the value of x_3.

 Solution:
 Since $f'(x) = \frac{1}{2}x^4 - \frac{1}{2}$, we can compute that $f'(1.502) \approx 2.045$. Furthermore, $f(1.502) \approx 0.113$. Taken together, we can write the equation of the tangent line as $y - 0.113 = 2.045(x - 1.502)$. Then, we can calculate the value of x_3 by finding the x-intercept of this line:

$$\begin{aligned}
0 - 0.113 &= 2.045(x_3 - 1.502) \\
-0.113 &= 2.045x_3 - 3.072 \\
2.959 &= 2.045x_3 \\
x_3 &= 1.447
\end{aligned}$$

 Commentary:
 - As you circulate the room identify groups that could present their responses to the class (ideally, use a document camera to showcase their work). Alternatively, you can use undergraduates' work to reconstruct the tangent lines and zeroes using a computer graphing tool (e.g., Desmos) to share with the class.
 - Encourage undergraduates to sketch a tangent line to $f(x)$ at $(x_2, f(x_2))$ and label the zero of the tangent line x_3. They should also (algebraically) find an equation for the tangent line to $f(x)$ at $(x_2, f(x_2))$ and determine the x-intercept to find x_3.

After a class discussion of Problems 2 and 3, review Nnamdi's, Mari's, and Amy's process. Then, ask your undergraduates why Nnamdi might have thought to use a line to begin with. From our experience, this question has prompted the following ideas from undergraduates:

- We can easily solve for the zero of a line (i.e., a first-degree polynomial function).
- Lines are "simpler" functions that, when zoomed in, approximate a continuous function.

Emphasize (or remind undergraduates) that linear approximations are generally much more accurate near the point of tangency, which is one reason that a good first guess will improve the likelihood that Newton's method works as expected. To supplement this idea, make sure to address the following connection to teaching during this discussion:

Discuss This Connection to Teaching

Linear functions (where "linear" refers to the fact that the graph of the function is a straight line) are an essential topic in a high school mathematics curriculum. Underscoring the importance of linear functions as a fundamental tool in simplifying complex problems in calculus and higher levels of mathematics provides prospective teachers insight into the many ways that linearization plays a critical role in mathematics and statistics. Moreover, they see that developing students' fluency with linear functions and related concepts will enhance students' capacity to use these ideas flexibly in future courses such as calculus.

We include a table in Problem 4 so that undergraduates can organize their work from Problems 1–3. The table helps undergraduates organize and collate their data in a manner that may better facilitate drawing conclusions based upon their calculations.

Class Activity Problem 4

4. Fill in the following table with the values of x_1, x_2, and x_3 that you found above. Describe what you notice about these values.

x_0	x_1	x_2	x_3
1.2			

Solution:
Filled in table:

x_0	x_1	x_2	x_3
1.2	1.667	1.502	1.447

Descriptions of what undergraduates notice will vary. Sample responses include:

- The first zero moves away from C, but afterwards the approximations move closer to C.
- From graphing, it seems like x_3 is very close to the real zero—it would be hard to draw the next iteration.

Commentary:
Undergraduates should be able to quickly fill in the table in Problem 4 using their work from Problems 1–3. You might give them a few minutes to observe patterns in the data in small groups before beginning a classroom discussion.

Advice on Delivering the Lesson Over Two Class Sessions

If you are teaching this lesson over two class sessions, this may be an appropriate place to end Day 1. See Chapter 1 for guidance on using exit tickets to facilitate instruction in a two-day lesson.

Class Activity: Problem 5 (15 minutes)

Pass out **Problems 5 and 6** of the Class Activity. Before instructing undergraduates to work in small groups on Problem 5, summarize the following connection to teaching:

Discuss This Connection to Teaching

For Problem 5, undergraduates will consider the details of Newton's method both graphically and algebraically. Using multiple representations is a key idea in high school mathematics. Links between algebraic and graphical representations of functions, for instance, are especially important in studying relationships and change.

As you circulate your classroom, you may want to identify groups to present their work to the class. (See chapter 1 for advice about selecting and sequencing student work for use in class.)

Class Activity Problem 5

5. The iterative process Amy follows from the work of Mari and Nnamdi is called Newton's method. To apply Newton's method, the process of "finding a tangent line at the point on the graph corresponding to the guess for the zero, finding its x-intercept, and using this x-intercept as the next guess for the zero" is repeated. These x-intercepts (usually denoted $x_0, x_1, x_2, x_3, \ldots$) provide successive approximations of the value of a zero of a function.

(a) Describe this process graphically.

Solution:
Make an initial estimate x_0 that is close to a zero of a given function. Sketch the tangent line to the curve at x_0. Then label the x-intercept of that tangent line. This x-intercept becomes x_1. Repeat this process of sketching the tangent line to the curve at x_1 and finding the x-intercept of the tangent line.

Commentary:

- Make sure undergraduates clearly identify the point on the function to which their line is tangent.
- After undergraduates share their ideas with the class, the dynamic sketch, found at `https://www.desmos.com/calculator/revqc4ybqz`, can help them visualize their ideas.

(b) Describe this process algebraically. Write out a formula to find x_{n+1}, the x-intercept of the tangent line created from the previous guess, x_n.

Solution:

1. Make an initial estimate x_0 that is close to a zero of a given function.
2. Write the equation of the tangent line,
$$\begin{aligned} y - f(x_n) &= f'(x_n)(x - x_n) \\ y &= f'(x_n)(x - x_n) + f(x_n) \end{aligned}$$
3. Determine a new approximation by solving for the zero of this tangent line. That is, since x_{n+1} is the x-intercept of this line, we solve the equation $0 = f'(x_n)(x_{n+1} - x_n) + f(x_n)$ for x_{n+1}:
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Commentary:
If undergraduates have difficulty getting started, ask them to reconsider their work in Problems 1–3 and think about the process they used. If further scaffolding is needed, help them notice that for the general process they should consider using $f(x)$ instead of the specific function used, say, in Problems 1–3.

(c) How do you know when to stop this iterative process? That is, when is your approximation of a zero "good enough?"

Sample Responses:

- You have to determine the appropriate accuracy. You repeat the process until the difference between iterations is smaller than the desired accuracy. For instance, if you need to be accurate to the thousandth, you need to repeat the process until $|x_{n+1} - x_n| < 0.001$.
- Referring to limitations of drawing on the graph: We are done after a couple of iterations.
- Never. You won't ever reach the zero.

Commentary:

- We have found that undergraduates tend to focus on the limitations of sketching successive lines on the graph or on the notion that somehow this process may never produce an exact zero. Emphasize the difference between closed form algebraic formulas for solving for zeroes versus approximations based upon iterative process such as this one.
- To further facilitate group interactions as undergraduates work on this problem, you might consider asking them to revisit Problem 4 and to then compute x_4 and x_5. If you choose to do this, consider the following questions:
  - Does the pattern you observed in Problem 4 hold? Do you see any new patterns?
  - What do you expect from the next iteration? Why?
  - What do you think it means to be "good enough" in this context?
  - When might we need a solution to be accurate to 3 decimal places? 8? 100?

Discussion: Newton's Method (5 minutes)

Formally define Newton's method according to the instructional material of your course. Undergraduates will calculate several iterations of Newton's method in Problem 6, so this may also be a good time to discuss or demonstrate ways in which technology can be used to ease calculations of successive Newton's method approximations.

Newton's method connects undergraduates' prior experiences with linear functions and their recent understandings about the relationship between the derivative of a function at a point and the slope of the tangent line to the graph of the function at that point. As undergraduates rediscover or develop Newton's method, they gain more experience with iterative processes in a context that builds upon these ideas. Their experiences in this lesson reinforce ideas about derivatives introduced early in the calculus course and rely on their having developed fluency with high school algebra concepts related to writing an equation of a line given a point and a slope and finding x-intercepts.

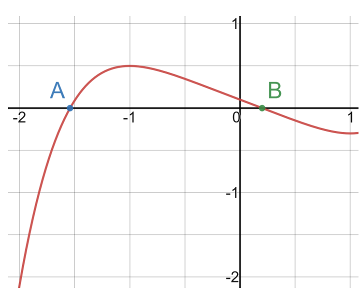
Class Activity: Problem 6 (15 minutes)

Instruct undergraduates to work in small groups on Problem 6. As a class, it may be helpful to first discuss 6(a)ii and compute the corresponding iterations for 6(b) if undergraduates seem to be unsure of how to proceed on this problem.

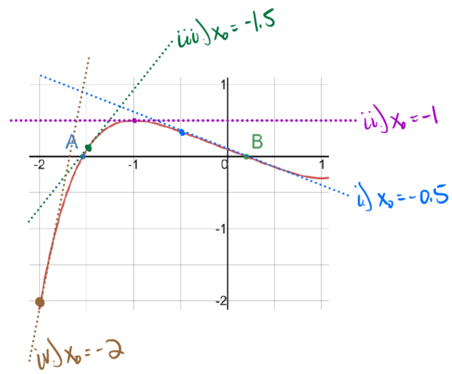
This problem focuses on the other two zeroes of $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$. The aim is to generate discussion about ways in which Newton's method can fail, how initial guesses can lead to different zeroes or none at all. Additionally, undergraduates engage in further practice in applying Newton's method.

Class Activity Problem 6

6. Reconsider $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$. Nnamdi now wants to use Newton's method to approximate the zero, A. He wonders what will happen if he uses the following initial guesses: -0.5, -1, -1.5, and -2.



For Problem 6(a), make sure undergraduates state what zero (A or B) they think each initial guess will lead to *before* they draw tangent lines on the graph. *After* drawing tangent lines, they should also note what zero (if any) they found. Below graphically shows (with tangent lines) what happens when you apply Newton's method using these initial guesses.



Class Activity Problem 6 : Part a

(a) Without doing any calculations, which zero of f do you expect each of these initial guesses to lead? Explain your reasoning. Use the graph above to graphically show (by drawing tangent lines) what happens when you apply Newton's method using these initial guesses.

i. $x_0 = -0.5$

Solution:

This guess leads to B. Sample explanation: The function f is very straight between $x = -0.5$ and the zero B, so the tangent line at $(-0.5, f(-0.5))$ will be a good linear approximation of the function and will cross the x-axis very close to B.

ii. $x_0 = -1$

Solution:

This guess does not lead to zero. The tangent line at $x = -1$ is horizontal and won't cross the x-axis.

Commentary:

It will help with the homework problems if undergraduates recognize that there is a horizontal tangent line at $x = -1$.

iii. $x_0 = -1.5$

Solution:

This guess leads to A. Sample explanation: -1.5 is already very close to the zero at A, so Newton's method will approximate that zero.

iv. $x_0 = -2$

Solution:

This guess leads to A. Sample explanation: Because of the shape of the graph around $x = -2$, the tangent lines won't ever cross the x-axis near any of the other zeroes.

In Problem 6(b), we recommend that undergraduates use technology to quickly apply Newton's method to find the zeroes. Problem 6(c) provides an opportunity to discuss connections between the graphical and algebraic representations of Newton's method seen in 6(a) and 6(b).

Class Activity Problem 6 : Parts b & c

(b) Use Newton's method with all four initial guesses to calculate a zero of f. Give your answer to three decimal places, when applicable.

Solution:

i. $x_0 = -0.5 \to x_1 = 0.24 \to x_2 = 0.200$
ii. $x_1 = -1 \to$ No. There's a horizontal tangent line at x_1.
iii. $x_0 = -1.5 \to x_1 = 1.545 \to x_2 = -1.542$
iv. $x_0 = -2 \to x_1 = -1.72 \to x_2 = -1.579 \to x_3 = -1.544 \to x_4 = -1.542$

Commentary:
For each initial guess, ask undergraduates how many iterations it took to get a sufficiently accurate estimate of the zero. Compare the number of iterations required to get a "close enough" approximation when the initial guesses led to the same zero (i.e., for $x_0 = -1.5$ and $x_0 = -2$).

(c) Summarize to Nnamdi what you observe in the graph of f that indicates what zero you will approximate given your initial guess.

Solution:
Answers will vary. Key points that you might look for in undergraduate responses include:

- Choosing an initial guess very close to the desired zero is most effective for approximating that zero (compare speed of convergence for $x_0 = -1.5$ versus $x_0 = -2$).
- Choosing an initial guess that creates a horizontal tangent line causes Newton's method to fail (see: $x_0 = -1$).
- Choosing an initial guess that produces a tangent line that is a good linear approximation of the function near the zero is very effective for approximating that zero (see: $x_0 = -0.5$)

Commentary:
To discuss connections between the graphical and algebraic representations of Newton's method, consider asking the class the following prompts:

- Do all initial guesses lead to the same zero? Why or why not?
- Which initial guesses did (or did not) lead to the same zero? When two or more guesses led to the same zero, which of them reached that zero faster? Why do you think this is?
- We saw that choosing an initial guess which is also local extremum causes Newton's method to fail. How can you explain this failure graphically? Algebraically?

Wrap-Up (5 minutes)

Discuss what undergraduates explored throughout the Class Activity and how it relates to other ideas in calculus:

- During this lesson we applied Newton's method to estimate zeroes of a function when algebraic techniques are insufficient.
- Newton's method is an application of the idea of linear approximation. In order to work with tangent lines in this way, we must be able to find their slopes by taking a derivative.
- The iterative process and algorithm associated with Newton's method can be easily programmed into a programmable computing device.

Homework Problems

At the end of the lesson, assign the following homework problems.

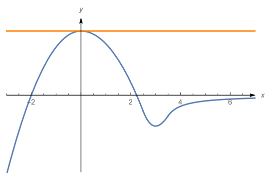
In the Class Activity, undergraduates saw that horizontal tangent lines lead to Newton's method failing to reach a zero. In Problem 1, undergraduates explore other ways that Newton's method can fail to converge. Having undergraduates formulate ideas about productive initial guesses versus unproductive ones engages undergraduates in the practice of using appropriate tools strategically and constructing viable reasons for their conclusions.

Homework Problem 1

1. The graph of $y = f(x)$ is shown here. Use the initial guesses given to determine which successfully lead to an approximation of a zero of the function f when using Newton's method. For each initial guess, graphically (by drawing tangent lines) support your conclusion based upon using Newton's method and explain your reasoning.

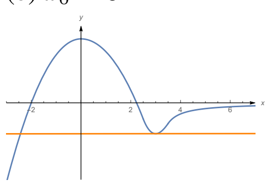
   Solutions:

(a) $x_0 = 0$



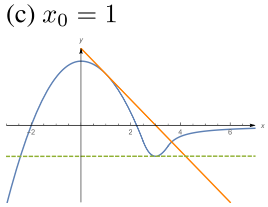
If $x_0 = 0$, the tangent line is horizontal. Thus, Newton's method fails.

(b) $x_0 = 3$



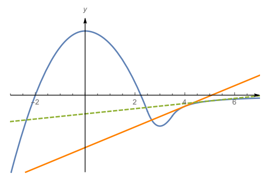
If $x_0 = 3$, the tangent line is horizontal. Thus, Newton's method fails.

(c) $x_0 = 1$



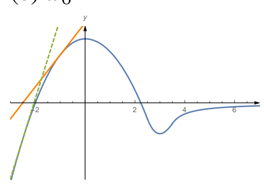
If $x_0 = 1$, then $x_1 = 3$ and we are now in the same situation as part (b). Thus, Newton's method fails.

(d) $x_0 = 4$



The iterations increase without bound. The sequence of iterations does not converge to a zero of the function. Thus, Newton's method fails.

(e) $x_0 = -1$

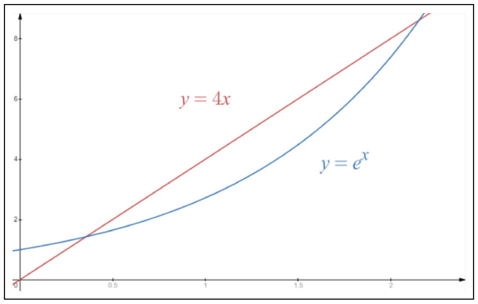


With this initial guess, the sequence of approximations will converge to -2. Thus, Newton's method works!

Problem 2 connects Newton's method to another area of secondary school mathematics: solving systems of equations. Just as we sometimes cannot solve for zeroes explicitly, we might also not be able to solve systems of equations using analytical methods. Instead, Newton's method gives us a way to find the solutions.

Homework Problem 2

2. Consider the system of equations given below.



$$y = e^x$$
$$y = 4x$$

(a) Explain how you could use Newton's method to approximate the two solutions to the system of equations.

Solution:
We create the function $h(x) = e^x - 4x$. The roots of this function are where $h(x) = 0$; that is, where $e^x - 4x = 0 \Rightarrow e^x = 4x$. Thus, when we apply Newton's method to find the roots of $h(x)$, we are also finding the solutions to the system.

(b) Choose any initial guess and calculate one iteration of Newton's method for each solution. Record your approximations up to six decimal places.

Sample Response:
First, note that $h'(x) = e^x - 4$. Using an initial guess of $x_0 = 0.5$ for the left-hand solution:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0.5 - \frac{-0.351279}{-2.351279} \approx 0.350601$$

Using an initial guess of $x_0 = 2$ for the right-hand solution:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 2 - \frac{-0.610943}{3.389056} \approx 2.180270$$

The next problem provides another example of Newton's method failing even though the tangent line is not horizontal—but this time, the reason for the failure is only obvious when explored algebraically rather than graphically. In Problem 3, undergraduates examine a hypothetical student's work and analyze questions one might ask the student to help guide them towards understanding how to adjust their work.

Homework Problem 3

3. Madalena is trying to use Newton's method to find the zeroes of the function $f(x) = x^3 - 2x + 2$. After making sure that $f'(1) \neq 0$, she chooses $x_0 = 1$. Then, she computes the first few iterations of Newton's method and begins a table of values:

x_0	x_1	x_2	x_3
1	0	1	

Madalena sees that this pattern will continue and comes to you for help.

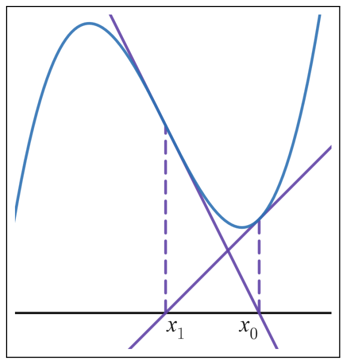
(a) Show that Madalena's computations for x_1 and x_2 are correct.

Sample Response:
With $f(x) = x^3 - 2x + 2$, $f'(x) = 3x^2 - 2$. Using Madalena's initial guess of $x_0 = 1$:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{1}{1} = 0$$
$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0 - \frac{2}{-2} = 1$$

(b) Create a graph of f and sketch tangent lines to explain why Newton's method has failed.

Solution:
We see below that Newton's method fails because the x-intercept of the tangent line to the graph at x_1 is, again, x_0; i.e., $x_2 = x_0 = 1$. But because we use the same iterative process at each step, this next tangent line will take us back to x_1 again; i.e., $x_1 = x_3 = 0$. This will continue indefinitely: $x_{2n} = 1$ and $x_{2n+1} = 0$ for all non-negative integers n.



(c) Consider the following questions that you might ask Madalena:

i. Explain how the question below might help you assess what Madalena understands about Newton's method:

Given a graph of f and an initial guess x_0, how could you find x_1 without making the calculations you have already tried?

Sample Response:
It's not clear that Madalena knows what Newton's method looks like graphically. This question would help to assess whether or not Madalena understands what her calculations represent and how they (usually) produce closer approximations.

ii. Explain how following up with the next question might help Madalena to advance in her understanding of Newton's method:

Considering your graphical explanation of Newton's method, how could two iterations of Newton's method have the same value?

Sample Response:
This question would lead Madalena to the fact that the tangent line to the graph drawn at x_1 would have to "point back at" x_0. Then, she might understand that Newton's method has failed because it is stuck in a loop.

iii. Explain why the question below might not help Madalena:

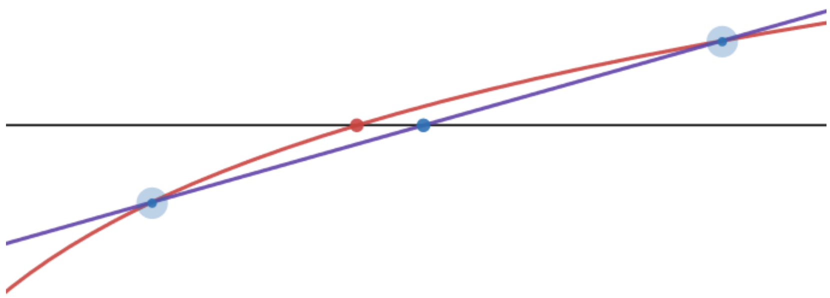
What is the formula for Newton's method?

Sample Response:
Madalena has already been using the correct formula to calculate Newton's method. Asking her to reproduce it wouldn't help her understand what's going on.

The purpose of Problem 4 is twofold: first, it allows undergraduates to connect Newton's method to the procedure you use to find the zero of a function with a graphing calculator (e.g., left bound, right bound, guess). This creates a connection to high school that may give prospective teachers a unique perspective on technology used in their classroom. Simultaneously, this question asks undergraduates to consider and respond to a reasonable (but flawed) suggestion from a hypothetical student.

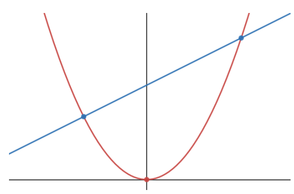
Homework Problem 4

4. Terrance, a high school algebra student, is using his TI-84 graphing calculator to find the zero of a function. To do so, the calculator requires him to choose a left bound (a point on the graph to the left of the zero), a right bound (a point on the graph to the right of the zero), and a guess (a point on the graph very close to the zero). Terrance thinks that the calculator is using the bounds he has chosen to construct a secant line, which it uses to approximate the zero. He draws the following example to illustrate his idea.



   (a) Under what circumstances would Terrance's "secant method" fail to approximate a zero? Create a graph of one such example.

   Solution:
   This method will not work if the graph of the function touches the x-axis, but does not cross the x-axis (see graph below). In this case the secant line will not have a zero between the two bounds.



   (b) Using your understanding of Newton's method (*but language that a high school algebra student would understand*), explain to Terrance why the calculator might need a left bound, right bound, and guess.

   Sample Response:
   When you give a calculator an initial guess, it uses that point to calculate an even closer guess afterwards; it does this by drawing a line that just touches the graph above your initial guess, then looking at where that line crosses the x-axis. The calculator does this many times to create a really good approximation, so the better your initial guess is, the faster and easier it is to get a really close estimate of the zero. Sometimes, even with a good guess, this process can lead to an unexpected zero. So, just to make sure that we don't accidentally find a different zero, the left and right bounds show the calculator where it should look.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Assessment Problem 1

1. Use Newton's method with initial guess $x_0 = 3$ to calculate the first three approximations of a zero of the function $f(x) = x^2 - 5$. Be sure to use at least six decimal places.

   Solution:
   For $f(x) = x^2 - 5$ we know that $f'(x) = 2x$, and so

   $$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 3 - \frac{4}{6} = \frac{7}{3} \approx 2.333333$$
   $$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = \frac{7}{3} - \frac{4/9}{14/3} = \frac{47}{21} \approx 2.23810$$
   $$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = \frac{47}{21} - \frac{4/441}{84/21} = \frac{2207}{987} \approx 2.23607$$

Assessment Problem 2

2. Jack and Raven are working together to find the zeroes of the function $f(x) = x^3 - 3x + 1$ using Newton's method. Jack suggests they begin with an initial guess of $x_0 = 1$. Raven says that won't work.

   (a) Why do you think Raven claims Jack's initial guess will not work? Using tangent lines, explain what Raven may have noticed.

   Solution:
   Raven sees that Jack's guess will not yield results because the tangent line at $x = 1$ is horizontal. Therefore, this tangent line does not cross the x-axis and Newton's method fails.

   (b) Write two questions Raven can ask Jack to help him revise his initial guess. Explain how Raven's questions might help Jack.

   Sample Responses:
   - Raven could ask Jack to draw the tangent line at $x = 1$. This would help Jack visualize that this tangent line does not cross the x-axis, which means they cannot proceed with Newton's method. Raven could then discuss how choosing an initial guess where a local maximum or local minimum occurs causes Newton's method to fail.
   - Raven could also ask Jack why he chose $x = 1$ as an initial guess and if there is a "better" initial guess. This might help Jack realize why $x = 1$ is an unproductive first guess and that he can choose an initial guess that is closer to one of the zeroes of the function.

3.6 References

[1] Conference Board of the Mathematical Sciences (2012). *The mathematical education of teachers II*. American Mathematical Society and Mathematical Association of America.

[2] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from `http://www.corestandards.org/`

3.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** NEWTON'S METHOD (page 1 of 2)

1. Write an equation of **a line** with slope 3 that passes through the point $(2, 1)$ in point-slope form. Then, write an equation of this line in slope-intercept form.

2. Write an equation of the **tangent line** to the graph of $f(x) = x^2 + 2$ at the point $(1, 3)$ in point-slope form. Then, write an equation of this tangent line in slope-intercept form.

3. More with tangent lines.

   (a) For a given function f, describe how to find an equation of the tangent line to the graph of f at $x = a$.

   (b) Now, write an equation of the tangent line to the graph of f at $x = a$.

PRE-ACTIVITY: NEWTON'S METHOD (page 2 of 2)

4. Find the zeroes of the following functions.

   (a) $f(x) = x^2 - 4$

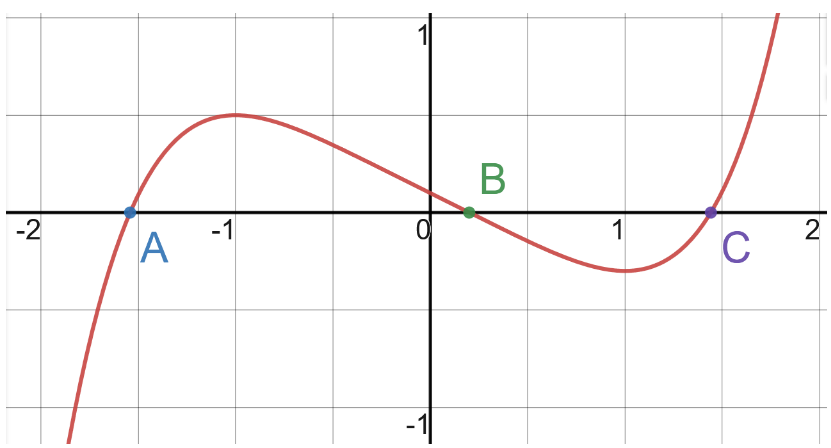
   (b) $g(x) = 3x^2 + 7x - 2$

   (c) $h(x) = x^3 + x^2 - 2x$

5. Consider the function, $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$

   (a) Nnamdi has excellent algebra skills, yet he tries to find the zeroes algebraically and gets stumped. Explain why he is having trouble.

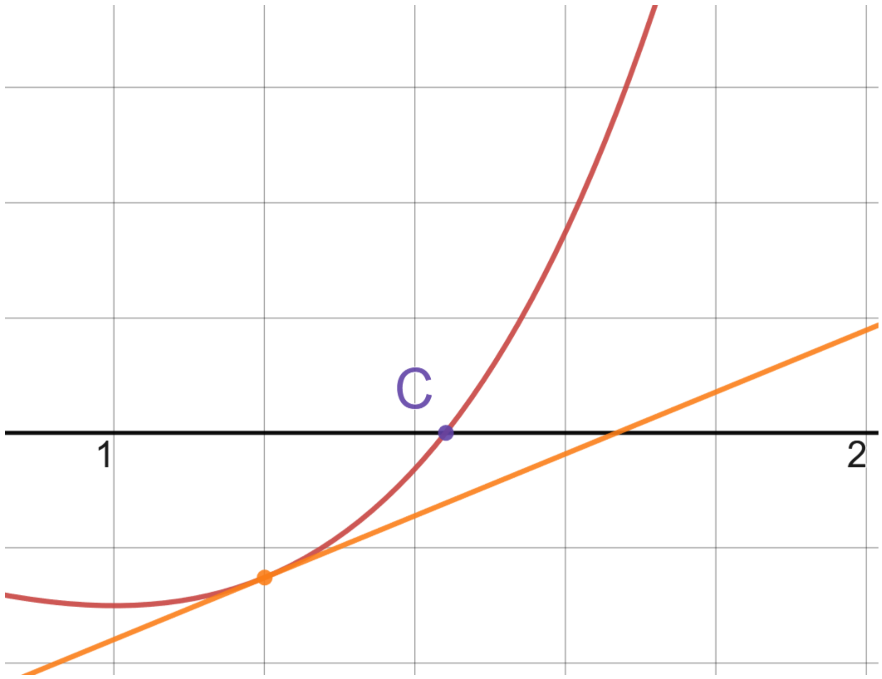
   (b) Nnamdi decides to graph f to find the zeroes. The zeroes are indicated on the graph as A, B, and C. Estimate the value of C.



NAME: **CLASS ACTIVITY:** NEWTON'S METHOD (page 1 of 4)

Recall the function from Problem 5 on the Pre-Activity, $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$, for which Nnamdi wanted to find the zeroes of the function. Nnamdi initially thinks that $x = 1.2$ is a good estimate of the zero, C, but when he zooms in on the graph he realizes that C is further to the right. He starts to experiment with linear functions to try to find a better estimate for C.

1. Nnamdi zooms in on the graph and sketches the tangent line at $x_0 = 1.2$ (see graph below).



   (a) Label the x-intercept of Nnamdi's tangent line as x_1.

   (b) Write an equation of Nnamdi's tangent line in point-slope form and find the value of x_1.

2. Taking inspiration from Nnamdi's idea, Mari decides to sketch another tangent line to the graph of $f(x)$ at the point $(x_1, f(x_1))$. She claims that the x-intercept of her tangent line will be closer to the zero C than x_1.

   (a) Do you agree with Mari's claim? Explain why or why not.

(b) Sketch in Mari's tangent line. Label the x-intercept of her tangent line as x_2.

(c) Write the equation of Mari's tangent line in point-slope form and find the value of x_2.

3. Amy uses both Mari's and Nnamdi's ideas to find a point, x_3, even closer to the zero C.

(a) What do you think she did? Explain.

(b) Find the value of x_3.

4. Fill in the following table with the values of x_1, x_2, and x_3 that you found above. Describe what you notice about these values.

x_0	x_1	x_2	x_3
1.2			

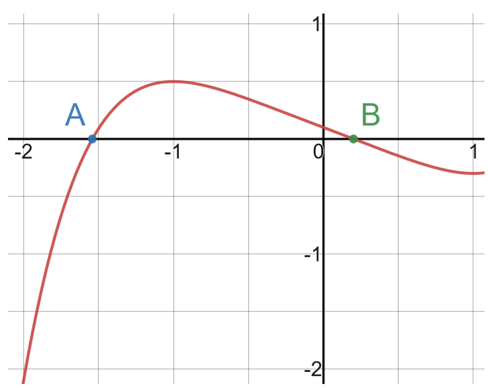
5. The iterative process Amy follows from the work of Mari and Nnamdi is called Newton's method. To apply Newton's method, the process of "finding a tangent line at the point on the graph corresponding to the guess for the zero, finding its x-intercept, and using this x-intercept as the next guess for the zero" is repeated. These x-intercepts (usually denoted x_0, x_1, x_2, x_3, $\ldots$) provide successive approximations of the value of a zero of a function.

   (a) Describe this process graphically.

   (b) Describe this process algebraically. Write out a formula to find x_{n+1}, the x-intercept of the tangent line created from the previous guess, x_n.

   (c) How do you know when to stop this iterative process? That is, when is your approximation of a zero "good enough?"

6. Reconsider $f(x) = \frac{1}{10}x^5 - \frac{1}{2}x + \frac{1}{10}$. Nnamdi now wants to use Newton's method to approximate the zero, A. He wonders what will happen if he uses the following initial guesses: -0.5, -1, -1.5, and -2.



(a) Without doing any calculations, which zero of f do you expect each of these initial guesses to lead? Explain your reasoning. Use the graph above to graphically show (by drawing tangent lines) what happens when you apply Newton's method using these initial guesses.

i. $x_0 = -0.5$

ii. $x_0 = -1$

iii. $x_0 = -1.5$

iv. $x_0 = -2$

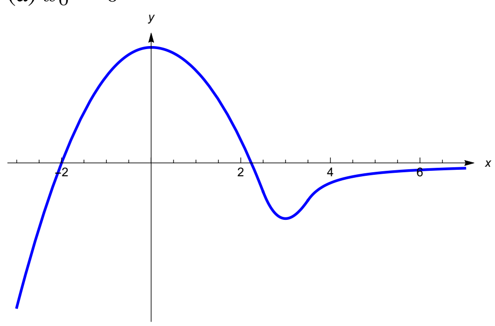
(b) Use Newton's method with all four initial guesses to calculate a zero of f. Give your answer to three decimal places, when applicable.

(c) Summarize to Nnamdi what you observe in the graph of f that indicates what zero you will approximate given your initial guess.

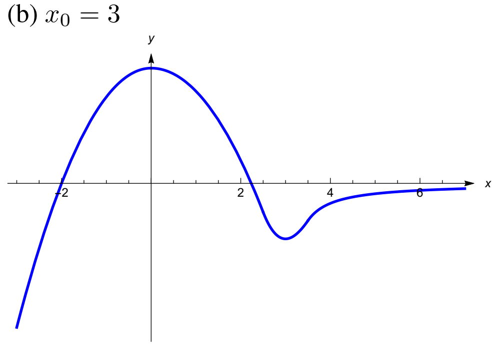
NAME: **HOMEWORK PROBLEMS:** NEWTON'S METHOD (page 1 of 3)

1. The graph of $y = f(x)$ is shown here. Use the initial guesses given to determine which successfully lead to an approximation of a zero of the function f when using Newton's method. For each initial guess, graphically (by drawing tangent lines) support your conclusion based upon using Newton's method and explain your reasoning.

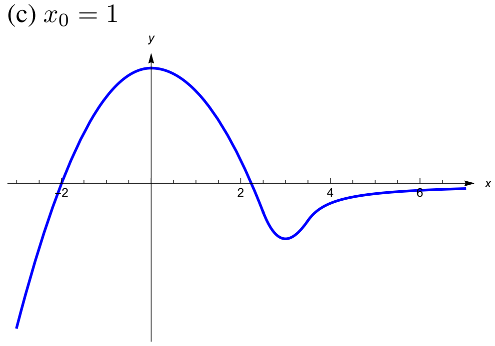
(a) $x_0 = 0$



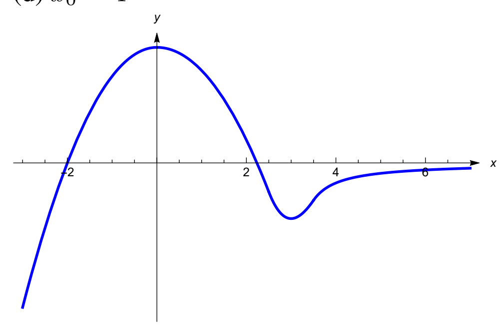
(b) $x_0 = 3$



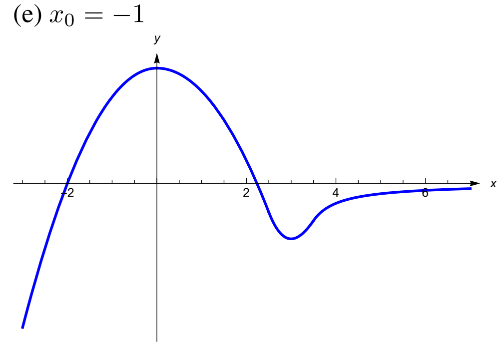
(c) $x_0 = 1$



(d) $x_0 = 4$

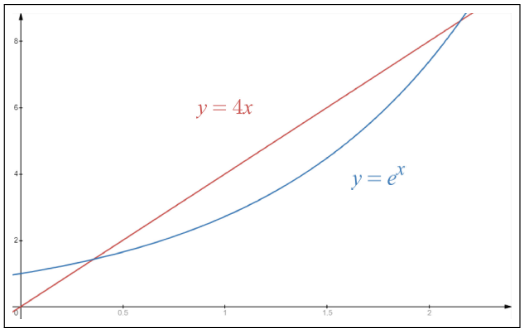


(e) $x_0 = -1$



HOMEWORK PROBLEMS: NEWTON'S METHOD (page 2 of 3)

2. Consider the system of equations given below.



$$y = e^x$$
$$y = 4x$$

   (a) Explain how you could use Newton's method to approximate the two solutions to the system of equations.

   (b) Choose any initial guess and calculate one iteration of Newton's method for each solution. Record your approximations up to six decimal places.

3. Madalena is trying to use Newton's method to find the zeroes of the function $f(x) = x^3 - 2x + 2$. After making sure that $f'(1) \neq 0$, she chooses $x_0 = 1$. Then, she computes the first few iterations of Newton's method and begins a table of values:

x_0	x_1	x_2	x_3
1	0	1	

   Madalena sees that this pattern will continue and comes to you for help.

   (a) Show that Madalena's computations for x_1 and x_2 are correct.

   (b) Create a graph of f and sketch tangent lines to explain why Newton's method has failed.

   (c) Consider the following questions that you might ask Madalena:

      i. Explain how the question below might help you assess what Madalena understands about Newton's method:

      Given a graph of f and an initial guess x_0, how could you find x_1 without making the calculations you have already tried?

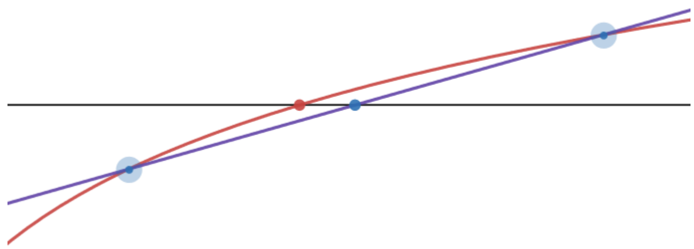
      ii. Explain how following up with the next question might help Madalena to advance in her understanding of Newton's method:

      Considering your graphical explanation of Newton's method, how could two iterations of Newton's method have the same value?

      iii. Explain why the question below might not help Madalena:

      What is the formula for Newton's method?

4. Terrance, a high school algebra student, is using his TI-84 graphing calculator to find the zero of a function. To do so, the calculator requires him to choose a left bound (a point on the graph to the left of the zero), a right bound (a point on the graph to the right of the zero), and a guess (a point on the graph very close to the zero). Terrance thinks that the calculator is using the bounds he has chosen to construct a secant line, which it uses to approximate the zero. He draws the following example to illustrate his idea.



   (a) Under what circumstances would Terrance's "secant method" fail to approximate a zero? Create a graph of one such example.

   (b) Using your understanding of Newton's method (*but language that a high school algebra student would understand*), explain to Terrance why the calculator might need a left bound, right bound, and guess.

NAME: **ASSESSMENT PROBLEMS:** NEWTON'S METHOD (page 1 of 1)

1. Use Newton's method with initial guess $x_0 = 3$ to calculate the first three approximations of a zero of the function $f(x) = x^2 - 5$. Be sure to use at least six decimal places.

2. Jack and Raven are working together to find the zeroes of the function $f(x) = x^3 - 3x + 1$ using Newton's method. Jack suggests they begin with an initial guess of $x_0 = 1$. Raven says that won't work.

   (a) Why do you think Raven claims Jack's initial guess will not work? Using tangent lines, explain what Raven may have noticed.

   (b) Write two questions Raven can ask Jack to help him revise his initial guess. Explain how Raven's questions might help Jack.

4

Variability: Mean Absolute Deviation and Standard Deviation

Introduction to Statistics

Elizabeth G. Arnold, *Colorado State University*
Elizabeth W. Fulton, *Montana State University*
Katharine M. Banner, *Montana State University*
Rachel Tremaine, *Colorado State University*

4.1 Overview and Outline of Lesson

Variability in data is a central concept in statistics, addressed throughout middle school, high school, and college statistics curricula. This lesson focuses on how two measures of variability—mean absolute deviation (MAD) and standard deviation (SD)—can be used to quantify the variation in a dataset. Both MAD and SD measure variability in terms of "average distance from the mean," although they use different methods to calculate "average distance." This lesson highlights how SD builds upon the understanding of variability students are introduced to in middle school with MAD. Understanding how to measure variation is an essential component of making statistical inferences, and statistical inference plays a large role in any undergraduate introduction to statistics course. Undergraduates have likely had different experiences learning about variability in their K–12 schooling, and this lesson respects what undergraduates may have previously encountered and presents a unified framework for developing a deeper understanding of "average distance from the mean" as a measure of variability.

1. Launch—Pre-Activity
   Prior to the lesson, undergraduates complete a Pre-Activity which introduces the importance of describing variability in data. In the Pre-Activity, undergraduates examine three dotplots to compare the center and variability of three different datasets. Instructors can launch the lesson by reviewing the solutions to the Pre-Activity.

2. Explore—Class Activity

   - *Problem 1*:
     Undergraduates quickly play a memory game online and record their time in order to generate a class dataset. Undergraduates then create a graphical summary of the data and describe what they notice about the data. The context of the memory game is used throughout the Class Activity.
   - *Problems 2 & 3*:
     Undergraduates analyze hypothetical student work to make sense of how to compute and interpret the mean absolute deviation of a dataset. The instructor provides a brief discussion to formally define and interpret the mean absolute deviation.

- *Problem 4*:
  Undergraduates analyze hypothetical student work to make sense of how to compute and interpret the standard deviation of a dataset. The instructor provides a brief discussion to formally define and interpret the standard deviation.
- *Problem 5*:
  Undergraduates return to their class dataset from Problem 1 and compute and interpret the mean absolute deviation and standard deviation of their data.

3. Closure—Wrap-Up
   The instructor concludes the lesson by revisiting the calculations and interpretations of mean absolute deviation and standard deviation and discussing how the concept of standard deviation compares to and builds on the concept of mean absolute deviation.

4.2 Alignment with College Curriculum

Given that variability plays a central role in teaching and learning statistics, a deep focus on this topic at the collegiate level will serve all undergraduates, including prospective teachers, exceptionally well. This lesson offers undergraduates an opportunity to develop an understanding of the concept, "average distance from the mean" by examining two different measures of variability (mean absolute deviation and standard deviation). The lesson fits well after instructors have taught different ways to visualize univariate, quantitative data and have discussed different measures of center (such as mean). Because this lesson only focuses on two measures of variability, instructors may wish to implement another lesson to address other measures of variability, such as range and interquartile range.

4.3 Links to School Mathematics

Statistics content standards are integrated throughout middle school and high school mathematics curricula, and it is important for prospective teachers to understand how and why variability is fundamental to teaching and learning statistics. By studying connections between the mean absolute deviation and the standard deviation, prospective teachers will develop a deeper understanding of variability and why the study of mean absolute deviation in middle school serves as a precursor to the study of standard deviation in high school.

This lesson highlights:

- Computing and interpreting the mean absolute deviation and the standard deviation of a dataset;
- Connections between the mean absolute deviation and the standard deviation.

This lesson addresses several statistical knowledge and mathematical practice expectations in common high school standards documents, such as the Common Core State Standards for Mathematics (CCSSM, 2010). Middle school students are expected to understand mean absolute deviation as one way to quantify the amount of variability present in data. Middle school students learn how to compute and interpret the mean absolute deviation of a dataset in the context of a problem (c.f. CCSS.MATH.CONTENT.6.SP.B.5.C). High school students are expected to build on their understanding of mean absolute deviation to develop an understanding of how to compute and interpret the standard deviation of a dataset, which is a more common measure of variability used in practice (c.f. CCSS.MATH.CONTENT.HSS.ID.A.2 and CCSS.MATH.CONTENT.HSS.ID.A.3). This lesson also provides opportunities for prospective teachers to think about the reasoning of others and construct sound statistical arguments.

4.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- Measures of central tendency (mean, median, mode);
- How to visualize univariate, quantitative data with dotplots, boxplots, and histograms.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Compute the mean absolute deviation and the standard deviation of a dataset;
- Interpret the mean absolute deviation and the standard deviation of a dataset in the context of a problem;
- Describe how the concept of standard deviation builds on the concept of mean absolute deviation;
- Examine hypothetical student work to make sense of "average distance from the mean";
- Evaluate and pose questions to help guide students' understanding about mean absolute deviation and standard deviation.

Anticipated Length

Two 50-minute class sessions.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity (print Problems 1, 2–3, 4, and 5 to pass out separately)
  - Computer/Tablet/Phone (for undergraduates) to play the memory game at the beginning of the Class Activity
  - Computer (for instructor) to compile the class dataset and create a graphical summary of the data during the Class Activity
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

4.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework for undergraduates to complete in preparation for the lesson, and ask undergraduates to bring their solutions to class on the day you start the Class Activity. At your discretion, allow undergraduates to use technology to compute the mean number of pets for each class in part (a).

> *The goal of the Pre-Activity is to introduce the importance of accounting for variability in data, so we purposely constructed each dataset to have the same mean number of pets (i.e., 1.5 pets) but the variability (and the distributions) are different.*

Pre-Activity Review (10 minutes)

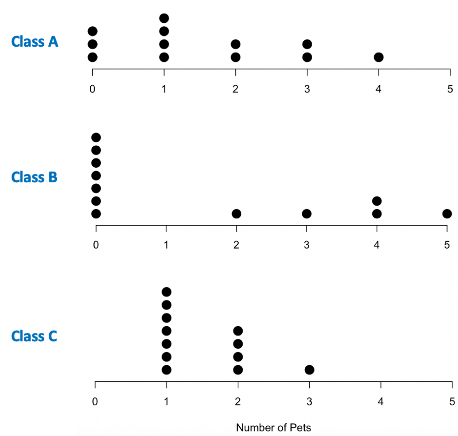
As a class, review the solutions to the Pre-Activity. Focus most of the discussion on part (c), define variability (e.g., "variability, sometimes referred to as spread, is commonly described by specifying how far the data are from a measure of center"), and discuss the following connection to teaching. Undergraduates' answers in the Pre-Activity may vary but the main point is for them to recognize that the amount of variability present in each distribution is different.

Discuss This Connection to Teaching

“Learning from data and making informed choices on the basis of data depend on understanding and describing variability” (Peck et al., 2013, p. 25). It is important for prospective teachers to understand that finding the mean of a dataset is often not enough to summarize quantitative data or characterize a distribution. In K–12 mathematics, students will be asked to compare distributions and they will need to consider the shape, center, and spread of the distributions.

Pre-Activity

1. Students from three different classes reported the number of pets in their household. The results are summarized graphically as dotplots and in a frequency table below.



Class A	Class B	Class C
1	0	1
0	0	1
4	0	1
3	4	1
2	4	2
1	3	1
1	0	2
3	5	2
0	0	3
2	0	2
1	0	1
0	2	1

(a) Compute the mean number of pets for each class.

Solution:

For Classes A, B, and C: $\bar{x}_A = \bar{x}_B = \bar{x}_C = 1.5$ pets.

(b) What is similar about the three dotplots?

Sample Responses:

- The mean number of pets in each class is the same.
- There are 12 dots in each dotplot.
- All have the same mean of 1.5 pets.

(c) What is different about the three dotplots?

Sample Responses:

- The shape of each dotplot is different.
- The medians are different. Class A and Class C have a median of 1 pet but Class B has a median of 0 pets.
- The modes are different. Most of the students in Class B have 0 pets and most of the students in Class A and Class C have 1 pet.
- The spread of each dotplot is different.
- The standard deviation of each dotplot is different. $s_A = 1.31$ pets, $s_B = 1.98$ pets, and $s_C = 0.67$ pets.

Conclude the Pre-Activity Review by discussing the following connection to teaching and letting undergraduates know that the purpose of this lesson is to explore two different measures of variability: mean absolute deviation (MAD) and standard deviation (SD).

Discuss This Connection to Teaching

National and state standards now emphasize statistics content standards in middle school and high school with the central theme of variability being developed throughout. The Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education (GAISE II) Report states that "Statistical thinking, in large part, must deal with the omnipresence of variability in data (e.g., variability within a group, variability between groups, sample-to-sample variability in a statistic). Statistical problem solving and decision making depend on understanding, explaining, and quantifying variability in the data within the given context" (Bargagliotti et al., 2020, p. 7). Because the concept of variability is foundational to teaching and learning statistics, undergraduates are expected to develop a deep understanding of this concept, and this is especially true for prospective teachers who will teach their students statistics.

Class Activity: Problem 1 (10 minutes)

Pass out **Problem 1** of the Class Activity. Ask each undergraduate to access the *Census at School* webpage and play the *Memory Game* (`http://ww2.amstat.org/education/cas/1.cfm`). Explain to the undergraduates why you are using a game from this particular website and its connection to teaching.

Discuss This Connection to Teaching

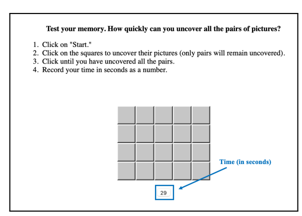
The *Census at School* site is an international resource that engages students in statistical problem solving. It is a useful resource for prospective teachers to be aware of and use in their future classrooms for class activities and projects. The site also serves as a good source for obtaining authentic data for statistical investigations.

Give undergraduates about a minute to play the game. Remind them to record their time (in seconds) from their **first attempt only**. As undergraduates play the game, set up a way to compile, share, and present the class dataset of memory game times (e.g., GoogleSheets, StatKey, etc.) so that you (or your undergraduates) can create a graphical summary of the data.

Class Activity Problem 1 : Part a

1. **Test Your Memory!**

   Play the *Census at School Memory Game* where you will need to uncover and match 10 pairs of pictures. The time it takes you to complete the game will be tracked. Go to the following link to access the game:
   `https://ww2.amstat.org/education/cas/1.cfm`



   (a) Play the game once and record your time (in seconds) as a number.

   Solution:
   Answers will vary. Most undergraduates complete the game in less than a minute.
   Commentary:
   If undergraduates play the game more than once, consider having a class discussion about why, from a statistical perspective, we should only use the time from their first attempt of the game. This may include the following ideas:

- You may get quicker at completing the game by playing it more than once, so using times from everyone's first attempt is more standardized among everyone.
- If we gave everyone 5 minutes to play the game and used all of the attempts, then those who were faster at completing the game would have more of their times recorded in the dataset. This would result in repeated measures (i.e., non independent observations from the same person), which might make the measure of variability smaller than what it is expected.

Complete Problems 1(b) and 1(c) as a class. Start by asking undergraduates to identify and explain what type of graph would be appropriate to summarize the class data. Next, create the type of graph recommended by undergraduates, display it to the class, and ask them to sketch the graph on their Class Activity handout for Problem 1(c).

Class Activity Problem 1 : Parts b & c

(b) As a class, compile everyone's time in a dataset. What graphical summary would be appropriate to visualize the class's distribution of times on the memory game? Explain your reasoning.

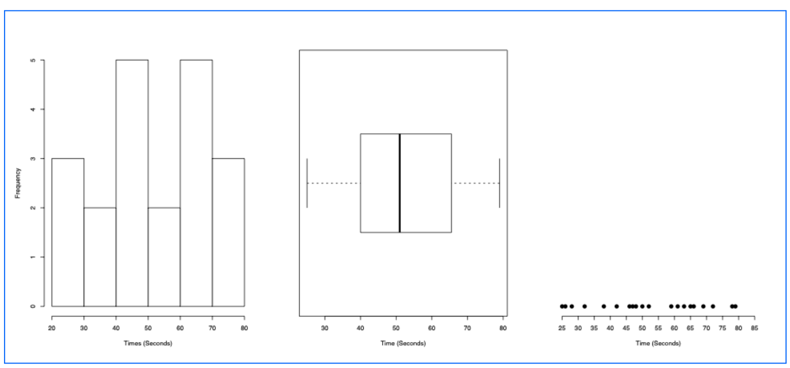
Sample Responses:

- A histogram because the times are quantitative.
- We can make a dotplot of the memory game times since the times are quantitative.
- I would make a boxplot because we have a single dataset of quantitative values.
- A stem-and-leaf plot would appropriately display the class's memory game times.

Commentary:
If all undergraduates suggest the same type of plot, ask them to think of other appropriate ways to visualize the data. If undergraduates suggest inappropriate types of graphs (such as a scatterplot), discuss why they are not appropriate to visualize these data.

(c) As a class, create a graphical summary to visualize the class's distribution of times on the memory game and sketch it below.

Sample Graphical Summaries:



Commentary:
Dotplots and histograms are the primary types of graphs used throughout the lesson so it may be worthwhile to quickly create both of these types of plots. If you create a dotplot, ask undergraduates, *What does one dot on this dotplot represent?* and lead them to an appropriate understanding (e.g., "one dot represents how long a single student in the class took to successfully complete the memory game.").

Finish Problem 1 by asking undergraduates to document what they notice about the class's distribution of memory game times, based on the graphical summary, in Problem 1(d).

Class Activity Problem 1 : Part d

(d) Describe what you notice about the class's distribution of times on the memory game.

Sample Responses:
Undergraduates may describe features about

- The center, shape, and spread of the distribution.
- The mode(s) of the distribution.
- Any noticeable outliers.

Class Activity: Problems 2 & 3 (30 minutes)

Pass out **Problems 2 and 3** of the Class Activity, which focus on computing and interpreting the MAD of a dataset. Undergraduates will work on these two problems in groups. (See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion.) Before they begin working, discuss the following connection to teaching.

Discuss This Connection to Teaching

These next several problems in the Class Activity focus on analyzing hypothetical students' thinking in order to develop undergraduates' skills in understanding school student thinking and developing questions to guide school students' understanding. It is important for undergraduates, especially prospective teachers, to understand how others use, reason with, and communicate statistics. These problems also give prospective teachers (and tutors and future graduate students) an opportunity to think about how they would respond to student work in ways that nurture students' assets and understanding and in ways that help develop students' statistical understanding.

Emphasize to undergraduates that the hypothetical students' work they will be examining is procedurally correct so they do not need to be looking for any arithmetic errors. Rather, encourage undergraduates to focus on the reasoning behind the calculations that Jasmine is using to compute the MAD. We intend for Problem 2 to be an informal introduction to MAD, and you will formally define it after undergraduates complete this problem. From previous implementations of this lesson, we have noticed that some undergraduates are familiar with MAD, but most do not recall learning MAD prior to this lesson.

In Problem 2, Jasmine drew a histogram and is computing the MAD. She first computes the mean. Then she computes the absolute value of the differences between each data point to the mean (i.e., the distances). Finally, she computes the average of those distances to compute the MAD.

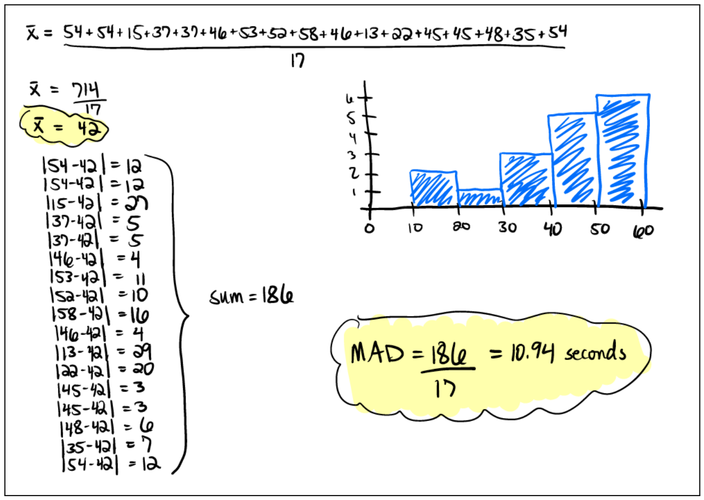
Class Activity Problem 2

2. **Quantifying Variability with Mean Absolute Deviation**

Amaury is teaching a high school intermediate algebra class and his students are learning about different measures of variability. The students in his class played the *Census at School Memory Game* and recorded their times, in seconds:

54, 54, 15, 37, 37, 46, 53, 52, 58, 46, 13, 22, 45, 45, 48, 35, 54

Amaury asked his students to create a graphical summary of the data, compute the mean, and quantify the amount of variability present. One of his students, Jasmine, did the following:



Jasmine, recalling what they learned in middle school, quantified the amount of variability by computing the **mean absolute deviation (MAD)**. All of their calculations are correct. Describe mathematically what Jasmine did to compute the MAD.

Sample Responses:

- Jasmine subtracted the value of each data point from the mean and took the positive value of that result by taking the absolute value. Then they averaged all of these values.
- Jasmine found the mean (42) and calculated the distance each data value was from the mean. Jasmine added all those distances and then divided by 17 (# of data values).
- Jasmine first got the mean of the data by adding all the numbers together, then dividing by the total number. In order to get the mean absolute deviation, Jasmine used absolute value and subtraction in order to find the distance from each data point and the mean. They then computed the average of those distances.

Commentary:
If undergraduates have difficulty recognizing how deviations from the mean are used in Jasmine's calculation, focus their attention to these deviations and how absolute value is used.

Discussion: Mean Absolute Deviation (MAD)
Formally define MAD, discuss how it is calculated, and emphasize the following connection to teaching. At your discretion, show undergraduates how to compute the MAD with technology. Note that R has an `MAD` command, but this computes the median absolute deviation.

To compute the mean absolute deviation (in R) of a dataset stored as the object "x", you can use the following code:

```
meanAD <- function(x){
  avg <- mean(x)
  mad <- mean(abs(x) - avg)
  return(mad)
}
```

Example:

```
x <- c(0.84, -0.18, -1.28, -1.07, -0.05)
meanAD(x)
# [1] 1.032
```

Discuss This Connection to Teaching

MAD is a measure of variability typically taught to middle school students and it is a precursor to the study of standard deviation in high school.

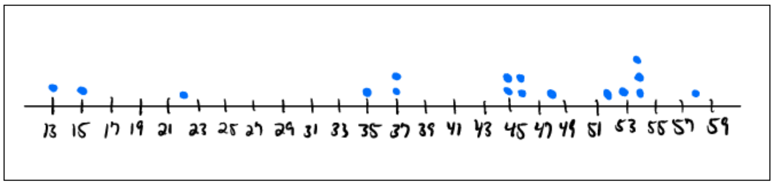
Definition:

- MAD stands for the **mean absolute deviation** (not "mean average deviation" or "median absolute deviation").
- Informally, the MAD measures how spread out the data are and provides a numerical value to quantify the amount of variability present in data.
- Specifically, the MAD measures the average distance the data values are from the mean.

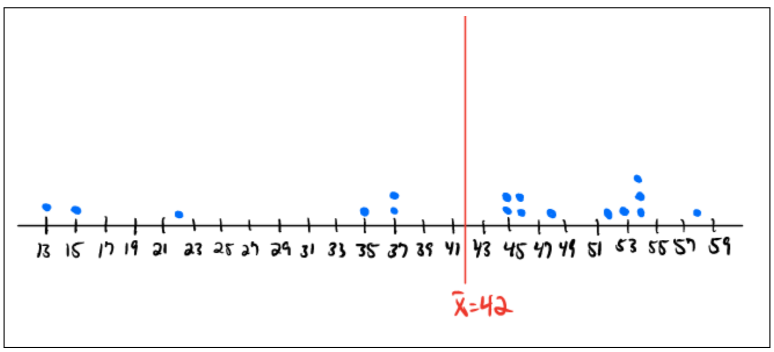
Conceptual Understanding:

Construct the following sequences of dotplots to help undergraduates conceptually understand MAD; this includes an understanding of how deviations from the mean are treated in the formula for MAD. Note that one of the Assessment Problems includes a dotplot with deviations drawn in.

1. Draw a dotplot of the high school class's data.

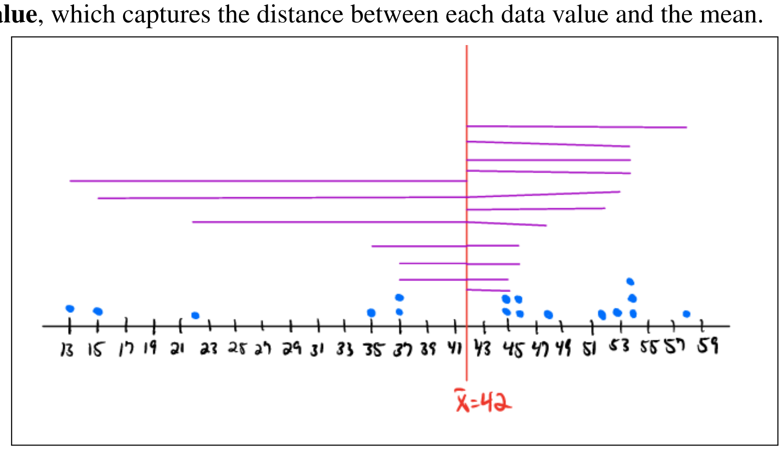


2. Indicate the mean (42 seconds) by drawing a vertical line through 42.



3. Ask undergraduates how far each data point is from the mean and then draw in each deviation from the mean using horizontal lines. Label them as indicated by the undergraduates.

   Some undergraduates may give negative values for points to the left of the mean and positive values for points to the right of the mean. If this occurs, ask undergraduates what happens when we sum those positive and negative values. Here, they will notice the sum will be zero which may help them understand why the formula for MAD uses **absolute value**, which captures the distance between each data value and the mean.



Emphasize that to compute the MAD, we find the distance between each data point and the mean, and then take the average of those distances.

Thus, MAD represents the "average distance from the mean." A smaller MAD generally indicates that the data are closer to the mean. A larger MAD generally indicates that the data are farther from the mean.

Formula:

- The formula for MAD is given by:

$$\text{MAD} = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

Problem 3 focuses on interpreting the MAD in the context of a problem, which can be challenging for undergraduates. This problem is intended to guide undergraduates in developing their knowledge about what a correct (and complete) interpretation of MAD includes. All three of the interpretations in Problem 3 are real examples from students. Tarryn's interpretation is correct. Jasmine's and Benny's interpretations are a good start, but they are incomplete and need to be expanded upon.

Class Activity Problem 3 : Parts a & b

3. **Interpreting Mean Absolute Deviation**

Two other students, Tarryn and Benny, also correctly computed the MAD. When Amaury asked his students to write a sentence interpreting their measure of variability in the context of the problem, Jasmine, Tarryn, and Benny wrote the following sentences:

Jasmine	The MAD is 10.94 seconds.
Tarryn	On average, the memory game times were 10.94 seconds away from the mean of 42 seconds.
Benny	A memory game time is 10.94 from the mean.

(a) One student correctly (and completely) interpreted the MAD in the context of the problem. Identify who it was, and describe what components of their interpretation make it correct and complete.

Sample Response:
Tarryn's interpretation is correct. She focuses on "average distance from the mean" in her interpretation and includes units (seconds) in her response.

Commentary:
From our experience, most undergraduates correctly identify Tarryn's interpretation as correct. Take time to highlight why her interpretation is correct, focusing on accounting for "average distance from the mean" in the interpretation.

(b) The other two students gave incomplete interpretations of the MAD. Based on their interpretations, describe what each

i. may understand about interpreting the MAD, and

ii. may not yet understand about interpreting the MAD.

Sample Responses:

- Jasmine
  - Jasmine understands that the MAD has a unit (seconds).
  - They may not fully understand what MAD measures because they did not write out what MAD measures in the context of the problem.

- Benny
  - Benny understands that MAD is a measure of distance from the mean.
  - He doesn't yet understand that MAD is an average distance from the mean.

Commentary:
In previous implementations of this lesson, we have noticed that undergraduates tend to focus only on what the hypothetical students do not yet understand. Make sure that undergraduates also attend to what the students do understand.

After discussing the solutions to Problems 3(a) and 3(b), emphasize the following connection to teaching:

Discuss This Connection to Teaching

The practice of evaluating student work requires knowledge to assess what a student does and does not yet understand, and prospective teachers need practice with the skill of nurturing students' assets and understanding in the classroom. Addressing different perspectives in a manner that conveys respect for student thinking and reasoning, both when a student's work is correct and when it is incorrect, is a critical practice for teachers.

Undergraduate responses to Problem 3(c) will vary. A key idea to include in an answer is that MAD is a measure of variability that describes the average distance from the mean.

Class Activity Problem 3 : Part c

(c) In a general context, describe what MAD measures.

Sample Response:
MAD tells us, on average, how far the data are from their sample mean. The higher the number, the more spread out the data are.

After discussing the solutions to Problem 3, emphasize the importance of statistical interpretations in this connection to teaching.

Discuss This Connection to Teaching

Part of the statistical problem-solving process involves interpreting results, and interpretations should integrate the context of the problem. Statistical thinking is different from mathematical thinking because of the omnipresence of variability and because of the central role context plays in understanding and interpreting data. All undergraduates benefit from having to interpret statistical concepts in the context of the problem. This skill is especially important for prospective teachers who will teach their students how to compute and interpret various measures of variability.

Advice on Delivering the Lesson Over Two Class Sessions

If you are teaching this lesson over two class sessions, stopping around Problem 3 is a good place. See Chapter 1 for guidance on using exit tickets to facilitate instruction in a two-day lesson.

Class Activity: Problem 4 (20-25 minutes)

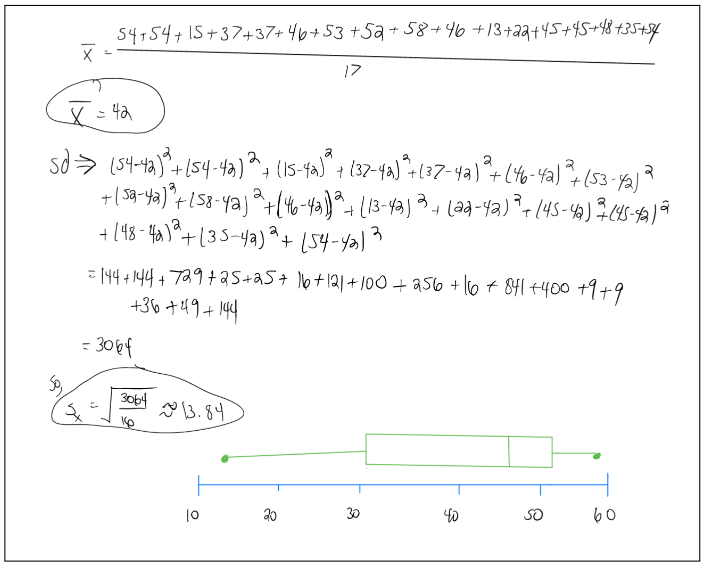
Pass out **Problem 4** of the Class Activity, which focuses on calculating and interpreting the standard deviation (SD) of a dataset. Instruct undergraduates to work in small groups on Problem 4 and let them know they will continue to examine hypothetical students' work. For Problem 4(a), emphasize that the student's work is procedurally correct so they do not need to be looking for any arithmetic errors. Rather, encourage undergraduates to focus on the reasoning behind the calculations that Josief is using to compute the SD. We intend for Problem 4(a) to be an informal introduction to SD, and you will formally define it after undergraduates complete this problem.

In Problem 4, Josief drew a boxplot and is computing the standard deviation (SD). He is finding deviations from the mean and squaring them. He adds these values, divides the sum by 1 less than the total number of data values, and concludes by taking the square root to compute the SD.

Class Activity Problem 4 : Part a

4. **Quantifying Variability with Standard Deviation and Interpreting Standard Deviation**

   Josief, another student in Amaury's class, did something different, as shown below:



(a) Josief quantified the amount of variability by computing the **standard deviation (SD)**. All of his calculations are correct. Describe mathematically what he did to compute the SD.

Sample Responses:

- Josief first found the mean. Then he computed the distance from each data point and the mean, squared those distances, and then added all of those numbers. With the sum of those numbers, he then divided by the amount of numbers minus one and then took the square root.
- To calculate standard deviation, calculate the difference between each value and the mean, square those differences and add them together, divide by one less than the total number of data points, and take the square root of that value.
- Josief computed the mean. Then he calculated $(x_i - \bar{x})^2$ for each data value, added them up, and divided by 16. Finally, he took the square root of that number to get the SD.
- He first found the mean. Then subtracted the mean from each data point and squared those values. He divided the sum of those values by 16 and took the square root.

Commentary:

If undergraduates have difficulty recognizing how deviations from the mean are used in Josief's calculation, focus their attention to these deviations and how they are squared.

Discussion: Standard Deviation (SD)

Formally define SD, discuss how it is calculated, and discuss the following connection to teaching. At your discretion, show undergraduates how to compute the SD of a dataset with technology.

Discuss This Connection to Teaching

SD is a measure of variability taught to high school students, and it builds on the concept of MAD. Both MAD and SD measure the same concept but they are calculated differently. SD is one of the more common measures of variability used in practice.

Definition:

- SD represents the **standard deviation** of a dataset.
- Informally, the SD measures how spread out the data are and provides a numerical value to quantify the amount of variability present in data.
- Specifically, the SD measures the average distance the data values are from the mean.

Formula:

- The formula for SD is given by:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

- To compute the SD, you first compute the mean of the dataset. Then, you compute the deviations from the mean by subtracting the mean from each data value. You will square all of those values, add them up, and divide the sum by one less than the number of total data values. Finally, you take the square root of that number to compute the SD.

> *Some undergraduates may ask about taking the square root and dividing by $n-1$ instead of n when computing the SD. Discuss this idea as you normally would in your class (e.g., The reason for using $n-1$ instead of n is beyond the scope of this course, but is related to statistical theory. In short, the sampling distribution for s^2 can be approximated by a known distribution if we divide by $n-1$ instead of n (as long as n is large), which makes statistical inference more straightforward for many applications.).*

We have found that undergraduates find it challenging to interpret the SD in the context of a problem. Problem 4(b) focuses on guiding undergraduates in developing their knowledge about what a correct (and complete) interpretation of SD includes. Problem 4(c) focuses on giving undergraduates an opportunity to think about how they would respond to student work in ways that help the student develop an understanding of interpreting SD in the context of a problem.

Class Activity Problem 4 : Parts b & c

(b) When asked to write a sentence to interpret the SD in the context of the problem, Josief wrote the following:

> The memory game times varied by 13.84 seconds.

Describe why Josief's interpretation is not completely correct.

Sample Response:
Josief is not attending to the fact that standard deviation measures an average distance from the mean.

Commentary:
The main point is that Josief's interpretation is missing the "average from the mean" aspect. More specifically, not all memory game times will vary exactly 13.84 seconds from the mean but that "on average" the times were about 13.84 seconds faster or longer than the average time of 42 seconds.

(c) Consider the following questions that one might ask Josief to help him with his interpretation of the standard deviation (in the context of the memory game times).

i. Explain how the following question might help Josief to advance in his understanding of interpreting standard deviation in the context of a problem:

Can you say more about how the memory game times varied?

Sample Response:
Asking Josief to describe how the memory game times varied will push him to explain what he means by "varied" in his interpretation. This may help him understand that standard deviation is one particular measure of variability that measures average distance to the mean, but its interpretation differs from other measures of variability. For instance, the range also describes how the data vary but in a way that is different from how the standard deviation describes variability.

ii. Explain how the following question might help you assess what Josief understands about interpreting the standard deviation in the context of a problem:

What does standard deviation measure?

Sample Response:
Asking this will help me know if he wrote the interpretation based on his understanding of what standard deviation measures or if he fully understands what standard deviation measures and is having difficulty translating that interpretation to the context of the problem.

iii. Explain why the following question might not help Josief:

Why is your interpretation incorrect?

Sample Response:
Josief probably thinks his answer is correct, so asking this question isn't helpful.

Undergraduate responses to Problem 4(d) will vary. A key idea to include in an answer is that SD is a measure of variability that describes the average distance from the mean.

Class Activity Problem 4 : Part d

(d) In a general context, describe what SD measures.

Sample Response:
SD tells us how far the data points deviate from their sample mean, on average.

Class Activity: Problem 5 (10 minutes)

Pass out **Problem 5** of the Class Activity, which asks undergraduates to return to the class's dataset from Problem 1 so they can compute the MAD and SD of that dataset. At your discretion, allow undergraduate to use technology for these computations. Focus on undergraduates' interpretations, making sure they are complete and correct. If you have run out of time, it may be helpful to assign this problem as homework.

Class Activity Problem 5

5. Return to the class dataset from Problem 1.

(a) Compute the mean absolute deviation of your class's data of memory game times and write a sentence interpreting the mean absolute deviation in the context of the problem.

Sample Response:
The average distance from the mean time of [*value of mean*] seconds is [*value of MAD*] seconds.

(b) Compute the standard deviation of your class's data of memory game times and write a sentence interpreting the standard deviation in the context of the problem.

Sample Response:
The average distance from the mean time of [*value of mean*] seconds is [*value of SD*] seconds.

Wrap-Up (15 minutes)

Conclude the lesson by revisiting the computations and interpretations of MAD and SD, discussing how the concept of SD compares to and builds on the concept of MAD. Further, it may be appropriate to discuss how MAD and SD are sensitive to outliers and why SD is used more often in practice. See below for specific considerations.

Importance of Computing and Interpreting MAD and SD:

- Variability is central to teaching and learning statistics because it is a foundational concept for understanding sampling distributions and sampling variability/error (i.e., the SD of the sampling distribution which is discussed later in the semester).
- In this lesson we focused on two different measures of variability, MAD and SD, both of which rely on deviations from the mean, and quantify how far, on average, data values are from the mean.
  - MAD uses the absolute value of the deviations and SD uses the squares of the deviations, both resulting in positive values.
- We also focused on interpreting statistics, a crucial part of the statistical problem-solving process.

MAD and SD in Middle School, High School, and College:

- MAD is a measure of variability taught in middle school while standard deviation is a measure of variability taught in high school and college.
- To conduct statistical inference, we need to quantify the background variability in the estimator of a population parameter of interest. In this lesson, undergraduates developed a deeper understanding of how two measures of variability can be used to quantify the variation in a dataset. The role of standard deviation in doing statistical inference will be present throughout the rest of an undergraduate introduction to statistics course, particularly when undergraduates quantify the standard deviation of a sampling distribution to find a margin of error and build a confidence interval.

How SD Builds on MAD:

- There's a reason MAD is a precursor to SD in school mathematics. MAD is easier to compute, and the concept of SD builds on the understanding students have of MAD.
- After students learn to compute and interpret the MAD of a dataset, they will learn that SD has a similar interpretation and a slightly more complex computation. Both MAD and SD have the same units and measure "average distance from the mean."
- The MAD is a good introduction to the concept of measuring "average distance from the mean" and more intuitive to understand compared to the SD because the computation involves taking a "true" average and the concept of average is taught in middle school.
  - MAD is a considered a more "true" average since it's calculated with the formula $\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$ (i.e., a sum divided by the total number of values), whereas the SD is calculated using $\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ (i.e., a sum divided by one less than the total number of values).

Sensitivity to Outliers:

- MAD and SD are both sensitive to outliers and are not robust measures of spread.

  - MAD and SD are sensitive to outliers because of the way they are calculated. MAD takes the absolute value of deviations from the mean and the SD squares those deviations. The SD then sums all of the individual squared deviations, divides by the number of observations minus one and then takes the square root of that quantity. This process makes large deviations from the mean have a greater influence in the calculation of SD compared to what happens in the calculation of MAD, so generally SD is a larger number than the MAD.

- MAD/SD measure deviations from the mean so it makes sense to report a SD/MAD when a mean is reported as a measure of center, and these measures are more appropriate when a distribution is unimodal and symmetric. When data contain an outlier, other measures of variability, such as IQR, are more appropriate to quantify the amount of variability present.

Discuss This Connection to Teaching

High school students are asked to relate the choice of measures of center and spread to the shape of the data distribution. Depending on the shape of the distribution, appropriate descriptors of center and spread may change. Prospective teachers should understand how different measures of variability correspond to different measures of center and the shape of a distribution.

Why SD is Used More in Practice:

- In practice, we often report the SD (more than the MAD) because much of statistical theory is based on using the SD. Undergraduates will learn more about the statistical theory if they continue with their statistics education and take a probability theory and mathematical statistics course sequence.
  - Working with squared deviations (that is, $(x_i - \bar{x})^2$) is computationally easier than working with absolute value deviations (that is, $|x_i - \bar{x}|$). In other words, it's much easier to take the derivative of squares than to take the derivative of an absolute value. For example, taking a derivative is necessary for finding the maximum likelihood estimator (MLE) for the true population variance.
  - SD has a tie to a variance estimator that's important for applied statistics. The variance is the SD squared, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$. When we have a random sample from a normal population (i.e., if $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$), this estimator of the population variance, σ^2, is unbiased (i.e., $E(s^2) = \sigma^2$), and as $n \to \infty$, s^2 of the approximate sampling distribution for the sample variance is known. This makes conducting statistical inference for the population variance mathematically convenient (see Wackerly et al. (2008) for more information).

You can ask undergraduates to complete an exit ticket, if you choose. See Chapter 1 for guidance on using exit tickets.

Homework Problems

At the end of the lesson, assign the following homework problems; we've focused on including problems that highlight connections to teaching. Assign any additional homework problems at your discretion.

Problem 1 prompts undergraduates to compute and interpret the MAD and the SD for a given dataset. At your discretion, allow undergraduates to use technology for these computations.

Homework Problem 1

1. Ten movies were randomly selected and the length of each movie (in minutes) is given below.

   152, 156, 98, 173, 68, 122, 92, 105, 138, 126

   (a) Compute the mean absolute deviation (MAD) and write a sentence that interprets the MAD in the context of the problem.

Solution:
$\bar{x} = (152 + 156 + 98 + 173 + 68 + 122 + 92 + 105 + 138 + 126)/10 = 123.$

Then,

$$\begin{aligned} \text{MAD} = (&|152 - 123| + |156 - 123| + |98 - 123| + |173 - 123| + |68 - 123| + \\ &|122 - 123| + |92 - 123| + |105 - 123| + |138 - 123| + |126 - 123|)/10 \\ = {}& 26 \text{ minutes} \end{aligned}$$

One sample interpretation is: "On average, the length of a movie is 26 minutes away from the mean of 123 minutes."

(b) Compute the standard deviation (SD) and write a sentence that interprets the SD in the context of the problem.

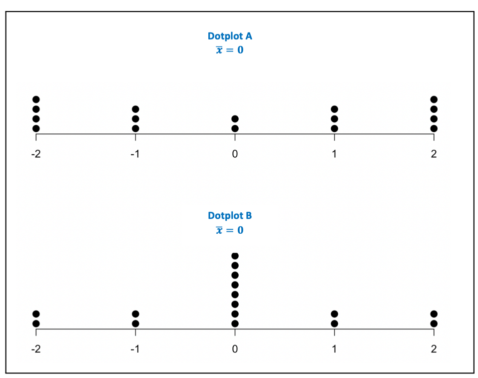
Solution:
The standard deviation (found using technology) is approximately 32.7 minutes. One sample interpretation is: "The difference between a movie length and the mean length of the movies is, on average, about 32.7 minutes."

High school students are asked to compare distributions in terms of their center and spread. It is important for future teachers to be able to visually assess and compare the center and spread of distributions so they can help their students do the same. The inspiration for Problem 2 came from delMas (2001).

Homework Problem 2

2. In this problem, you will *visually* compare the mean absolute deviation (MAD) between pairs of dotplots.

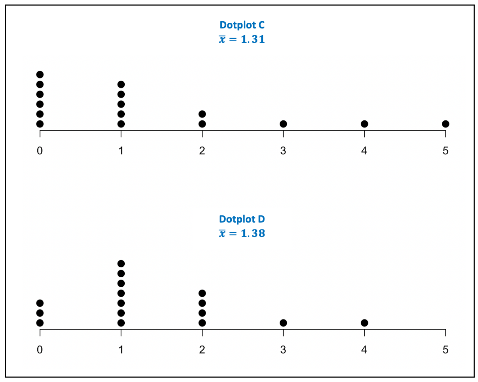
   (a) **Dotplots A and B**.
   Without doing any calculations, identify which dotplot (A or B) has the *larger* MAD. Explain your reasoning.



Sample Responses:
Dotplot A has the larger MAD. Sample responses are provided below.

- Dotplot A has more dots at the extreme ends (e.g., -2 and 2).
- The mean of Dotplot B is closer to its mode compared to the mean of Dotplot A and its mode. Thus, Dotplot B has less variability.

(b) **Dotplots C and D**.
Without doing any calculations, identify which dotplot (C or D) has the *larger* MAD. Explain your reasoning.

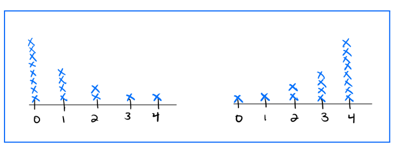


Sample Responses:
Dotplot C has the larger MAD. Sample responses are provided below.

- Dotplots C and D have similar means, but C has more skew than D. This means that Dotplot C has more variability.
- The dots in Dotplot D are more tightly clustered around its mean (compared to those in Dotplot C). Thus, the MAD of Dotplot D is smaller than the MAD of Dotplot C.

(c) Draw two different dotplots that have the same MAD. Describe your thought process when creating these two different dotplots.

Sample Response:



These are two different plots with the same MAD. I wanted to create mirror images that have the same range. Thus, the right dotplot is a reflection of the left dotplot. This makes it so that the distance between a data point and mean in the left dotplot is equivalent to distance between the reflected data point and reflected mean in the right dotplot. Thus, I know that the MAD of each will be the same.

Problem 3 prompts undergraduates to describe why the MAD is a statistical concept taught before the SD in K–12 schooling.

Homework Problem 3

3. Consider the mean absolute deviation (MAD) and the standard deviation (SD). Typically, MAD is first taught in middle school and SD is taught in high school. Describe why it is helpful for school students to learn MAD before SD.

Solution:
Answers will vary; key ideas to include in a correct solution are described below:

- Both the MAD and the SD have similar interpretations and measure, on average, how much quantitative data deviates from the mean. Because the MAD is simpler to compute, students can focus on understanding how to interpret the MAD and transfer that knowledge to interpreting the

SD.

- MAD is a more intuitive measure of spread because it describes the exact average distance each data point deviates from the mean.
- The MAD is easier to calculate than the SD. We don't need to square deviations and then take the square root when computing the MAD.

Sample Response:

- MAD is more obvious and understandable than SD. We have calculated the mean many times and when we do the MAD, the equation $\frac{\sum_{i=1}^{n}|x_i-\bar{x}|}{n}$ tells us exactly what we are doing. We're finding the average of the absolute deviations. The equation makes more sense to what we are finding. SD has additional components we need to comprehend when doing it. The equation $\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ has a square, an $n-1$, and a square root which we have to learn more about to understand. They are similar measures because they measure "average distance from the mean," but it is easier to understand MAD.

Problem 4 has been adapted from a problem created by the *LOCUS Project* (Levels of Conceptual Understanding in Statistics). You can view commentary and correct answers to the original problem at `https://locus.statisticseducation.org/professional-development/questions/analyze-data?type=prodev_multiple_choice_question&field_prodev_level_tid=8&page=5`. In Problem 4, undergraduates have the opportunity to examine hypothetical student work and to identify what the students do and do not yet understand. Undergraduates will also practice posing questions that may guide the students' statistical understanding.

Overall, the LOCUS Project is a useful resource for future teachers to be aware of as they can potentially use these kinds of questions in their future classrooms. The problems from the LOCUS Project have been developed to measure students' understanding across levels of development (elementary, middle school, high school) as identified in the Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II) Report (Bargagliotti et al., 2020), and they align with the Common Core State Standards for Mathematics.

Homework Problem 4

4. Four students (Daveed, Monica, Juliana, and Bryant) are working on the following problem together, but they all pick a different answer.

> The director of the City Transportation System is interested in the amount of time required for a bus to make the trip from Downtown Station to City Mall. After collecting data for several months by recording the time it takes to make the trip, she finds that the distribution of times has a standard deviation of 3 minutes.
>
> Which of the following is the best interpretation of the standard deviation?
>
> A. A bus that leaves from Downtown Station typically arrives at City Mall 3 minutes later than the scheduled time.
>
> B. A bus typically takes about 3 minutes to get from Downtown Station to City Mall.
>
> C. The time a bus takes to get from Downtown Station to City Mall never varies more than 3 minutes from the mean trip time.
>
> D. The difference between the actual time a bus takes to get from Downtown Station to City Mall and the mean trip time is, on average, about 3 minutes.

Daveed selects option A, Monica selects option D, Juliana selects option C, and Bryant selects option B.

(a) Who selected the correct answer?

Solution:

Monica (Choice D) selected the correct answer.

(b) For *each* student who selected an incorrect answer, examine the choice they selected and describe a statistical concept they do understand.

Sample Responses:

- Daveed (Choice A):
  - Daveed may understand that the typical difference from the expected time is 3 minutes.
- Bryant (Choice B):
  - Bryant may understand that the typical time can be used to characterize part of a distribution.
- Juliana (Choice C):
  - Juliana may understand that standard deviation quantifies the deviation from the mean.

(c) For *each* student who selected an incorrect answer, examine the choice they selected and describe a statistical concept they might not yet fully understand.

Solution:

Answers can vary. We provide sample responses below and quotes from the *Correct Answer and Commentary* link at the LOCUS Project website.

- Daveed (Choice A) and Bryant (Choice B):
  - Daveed's and Bryant's choice demonstrates that they may not yet understand the concept of standard deviation as a measure of spread.
- Juliana (Choice C):
  - Based on her choice, Juliana may be struggling with the concept that "the standard deviation is a statement about average variability from the mean and not maximum variability from the mean. This option is incorrect because it prescribes absolute bounds on the variability, e.g. 'never varies by more than 3 minutes.'" (LOCUS Project).
- "Options (A) and (B) both make statements regarding the typical time of the bus trip, which would be measures of center . . . " (LOCUS Project)

(d) For *each* student who selected an incorrect answer, write a question you could ask them to help guide their understanding of interpreting a standard deviation in the context of the problem. Briefly explain how your question may help guide their statistical understanding.

Sample Responses:

- Daveed (Choice A):
  - Can you tell me what standard deviation means in your own words?
  - This question will help me assess to what degree Daveed understands what standard deviation represents. Then, we can work backwards from his understanding to interpret the standard deviation in the context of the problem.
- Bryant (Choice B):
  - What property of the distribution is the standard deviation measuring?
  - Bryant seems to be interpreting 3 minutes as the mean and not the standard deviation. Based on his response to my question, I would follow up and ask him to explain what about his choice describes variability.
- Juliana (Choice C):
  - Let's assume the typical time for a bus to travel from Downtown Station to City Mall is 5 minutes. Is it possible for the bus to take 10 minutes to make that trip?

– Juliana is close to correctly interpreting the standard deviation in the context of the problem, but her use of "never" is problematic. By asking Juliana this question, it may help her see that it could take longer than 8 minutes (i.e., typical time + 3 minutes) due to unforeseen circumstances (like traffic, construction delays, accidents, etc.). Hopefully this will help her understand why the use of "never" is incorrect in the interpretation that she selected.

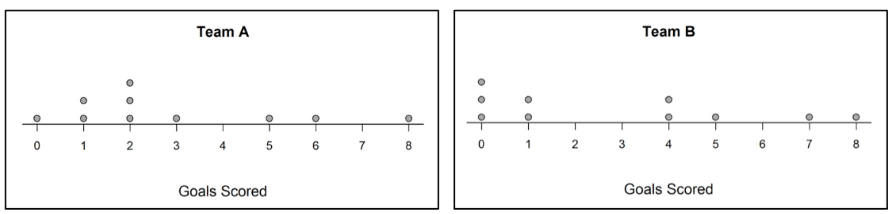
Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Parts of Problem 1 have been adapted from a problem created by the *LOCUS Project* (see `https://locus.statisticseducation.org/professional-development/questions/by-grade/grade-6?field_prodev_level_tid=All&page=6`). The primary purpose of this problem is to assess undergraduates' statistical understanding of computing and interpreting standard deviation.

Assessment Problem 1

1. Two soccer teams will be meeting in the city championship game. Each team played 10 games and averaged 3 goals scored per game for the season. The two dotplots below show the number of goals scored by each team per game for the season.



(a) Describe what one dot in the dotplot for Team A represents.

Solution:

One dot represents how many goals Team A scored in one of the games during the season.

(b) Compute the standard deviation (SD) for Team A.

Solution:

$$\begin{aligned}\bar{x} &= \frac{0+(2)1+(3)2+3+5+6+8}{10}\\ &= \frac{30}{10}\\ &= 3 \text{ goals}\end{aligned}$$

$$\begin{aligned}\text{SD} &= \sqrt{\frac{(0-3)^2+(2)(1-3)^2+(3)(2-3)^2+(3-3)^2+(5-3)^2+(6-3)^2+(8-3)^2}{10-1}}\\ &= \sqrt{\frac{58}{9}}\\ &\approx 2.54 \text{ goals}\end{aligned}$$

(c) Write a sentence that interprets the SD for Team A in the context of the problem.

Sample Responses:

- The SD tells us that Team A's goals scored per game differed from the mean by an average of 2.54 goals.
- The number of goals scored per game is typically 2.54 goals away from the mean, which is 3 goals.
- Since the SD for Team A is 2.54, then that tells us that on average, the goals scored were 2.54 away from the mean.

(d) Based on the dotplots (and without computing the SD for Team B), does Team A or Team B have more variability in the number of goals scored per game over the course of the season? Explain your reasoning.

Sample Responses:

- Team B has more games with goal counts further from the mean (3 goals), so it seems like Team B's season had more variability than Team A's.
- Team B has more variability. Visually, there are more data points at the extreme left, which indicates more variability from the mean.
- Team B has more variability because there are more numbers that are further away from the mean, making the SD larger.

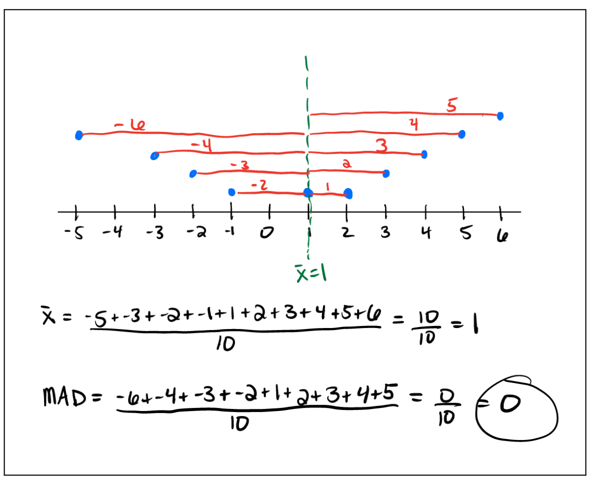
The purpose of Problem 2 is to assess undergraduates' knowledge of (1) examining hypothetical student work and identifying what the student does and does not yet understand about mean absolute deviation, and (2) posing questions to help guide the student's statistical understanding.

Assessment Problem 2

2. Delia was given the following dataset and asked to compute the MAD.

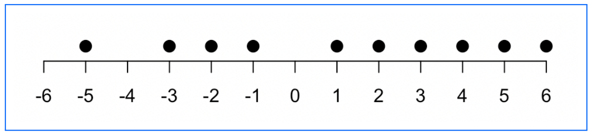
$$1, 5, 2, 6, -2, 3, 4, -1, -3, -5$$

She incorrectly computed the MAD, as shown in her work below.



(a) Draw a dotplot to graphically display the data. Explain how your dotplot shows that the MAD cannot equal 0.

Sample Response:



If the MAD is 0, then all dots on the dotplot should have the same value (like all 1's). But the dotplot clearly shows that variability is present in the data because the dots are spread out. Thus, the MAD can't be 0.

(b) Examine Delia's work and describe what she understands about MAD.

Sample Responses:

- Delia understands how to identify and compute deviations from individual data points to the mean of the dataset.
- Delia understands that computing the MAD involves using the mean of a dataset and deviations.
- Delia understands that MAD is an exact average of deviations.

(c) Identify the error(s) Delia made in her work.

Sample Response:

Delia incorrectly used the *signed* deviations (e.g., -3) rather than the absolute value of the deviations in her computation.

(d) Write one question you could ask Delia to help her correct her work. Briefly explain how your question might help guide Delia's statistical understanding of MAD.

Sample Responses:

- "What do you expect the sum of the deviations from the mean to be?" This will help me determine whether the student understands the mean as a balance point. Then I can assess what they conceptually understand about the mean (i.e., balance point) and the MAD (i.e., average distance from the mean).
- "If the MAD is 0, what does that mean in the context of the problem?" This question will hopefully guide Delia to recognize that if the MAD is 0, then no variability should be present in her dotplot.
- "What does MAD stand for and how are you using the deviations in your calculation?" Asking a student what MAD stands for will help me understand if they know they need to compute "absolute value deviations." Then I can help guide them to recognize that they are using signed deviations rather than taking the absolute value of the deviations in their calculation of MAD.
- "Can you draw a dotplot where the MAD is 0? How does that dotplot differ from your work here?" These questions will help me assess whether a student visually understands what a MAD of 0 looks like. If they correctly draw a dotplot with a MAD of 0, then they can visually see how the MAD of this dataset should not be 0.

4.6 References

[1] Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., Spangler, D. A. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education.* American Statistical Association.

[2] Census at School (Accessed April 7, 2023). `http://ww2.amstat.org/censusatschool/`

[3] delMas, R. C. (2001). What makes the standard deviation larger or smaller? *Statistics Teaching and Resource Library*. Available at `https://www.causeweb.org/cause/archive/repository/StarLibrary/activities/delmas2001/`

[4] *LOCUS Project* (Accessed April 7, 2023). `https://locus.statisticseducation.org/professional-development`

[5] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from `http://www.corestandards.org/`

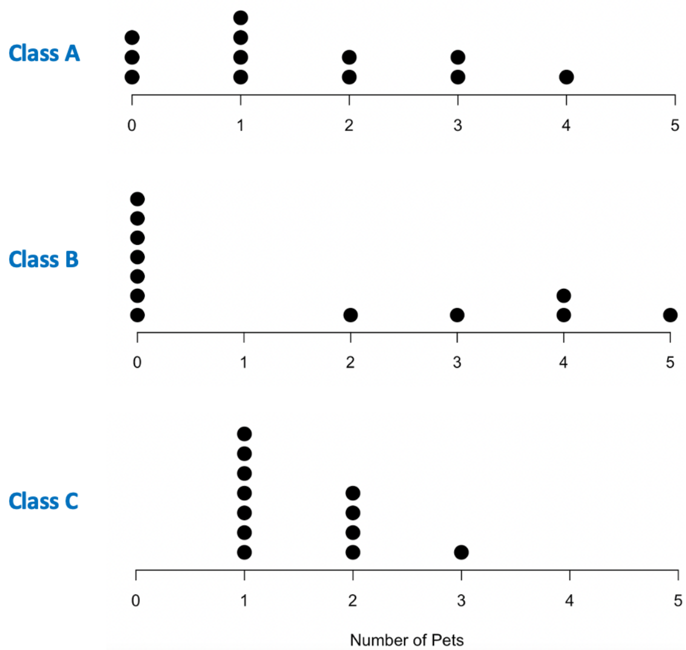
[6] Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L. (2008). *Mathematical statistics with applications*. Thomson Brooks/Cole.

4.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** VARIABILITY: MAD AND SD (page 1 of 1)

1. Students from three different classes reported the number of pets in their household. The results are summarized graphically as dotplots and in a frequency table below.



Class A	Class B	Class C
1	0	1
0	0	1
4	0	1
3	4	1
2	4	2
1	3	1
1	0	2
3	5	2
0	0	3
2	0	2
1	0	1
0	2	1

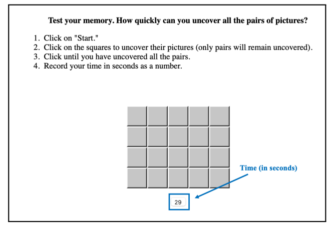
(a) Compute the mean number of pets for each class.

(b) What is similar about the three dotplots?

(c) What is different about the three dotplots?

NAME: **CLASS ACTIVITY:** VARIABILITY: MAD AND SD (page 1 of 6)

1. **Test Your Memory!**



Play the *Census at School Memory Game* where you will need to uncover and match 10 pairs of pictures. The time it takes you to complete the game will be tracked. Go to the following link to access the game: `https://ww2.amstat.org/education/cas/1.cfm`

(a) Play the game once and record your time (in seconds) as a number.

(b) As a class, compile everyone's time in a dataset. What graphical summary would be appropriate to visualize the class's distribution of times on the memory game? Explain your reasoning.

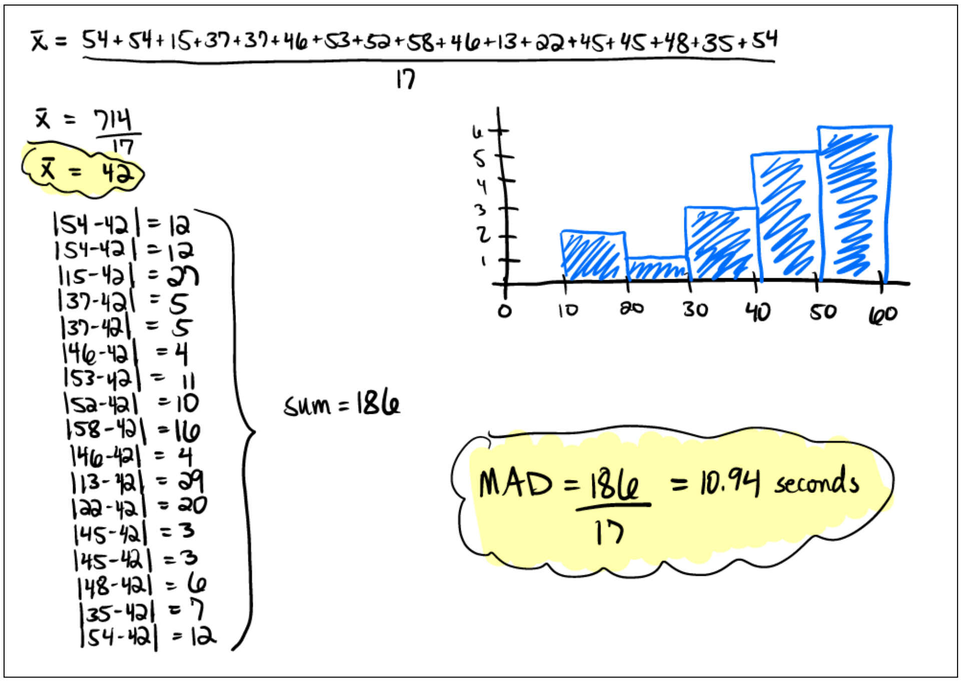
(c) As a class, create a graphical summary to visualize the class's distribution of times on the memory game and sketch it below.

(d) Describe what you notice about the class's distribution of times on the memory game.

2. **Quantifying Variability with Mean Absolute Deviation**
Amaury is teaching a high school intermediate algebra class and his students are learning about different measures of variability. The students in his class played the *Census at School Memory Game* and recorded their times, in seconds:

54, 54, 15, 37, 37, 46, 53, 52, 58, 46, 13, 22, 45, 45, 48, 35, 54

Amaury asked his students to create a graphical summary of the data, compute the mean, and quantify the amount of variability present. One of his students, Jasmine, did the following:



Jasmine, recalling what they learned in middle school, quantified the amount of variability by computing the **mean absolute deviation (MAD)**. All of their calculations are correct. Describe mathematically what Jasmine did to compute the MAD.

3. **Interpreting Mean Absolute Deviation**
Two other students, Tarryn and Benny, also correctly computed the MAD. When Amaury asked his students to write a sentence interpreting their measure of variability in the context of the problem, Jasmine, Tarryn, and Benny wrote the following sentences:

Jasmine	The MAD is 10.94 seconds.
Tarryn	On average, the memory game times were 10.94 seconds away from the mean of 42 seconds.
Benny	A memory game time is 10.94 from the mean.

(a) One student correctly (and completely) interpreted the MAD in the context of the problem. Identify who it was, and describe what components of their interpretation make it correct and complete.

(b) The other two students gave incomplete interpretations of the MAD. Based on their interpretations, describe what each
   i. may understand about interpreting the MAD, and
   ii. may not yet understand about interpreting the MAD.

(c) In a general context, describe what MAD measures.

4. **Quantifying Variability with Standard Deviation and Interpreting Standard Deviation**
Josief, another student in Amaury's class, did something different, as shown below:

$$\bar{x} = \frac{54+54+15+37+37+46+53+52+58+46+13+22+45+45+48+35+54}{17}$$

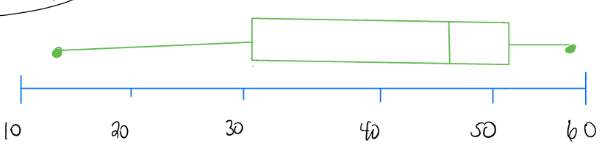
$$\bar{x} = 42$$

$$sd \Rightarrow (54-42)^2+(54-42)^2+(15-42)^2+(37-42)^2+(37-42)^2+(46-42)^2+(53-42)^2$$
$$+(52-42)^2+(58-42)^2+(46-42)^2+(13-42)^2+(22-42)^2+(45-42)^2+(45-42)^2$$
$$+(48-42)^2+(35-42)^2+(54-42)^2$$

$$= 144+144+729+25+25+16+121+100+256+16+841+400+9+9$$
$$+36+49+144$$

$$= 3064$$

SD,
$$s_x = \sqrt{\frac{3064}{16}} \approx 13.84$$



(a) Josief quantified the amount of variability by computing the **standard deviation (SD)**. All of his calculations are correct. Describe mathematically what he did to compute the SD.

(b) When asked to write a sentence to interpret the SD in the context of the problem, Josief wrote the following:

> The memory game times varied by 13.84 seconds.

Describe why Josief's interpretation is not completely correct.

(c) Consider the following questions that one might ask Josief to help him with his interpretation of the standard deviation (in the context of the memory game times).

i. Explain how the following question might help Josief to advance in his understanding of interpreting standard deviation in the context of a problem:

Can you say more about how the memory game times varied?

ii. Explain how the following question might help you assess what Josief understands about interpreting the standard deviation in the context of a problem:

What does standard deviation measure?

iii. Explain why the following question might not help Josief:

Why is your interpretation incorrect?

(d) In a general context, describe what SD measures.

5. Return to the class dataset from Problem 1.

   (a) Compute the mean absolute deviation of your class's data of memory game times and write a sentence interpreting the mean absolute deviation in the context of the problem.

   (b) Compute the standard deviation of your class's data of memory game times and write a sentence interpreting the standard deviation in the context of the problem.

NAME: **HOMEWORK PROBLEMS:** VARIABILITY: MAD AND SD (page 1 of 2)

1. Ten movies were randomly selected and the length of each movie (in minutes) is given below.

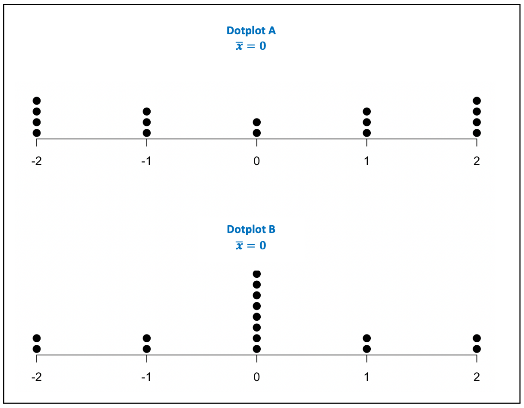
152, 156, 98, 173, 68, 122, 92, 105, 138, 126

(a) Compute the mean absolute deviation (MAD) and write a sentence that interprets the MAD in the context of the problem.

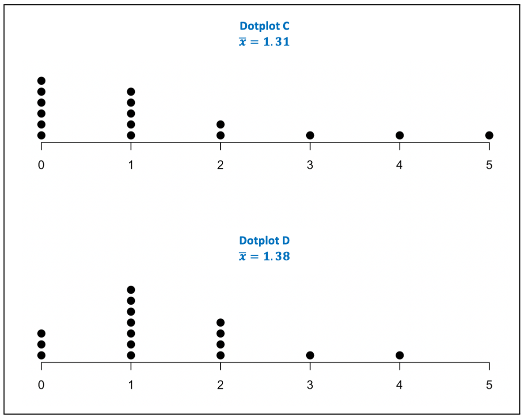
(b) Compute the standard deviation (SD) and write a sentence that interprets the SD in the context of the problem.

2. In this problem, you will *visually* compare the mean absolute deviation (MAD) between pairs of dotplots.

(a) **Dotplots A and B**.
Without doing any calculations, identify which dotplot (A or B) has the *larger* MAD. Explain your reasoning.



(b) **Dotplots C and D**.
Without doing any calculations, identify which dotplot (C or D) has the *larger* MAD. Explain your reasoning.



(c) Draw two different dotplots that have the same MAD. Describe your thought process when creating these two different dotplots.

3. Consider the mean absolute deviation (MAD) and the standard deviation (SD). Typically, MAD is first taught in middle school and SD is taught in high school. Describe why it is helpful for school students to learn MAD before SD.

4. Four students (Daveed, Monica, Juliana, and Bryant) are working on the following problem together, but they all pick a different answer.

> The director of the City Transportation System is interested in the amount of time required for a bus to make the trip from Downtown Station to City Mall. After collecting data for several months by recording the time it takes to make the trip, she finds that the distribution of times has a standard deviation of 3 minutes.
>
> Which of the following is the best interpretation of the standard deviation?
>
> A. A bus that leaves from Downtown Station typically arrives at City Mall 3 minutes later than the scheduled time.
>
> B. A bus typically takes about 3 minutes to get from Downtown Station to City Mall.
>
> C. The time a bus takes to get from Downtown Station to City Mall never varies more than 3 minutes from the mean trip time.
>
> D. The difference between the actual time a bus takes to get from Downtown Station to City Mall and the mean trip time is, on average, about 3 minutes.

Daveed selects option A, Monica selects option D, Juliana selects option C, and Bryant selects option B.

(a) Who selected the correct answer?

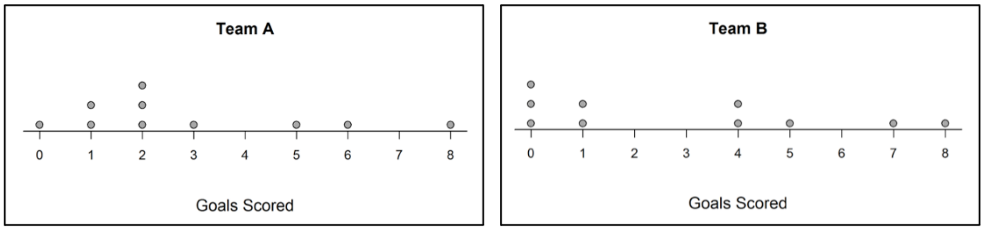
(b) For *each* student who selected an incorrect answer, examine the choice they selected and describe a statistical concept they do understand.

(c) For *each* student who selected an incorrect answer, examine the choice they selected and describe a statistical concept they might not yet fully understand.

(d) For *each* student who selected an incorrect answer, write a question you could ask them to help guide their understanding of interpreting a standard deviation in the context of the problem. Briefly explain how your question may help guide their statistical understanding.

NAME: **ASSESSMENT PROBLEMS:** VARIABILITY: MAD AND SD (page 1 of 3)

1. Two soccer teams will be meeting in the city championship game. Each team played 10 games and averaged 3 goals scored per game for the season. The two dotplots below show the number of goals scored by each team per game for the season.



(a) Describe what one dot in the dotplot for Team A represents.

(b) Compute the standard deviation (SD) for Team A.

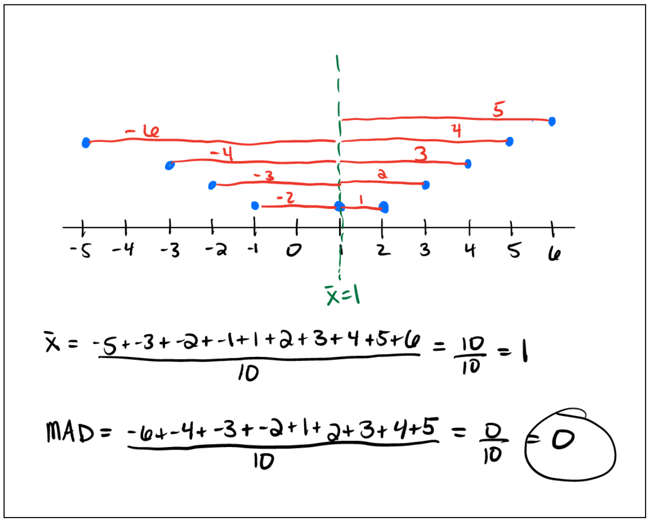
(c) Write a sentence that interprets the SD for Team A in the context of the problem.

(d) Based on the dotplots (and without computing the SD for Team B), does Team A or Team B have more variability in the number of goals scored per game over the course of the season? Explain your reasoning.

2. Delia was given the following dataset and asked to compute the MAD.

$$1, 5, 2, 6\ , -2, 3, 4, -1, -3, -5$$

She incorrectly computed the MAD, as shown in her work below.



(a) Draw a dotplot to graphically display the data. Explain how the dotplot shows that the MAD cannot equal 0.

(b) Examine Delia's work and describe what she understands about MAD.

(c) Identify the error(s) Delia made in her work.

(d) Write one question you could ask Delia to help her correct her work. Briefly explain how your question might help guide Delia's statistical understanding of MAD.

5

Using Sampling Distributions to Build Understanding of Margin of Error

Introduction to Statistics

Elizabeth G. Arnold, *Colorado State University*
Katharine M. Banner, *Montana State University*
Elizabeth W. Fulton, *Montana State University*
Rachel Tremaine, *Colorado State University*

5.1 Overview and Outline of Lesson

A goal of many statistical studies is to characterize interesting features of some population of interest (e.g., population mean, population variance). Although we can learn everything we would ever want to know about a population by examining its distribution, we usually cannot obtain the population distribution because we cannot observe every unit in the population. However, it may still be possible to learn something about the population by studying a subset (i.e., sample) of the population, as long as that sample is random. With some assumptions about the shape of the population distribution, sample statistics can be used to estimate population parameters with some margin of error. This lesson provides an opportunity for undergraduates to understand why we can use a sample mean (from a random sample) to estimate a population mean and how we can use the variability in a sampling distribution to help us estimate a margin of error.

1. Launch—Pre-Activity
   Prior to the lesson, undergraduates complete a Pre-Activity where they are presented with an example of a margin of error reported in the media and are asked to discuss why a margin of error might be reported.

2. Explore—Class Activity
   The context of the Class Activity focuses on investigating tests scores from a fake population of seniors at "Maplewood High School." Under a "fake world" premise where the population distribution is known, undergraduates work through the Class Activity to understand sampling variability in the context of using simulation to build an approximate sampling distribution for a sample mean. They then use the standard deviation of this sampling distribution (i.e., standard error) to build an understanding of margin of error.

   - *Problems 1 & 2*:
     Using 100 cards that represent the 100 test scores from a population of seniors at Maplewood High School, undergraduates engage in a hands-on simulation activity to get a feel for how sample means relate to the population mean. To do this, undergraduates randomly select samples of 10, compute the sample mean test score, create an approximate sampling distribution of sample means, and discuss sample-to-sample variability.

- *Problem 3*:
  Because technology provides a more efficient way of conducting a simulation, undergraduates use technology (e.g., StatKey, R, etc.) to conduct a simulation (instructors have access to the dataset of the 100 test scores). Undergraduates continue to simulate samples of size 10 and create an approximate sampling distribution of 500 sample mean test scores.
- *Problems 4–6*:
  Undergraduates explore two different methods (counting dots and the empirical rule) to identify the middle 95% of the sample mean test scores from their sampling distribution in Problem 3. Then, they use this information to construct a margin of error with a confidence level of 95% and relate this to $\pm 2\times$Standard Error. In Problem 6, undergraduates consider when it is more convenient to use either the counting dots method or the empirical rule to construct a margin of error with a confidence level of 68% and 90%.
- *Problem 7*:
  Problem 7 concludes the activity and prompts undergraduates to write a sentence that uses a point estimate (e.g., sample mean) and a margin of error to report a range of plausible values in the context of test scores for seniors at Maplewood High School.

3. Closure—Wrap-Up
   Conclude the lesson by emphasizing the main points of the lesson. That is, if we have a random sample, we can use sample statistics to estimate population parameters with some margin of error. We can place a margin of error around our sample estimate to provide us with a range of plausible values.

5.2 Alignment with College Curriculum

A simulation-based introduction to inference using appropriate technology supports the development of statistical reasoning and can aid in student understanding of sample-to-sample variation and how to make inference using sampling distributions (see Carver et al., 2016; Bargagliotti et al., 2020). Deriving a margin of error through simulation in collegiate courses provides prospective teachers an opportunity to develop a deeper understanding of concepts they may teach their future students, such as sampling variability and its connection to margin of error. Further, the use of simulation models may help all undergraduates better understand concepts like bootstrapping, which may appear later in your course.

Incorporate this lesson into your curriculum as you see fit. This lesson emphasizes the use of simulation to construct sampling distributions, which are used to build an intuitive understanding of margin of error; we conclude the lesson by indirectly asking undergraduates to find a range of plausible values. Although we do not explicitly refer to this range of plausible values as a confidence interval, a margin of error is directly tied to a desired confidence level so this lesson can naturally lead into a lesson about confidence intervals.

5.3 Links to School Mathematics

High school students are expected to conduct simulations to explore, describe, and summarize sample-to-sample variability of a statistic; that is, create an approximate sampling distribution of a statistic. This lesson addresses several statistical knowledge and mathematical practice expectations recommended by professional organizations (e.g., Common Core State Standards for Mathematics [CCSSM, 2010]; Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II [GAISE II; Bargagliotti et al., 2020]). For example, high school students are expected to understand that sample statistics (from a random sample of a population) can be used to estimate population parameters (c.f. CCSS.MATH.CONTENT.HSS.IC.A.1 and CCSS.MATH.CONTENT.7.SP.A.1). Further, they are expected to use simulation models to create an approximate sampling distribution and construct a margin of error (c.f. CCSS.MATH.CONTENT.7.SP.A.2 and CCSS.MATH.CONTENT.HSS.IC.4).

This lesson highlights:

- Making inferences about a population parameter based on a random sample from that population;
- Conducting a simulation to create an approximate sampling distribution;
- Constructing a margin of error using simulation models.

5.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- The difference between a population and a sample;
- How to create a dotplot and calculate summary statistics such as a sample mean;
- The empirical rule.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Use data from a random sample to estimate a population mean;
- Conduct a simulation to create an approximate sampling distribution and to construct a margin of error;
- Write a sentence that uses a point estimate (e.g., sample mean) and a margin of error to report a range of plausible values in the context of a problem;
- Examine hypothetical students' understanding of margin of error and evaluate questions to help guide students' statistical understanding about reporting a margin of error and using different methods for estimating a margin of error from a unimodal and symmetric sampling distribution.

Anticipated Length

Two 50-minute class sessions.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity (print Problems 1–2, 3, 4–6, and 7 to pass out separately)
  - 100 Test Score Cards (cut out cards and place in a bag for each group to use for Problems 1–2)
  - 100 Test Scores Dataset (.csv file to share with undergraduates when they conduct a computer simulation)
  - Appropriate technology and software for the instructor and undergraduates to conduct a simulation and create an approximate sampling distribution.
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on a quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files, along with the .csv file of Test Scores, can be downloaded from maa.org/meta-math.

5.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework for undergraduates to complete in preparation for the lesson, and ask undergraduates to bring their solutions to class on the day you start the Class Activity.

Pre-Activity Review (5–10 minutes)

As a class, discuss the solutions to the Pre-Activity.

Problem 1 illustrates how a margin of error (ME) is commonly reported in the media (i.e., Point Estimate $\pm$ ME), and by the end of the Class Activity, undergraduates will write a similar type of sentence that uses a sample mean and a margin of error to report a range of plausible values in the context of a problem.

Pre-Activity

1. Consider the following sentence from a statistical report, as presented in De Veaux, Velleman, and Bock (2012).

   Based on meteorological data for the past century, a local TV weather forecaster estimates that the region's average winter snowfall is 23 inches, with a margin of error of ± 2 inches.

   (a) If you lived in this region, would you want the margin of error to be large or small? Explain.

   Sample Responses:

   - I would want the margin of error to be small so that I could budget my snow removal costs more accurately. If there is a larger margin of error, it would be much more difficult to estimate how much I might need to pay for snow removal throughout the winter.
   - I'd want the margin of error to be small so that the estimate of "23 inches" is more likely to occur during the winter snowfall.

   (b) Why do you think a margin of error is reported?

   Sample Responses:

   - I think the margin of error is reported to show how much average variation there is in the amount of snow that falls each winter. If there is larger variation, you might have no snow at all or way more snow, so the 23 inch estimate would not be very helpful.
   - It's reported to illustrate the reliability of the "23 inches of snowfall" as a year-to-year expectation.
   - To give a sense for how much variation we can expect around 23 inches of annual snowfall in the region.

   Commentary:
   From our experience, it is common for learners to misinterpret the word "error" in margin of error and assume it means something is "wrong" (e.g., a measurement error). Address how the word "error" can be misleading and let undergraduates know that a margin of error does not represent some sort of "mistake" in the data collection process. Rather, it refers to "estimation error," which is directly related to sampling variability (or the variability of an estimator, when a random sample from a single population is assumed).

Conclude the Pre-Activity Review by letting undergraduates know that the purpose of this lesson is to use simulation models to develop an understanding of margin of error and discuss the following connection to teaching.

Discuss This Connection to Teaching

National and state content standards now emphasize statistics content standards to be integrated throughout grades 6–12. Within a high school intermediate algebra course, students are expected to construct a margin of error through the use of simulation models for random sampling. The goal of this lesson is to show undergraduates how they can conduct a simulation to construct an approximate sampling distribution and build up to a margin of error.

Class Activity: Problems 1 & 2 (15–20 minutes)

Pass out **Problems 1 and 2** of the Class Activity, and ask undergraduates to work in small groups. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion.

Introduce the purpose of these two problems by letting undergraduates know that

- The dotplot contains 100 test scores, which represents the entire population of test scores for seniors at Maplewood High School.

- Our population consists of 100 seniors and we are assuming that we know the population mean test score. Typically we do not have access to data from an entire population and thus do not know the population mean. However, for this activity we are going to assume we know what the population mean is so we can get a feel for how sample means relate to the population mean, characterize variability in the sampling process, and develop an understanding of how to construct a margin of error.
- In this activity, we are going to use simulation to take many random samples from the same population and compute each sample's mean. From this, we will look at the overall patterns of the sample means and see how they relate to the population mean.

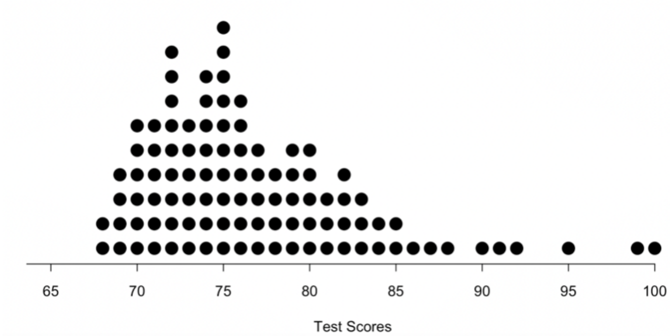
When a study in the real world is designed to learn about a population mean, we typically use the sample mean to estimate the population mean. In this lesson, undergraduates will build intuition about why the sample mean is a useful estimator of the population mean (when it is computed from a random sample). To help undergraduates build this intuition, we are applying a "fake world" context in this activity and assuming we know the population mean. This allows undergraduates to compare their sample means to the population mean and see how a sample mean from a random sample is usually a good guess of the population mean. Having each student (or small groups of students) take their own sample and compute their own sample mean also showcases the variability in the sampling process. Some samples will produce sample means very close to the population mean, while others will not. In the real world, we won't know if we have a sample mean "far from" or "close to" the true mean, so this motivates our need to understand how much variability we can expect in the sample means (of a fixed size from a fixed population).

We were intentional about generating this fake dataset in a way to not distract from concepts that build an understanding of margin of error. We chose a context that is straightforward to understand and created a dataset with exactly 100 data points whose values are "nice" in the sense that they are whole numbers, making all computations and methods for computing a margin of error as straightforward as possible.

Context for Class Activity

Investigating Test Scores for Seniors at Maplewood High School

Suppose that our entire population is 100 student test scores taken from the seniors at Maplewood High School (see the *population distribution* below), and we know the population mean test score (out of 100 points) is $\mu = 76.96$ points.



Before beginning Problems 1 and 2, ask undergraduates to briefly comment on the shape, center, and spread of the population distribution because they will make comparisons to this population distribution in Problem 3. Give each group a bag containing the 100 cards that represent the population of test scores, and demonstrate how to randomly draw a sample of size 10, as needed. Remind undergraduates to return the cards back to the bag (and to shuffle the cards within the bag) after they finish randomly selecting 10 of the cards. For Problem 1, focus the discussion on the sample-to-sample variation that is present among each group's sample of 10 test scores.

Class Activity Problem 1

1. You have a set of 100 cards that represent the 100 test scores from the entire population of seniors at Maplewood High School. Randomly draw a sample of **size 10** from these cards.

   (a) Write down the 10 test scores you randomly selected and compute the mean test score of your sample.

   Sample Response:
   Answers will vary. One sample is

   $$80, 72, 70, 82, 76, 79, 69, 74, 76, 83$$

   which has a mean of 76.1 points.

   Commentary:
   Make sure undergraduates record their sample mean on the Class Activity because they will use this value as their point estimate on Problem 7, where they will be prompted to write a sentence that uses a point estimate and a margin of error to report a range of plausible values in the context of the typical test score for seniors at Maplewood High School.

   (b) Share your sample mean test score with classmates and compare. Why do you think your sample mean test score is different from others?

   Sample Responses:
   An ideal response will include the concept of sample-to-sample variability.

   - My sample mean is different from others because of random chance. Different test scores may be selected when a different random sample is taken.
   - Our sample means are different because we did not all draw the exact same 10 test scores from our bags.

   Commentary:
   From our experience, some undergraduates may incorrectly explain that differences occurred as a result of having a "bad" sample or as a result of calculating the sample mean incorrectly. If this occurs, let undergraduates know that sampling variability is a characteristic of the real world, and it's not an indication of a "bad" sample or a computation mistake. The fact that not all random samples of 10 test scores are the same tells us that we need to expect some variation in our sample means. A single point estimate, like the sample mean computed in part (a) does not tell the whole story. We want to know "how good" this estimate is, and one way to report the quality of an estimate is to report the sample-to-sample (or sampling) variability. One way to convey the sampling variability is to report a range of plausible values.

In Problem 2, undergraduates continue to draw samples of size 10 from their bag of cards. Encourage undergraduates to take turns drawing samples from the bag. As undergraduates work on this problem, set up a way for you to compile the class's sample means so that you can display a dotplot of their sample means (i.e., an approximate sampling distribution of sample means). Alternatively, you may consider passing out 10 sticky notes to each group and asking them to write a sample mean test score on each sticky note. Then ask undergraduates to appropriately place their sticky notes on the board. This may help them visualize their "dots" in the sampling distribution.

Class Activity Problem 2

2. What happens if we repeat this process? From your set of 100 cards, continue to randomly draw a sample of size 10 and compute the mean test score of your sample. Repeat this process until you have a total of 10 sample means.

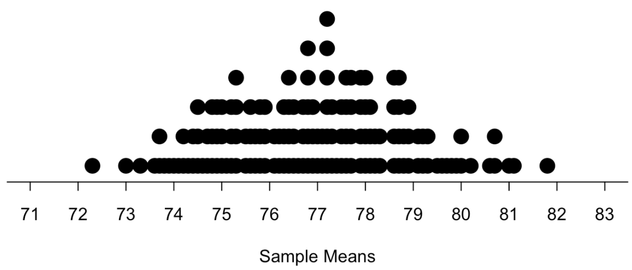
(a) Write the 10 sample mean test scores below.

Sample Response:
Answers will vary. Ten sample means (units = points) include:

$$76.1, 75.9, 75.9, 75.8, 76.4, 75.2, 76.6, 77.3, 79.7, 78.8$$

(b) As a class, create a dotplot of everyone's sample means and sketch it below. This is the class's *approximate sampling distribution* of sample mean test scores. Describe what one dot in the dotplot represents.

Solution:
Below is an example of a dotplot we created from 150 sample mean test scores. The mean of this approximate sampling distribution is 76.9 points and the standard deviation, referred to as the standard error, is 1.9 points.



The sample responses below describe what one dot in the dotplot represents:

- 100 cards were marked with test scores to represent the scores in the population. I sampled 10 of them and calculated the mean test score from that sample. The dot represents this sample mean.
- A dot represents the mean of one sample where 10 test scores were randomly selected.

Commentary:
Emphasize that you created an *approximate* sampling distribution of the sample mean test scores. We call it an approximate sampling distribution because it was constructed from roughly 150 sample means (exact number depends on how many groups of students you have in your class), and it was found empirically. True sampling distributions would include means from all possible samples of size 10 from the population of 100 test scores (i.e., the sample means resulting from the $\binom{100}{10}$ potential samples of size 10). "In real life, we get only one sample, but the sampling distribution gives us a mechanism for asking what would happen if we could take a random sample again and again" (Peck et al., 2013, p. 68). Simulation can be used to find approximate sampling distributions which provides us with a pretty good idea of what the true sampling distribution would look like.

"What is one dot?" is a challenging but useful question for undergraduates to think about. Describing what one dot represents will help undergraduates understand what the computer simulation is doing for them in Problem 3. If undergraduates are struggling with this question, ask them to find their dot on the sampling distribution and to describe what they did by hand with the cards to obtain the value of that dot.

After Problem 2, discuss the concept of a sampling distribution by sharing the following connection to teaching.

Discuss This Connection to Teaching

Sample-to-sample variation (or sampling variability) is a "big idea" concept in statistics (see Peck et al., 2013). The sampling distribution of a sample statistic (such as a sample mean) describes how sample statistics vary from one sample to the next. The sampling distribution of a sample statistic allows us to answer the following questions:

- How much will the value of a sample statistic tend to differ from one random sample to another?
- How much will the value of a sample statistic tend to differ from the corresponding population value? (Peck et al., 2013, p. 36–37)

The concept of a sampling distribution is fundamental to teaching and learning statistics, and implementing problems such as Problem 1 and 2 in the classroom is a way to allow students to notice that everyone's samples are similar but different. Further, by constructing an approximate sampling distribution, students can see sampling variability in the sample mean. All undergraduates will benefit from visualizing this sampling variability, and this is especially true for prospective teachers who will teach their future students about this big idea concept in statistics.

Class Activity: Problem 3 (20 minutes)

Pass out **Problem 3** of the Class Activity and instruct undergraduates to continue working in their small groups. Each group will need access to a computer and the dataset of 100 test scores so they can conduct a computer simulation. Instruct undergraduates on how to use your program of choice (e.g., StatKey, R) to conduct the simulation. As undergraduates use technology to build a sampling distribution, remind them to think about what they just did by hand and how it relates to what the computer is doing to conduct the simulated sampling distribution.

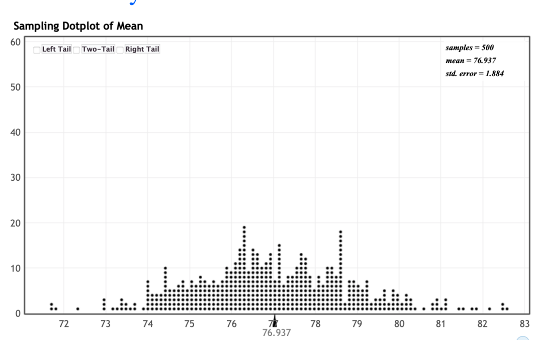
Class Activity Problem 3 : Parts a & b

We can continue to perform a simulation by hand with the cards, but technology provides us a more efficient way to do this! Our goal is to see what other sample means we might have gotten from different samples of size 10, and simulation is a tool we can use to see how variable the sample mean test scores might be.

3. Using the population of 100 test scores, use technology to conduct a simulation. Randomly draw a sample of **size 10** and compute the mean test score of your sample. Repeat this process for a total of 500 times.

   (a) Create a dotplot of your 500 sample mean test scores (i.e., another approximate sampling distribution) and sketch it below. Describe what one dot in the dotplot represents.

   Sample Response:
   Below is an example of a sampling distribution that was created in StatKey. The mean is 76.937 points and the standard error is 1.884 points. A single dot represents the mean of one sample where 10 test scores were randomly selected.



Commentary:

- Check that undergraduates' sampling distributions are reasonable—their sampling distribution should be centered around 76.96 (the population mean).
- We have found it useful to ask undergraduates to take a screen shot of their sampling distribution because they will continue to use this sampling distribution to answer Problems 4 and 5.

(b) Describe the shape, center, and spread of your sampling distribution from Problem 3(a). What is the mean and standard deviation of your sampling distribution? Note that the standard deviation of a sampling distribution is referred to as a *standard error*.

Sample Response:

- Shape = approximately unimodal and symmetric
- Center = 76.937 points (mean of the sampling distribution)
- Variability = 1.884 points (standard deviation of the sampling distribution)

Commentary:

Explain that the standard deviation of a sampling distribution is called a *standard error*. Emphasize that standard error is not an "error" in the sense of measurement error, but instead standard error represents the sample-to-sample variation in a population for a sample of size n. The standard error provides us with an estimate of the variability in the sample means that is due to the sampling process alone. That is, it provides us with a way to quantify how much background noise we can expect to see when taking a sample mean from a population (i.e., a guess at how much the value of a sample mean will differ from across samples). This is an essential piece of information for statistical inference because it allows us to provide a range of values instead of just one "best guess."

Problems 3(c) and 3(d) guide undergraduates to compare the approximate sampling distribution to the population distribution and to recognize that the mean of the approximate sampling distribution is centered near the population mean. Because the center of the sampling distribution is generally very close or at the population mean, the sample mean is an unbiased estimator of the typically unknown population mean. After discussing the solutions with the class, emphasize when a sample statistic can be used to estimate a population parameter (i.e., we have seen that when the sample is a random sample from the population of interest, it is often a good guess (i.e., "close to" the population mean). Though we won't know for certain if it's a good guess or a bad guess, we can use the standard error to help us understand how much variation there is between good and bad guesses in a given population.

Class Activity Problem 3 : Parts c & d

(c) Compare the shape, center, and spread of your sampling distribution from Problem 3(a) to the population distribution presented on page 1 of the Class Activity.

Sample Response:

- The population distribution and the sampling distribution are both unimodal. However, the population distribution is skewed right and the sampling distribution is unimodal and symmetric.
- The center of these two distributions is almost the same!
- The sampling distribution has smaller variability than the population distribution. The data in the population distribution range from about 67 points to 100 points but the data in the sampling distribution range from about 72 points to 82 points.

(d) Based on your sampling distribution from Problem 3(a), is a sample mean (computed from a random sample of size 10) a good way to estimate the population mean test score? Explain your reasoning.

Sample Responses:

- Yes, because most of the sample means in the sampling distribution are near the population mean of 76.9 points.
- We can see that a sample mean computed from a random sample of the population most often results in a sample mean that is close to the true population mean (i.e., the center and most of the mass of the sampling distribution is at or near the population mean).

Commentary:

- This is an important question. Although there were some samples that had sample means far away from the true population mean, most of the samples had sample means that were close to the true population mean.
- You may need to remind undergraduates that they are still working under the fake premise of knowing the population distribution. Further, the "sampling distribution gives us a mechanism for asking what would happen if we could take a random sample again and again" (Peck et al., 2013, p. 68). In real-life, we often get only one sample and if it was a random sample, we can use that sample mean as our "best guess" at the true population mean, and we can use the standard error to quantify how much variability we'd expect to see from sample mean to sample mean.

After discussing the solutions to Problem 3, emphasize the following connection to teaching, which focuses on the use of simulation-based method to teach and learn statistics.

Discuss This Connection to Teaching

High school teachers are expected to implement a simulation-based introduction to inference in their classes; their students will use data collected from a random sample and simulations to make inferences about a population (Peck et al., 2013). Prospective teachers need to be familiar with simulation as an instructional tool to develop their future students' conceptual understanding of many statistical ideas and to address expectations from content standards. For example, using a simulation to learn about sampling distributions and a margin of error provides students an opportunity to think about the process rather than trying to interpret theoretical approaches (Burrill, 2021). In addition, conducting a simulation by hand first and then following up with a computer simulation is a good pedagogical technique to use as students can connect what the computer is doing to what they did by hand.

Advice on Delivering the Lesson Over Two Class Sessions

If you are teaching this lesson over two class sessions, stopping around Problem 3 is a good place. See Chapter 1 for guidance on using exit tickets to facilitate instruction in a two-day lesson.

Discussion: Margin of Error and Range of Plausible Values (15 minutes)

The remainder of the Class Activity shifts its focus to building understanding of margin of error and writing a range of plausible values for a population parameter that we note are associated with different confidence levels. In this lesson, we do not spend time defining confidence or interpreting these ranges of plausible values as confidence intervals. We have found that instructors often follow this lesson with a lesson about confidence intervals and their interpretation.

Problems 4 and 5 do not rely on the standard, formal definition or formula of the margin of error that is commonly used when you are given only one sample. Because we are still operating under the fake premise of knowing the population distribution and being able to repeatedly sample from that population distribution, these problems prompt undergraduates to use their sampling distribution to construct a margin of error. Further, these problems focus on two methods ("counting dots" and the empirical rule) to understand the connection between the middle 95% of the sample means and a margin of error with a 95% confidence level. We leave it to the instructor to decide when they want to provide the formula of the margin of error; some teachers have found it works well to discuss the formula after this lesson and under the premise that we typically work with one random sample from an unknown population distribution.

Discuss This Connection to Teaching

Constructing a margin of error through the use of simulation models for random sampling is an explicit high school content standard that prospective teachers will be expected to teach.

Class Activity: Problems 4–6 (25 minutes)

Pass out **Problems 4–6** of the Class Activity, and instruct undergraduates to continue working in their small groups. Problem 4 focuses on "counting dots" to build a margin of error while Problem 5 focuses on using the empirical rule to build a margin of error. We intend that undergraduates will recognize that these two methods produce a similar answer for a margin of error with a confidence level of 95% because their sampling distribution is unimodal and symmetric.

Class Activity Problem 4

In Problem 3, we found that a sample mean (computed from a random sample) is usually a really good guess of the population mean. But, not all good guesses are created equally! We also saw in that sampling distribution that some of our good guesses were further from the population mean than others. In statistics, we care about how much variability is present among these sample means (estimated by the standard error you found in Problem 3), and we often report a range of plausible values for a population mean. Problems 4 and 5 guide you through two methods for estimating a margin of error from a unimodal and symmetric sampling distribution (use your sampling distribution from Problem 3 to answer Problems 4 and 5).

4. **Method 1: Counting Dots**

   (a) Based on counting dots in your sampling distribution, the **middle 95%** of the sample mean test scores land between ___________ points and ___________ points. Explain how you came up with these two values.

   Sample Response:
   The sample below is based on the sampling distribution presented in Problem 3(a).

   - To capture the middle 95% of the sample means I need to exclude 5% of the sample means. This means that I need to exclude 2.5% of the sample means from each tail end. Since we used 500 random samples of size 10 and 2.5% of 500 is 12.5, I need to exclude 12–13 dots from each tail end. When I do this, the middle 95% of the sample means land between 73.3 and 80.9.

   (b) The two values you found above in Problem 4(a) are both approximately ___________ points from the mean of your sampling distribution. Explain how you came up with your answer, which is a margin of error with a confidence level of 95%.

   Sample Response:
   Continuing from the sample response presented above in 4(a), we have that $76.9 - 73.3 = 3.6$ and $80.9 - 76.9 = 4$. These values are both about 3.8 points from the mean of the sampling distribution.

Commentary:
This method asks undergraduates to "count dots" (a technique they may be unfamiliar with) to determine where the middle (in this case) 95% of the sample means land. Under the assumption that a sampling

distribution is unimodal and symmetric, undergraduates could count the middle 95% of the dots on the sampling distribution, but a more efficient method would be to exclude 5% of the dots, or in other words, exclude 2.5% of the dots from each tail of the sampling distribution. Make sure both approaches are discussed in class.

Problem 5 prompts undergraduates to use the empirical rule to estimate a margin of error. From our experience, most undergraduates remember the empirical rule from high school, but remind them of the rule as needed. We have found that many undergraduates will recognize that they need to "go out" 2 standard deviations in the sampling distribution (i.e., 2 standard errors).

Class Activity Problem 5

5. **Method 2: Empirical Rule**
A student, Kyle, used the empirical rule instead of counting dots to estimate a margin of error with a confidence level of 95%. They said that using the empirical rule was quicker than counting dots and that they got a similar answer to their friends who counted dots.

(a) Use the empirical rule and your sampling distribution to estimate a margin of error with a confidence level of 95% (show your work). In other words, how many test score points do you need to go out from the mean of the sampling distribution to capture the **middle 95%** of the sample mean test scores? Compare your answer to your answer in Problem 4(b).

Sample Response:
The sample below continues to be based on the sampling distribution presented in Problem 3(a) as a sample response.

- The empirical rule tells me that the middle 95% of the data in the sampling distribution falls within 2 standard errors of the mean. So, we need to go out $2 \times 1.884 = 3.768$ points from the mean of 76.937 points.
- In 4(b), the margin of error was about 3.8 points which is very similar to 3.768 points found in 5(a)!

(b) Describe why Kyle might have thought to use the empirical rule to compute a margin of error in this situation.

Sample Response:
Kyle probably thought to use the empirical rule because they noticed that the sampling distribution is unimodal and symmetric and because we were looking to find the middle 95% of the sampling distribution.

(c) Explain why it's appropriate for Kyle to use the empirical rule to compute a margin of error in this situation.

Sample Response:
The empirical rule assumes that the distribution is unimodal and symmetric. Because our sampling distribution is approximately unimodal and symmetric, it is appropriate to use the empirical rule.

Commentary:
This method asks undergraduates to apply the empirical rule to determine where the middle 95% of the sample means land. We want undergraduates to recognize that this method is appropriate because (1) the sampling distribution is unimodal and symmetric and (2) we are looking for the middle 95% which directly ties to one of the three values (i.e., 68%–95%–99.7%) stated in the empirical rule.

Discuss This Connection to Teaching

The empirical rule is often taught in a high school algebra course. Students are expected to "use the mean and standard deviation of a data set to fit it to a normal distribution and to estimate population percentages" and "recognize that there are data sets for which such a procedure is not appropriate" (CCSSM, 2010, p. 81). Relating the empirical rule to a margin of error associated with a 95% confidence level is a way prospective teachers can help their students understand that statistical methods are only as strong as the assumptions upon which they are built, and critical assessment of those assumptions is one of the most important steps of a statistical analysis.

Problem 6 focuses undergraduates' attention on responding to others' statistical conjectures, which you can emphasize by discussing the following connection to teaching. Further, this problem helps undergraduates recognize when the "counting dots" method may be more efficient than the empirical rule to estimate a margin of error, assuming that a sampling distribution is approximately unimodal and symmetric.

Discuss This Connection to Teaching

Problem 6 focuses on analyzing other students' thinking in order to develop undergraduates' skills in understanding school student thinking and guiding school students' understanding. It is valuable for all undergraduates (especially prospective teachers) to think about how others use, reason with, and communicate statistics. These problems also give prospective teachers (and tutors and future graduate students) an opportunity to think about how they would respond to student work in ways that nurture students' assets and understanding, and in ways that help develop students' statistical understanding.

Class Activity Problem 6

6. Luis is another student in the same class as Kyle. He noticed what Kyle did and wonders if it will always work. He asks the teacher if they can always use the empirical rule in this situation to estimate a margin of error. The teacher recognizes that this is an opportunity to help her students understand when these methods can be used to estimate a margin of error and when one method is more convenient over the other. The teacher ask her students the following question:

   If you are given a sampling distribution that is approximately unimodal and symmetric, describe how you could estimate a margin of error with a confidence level of 68% and another margin of error with a confidence level of 90%.

   (a) Explain how this question can help Luis to understand more about under what conditions the empirical rule is a convenient method for estimating a margin of error from a unimodal and symmetric sampling distribution.

   Sample Response:
   This question helps Luis see that there are certain middle percentages of a unimodal and symmetric distribution that can be quickly described by the empirical rule, and there are others that are not so easy. The empirical rule is quick and easy for margin of errors with a confidence level of 68%, 95% and 99.7%, which correspond to the mean of the sampling distribution +/- 1, 2, 3 SD of the sampling distribution (i.e., the SE of the sample statistic), respectively. So, the 68% prompt connects to margin of error to 1 SD, but the 90% prompt doesn't have an exact match from the empirical rule.

(b) What other confidence levels could the teacher ask about to help students understand under what conditions the empirical rule is (or is not) a convenient method for estimating a margin of error from a unimodal and symmetric sampling distribution? Explain.

Sample Response:
The teacher could ask about the other two values in the empirical rule. That is, 95% and 99.7% to demonstrate when the empirical rule is a convenient method for estimating a margin of error. Any other percent besides 68%, 95% and 99.7% will likely be more convenient to use the counting dots method.

Class Activity: Problem 7 (5 minutes)

Pass out **Problem 7** of the Class Activity, which asks undergraduates to return to the context presented at the beginning of the Class Activity (investigating test scores for seniors at Maplewood High School) and write a sentence that summarizes what they discovered. The goal is for undergraduates to use their estimate of the population parameter (i.e., their sample mean from Problem 1(a)) and a margin of error (constructed using their approximate sampling distribution) to write a sentence similar to the following from the Pre-Activity: *"A local TV weather forecaster estimates that the region's average winter snowfall is 23 inches, with a margin of error of $\pm$ 2 inches."*

Class Activity Problem 7

7. Use the mean from your first random sample of 10 test scores (see Problem 1(a)) and the margin of error you found using your approximate sampling distribution (see Problem 4 or 5) to write a sentence that describes what you found about the typical test score for seniors at Maplewood High School. Your sentence can be modeled after the sentence describing snowfall in Problem 1 of the Pre-Activity.

Sample Response:
We estimate that the average test score for seniors at Maplewood High School is 76.1 points $\pm$ 3.8 points.

Commentary:

- Make sure undergraduates are using their sample mean from Problem 1(a) and not the center of their approximate sampling distribution as their point estimate. You can remind them that in the real-world, we often get only one sample. If that sample is a random sample, then the sample mean is our best guess at the true population mean.
- We have seen some undergraduates chose to write their answer in terms of a confidence interval (e.g., $(72.3, 79.9)$ points). If this occurs, then optionally and as appropriate for your class, focus on guiding undergraduates to correctly interpret their interval in the context of the problem. For example, you can have undergraduates compare their range of plausible values to the true mean. Then, undergraduates may notice that some of their peers' "intervals" contained the true mean and some did not.
- Alternatively, if undergraduates do not write their answer in terms of a confidence interval, this problem may be extended to introduce the concept of a confidence interval, and we leave it up to the instructor whether they want to spend time in this lesson or subsequent lessons defining and interpreting a confidence interval.

Wrap-Up (5 minutes)

Conclude the lesson by briefly discussing under what conditions a sample statistic is a good estimate of a population parameter, how undergraduates used simulation to construct a sampling distribution and build an understanding of margin of error, and why it is valuable to report a margin of error. This discussion may include the following prompts and ideas:

- In the Class Activity, how did the center of the approximate sampling distribution compare to the center of the population distribution? Do you think this will always be the case? What implications does this have in the real-world, where we often do not have access to the population distribution and can only rely on one random sample?
  - In the Class Activity, we saw (on average) how the sample means are really close to (if not exactly at) the population mean. This allows us to use a single sample mean (from a random sample) as our best guess at the true population mean. In other words, the sample mean is an unbiased estimator of the typically unknown population mean.
- Why is it valuable to report a margin of error?
  - A margin of error is valuable to report when estimating a population parameter because although a sample statistic may be a good guess at the population parameter, we know that there is inherent variability in the sampling strategy used. If we took another sample of the same size our "best guess" would change slightly (as we saw in the Class Activity). A margin of error lets us represent a range of plausible values we could expect to observe from a single sample of the population of interest.
- How did we use simulation to develop an understanding of margin of error?
  - Review how undergraduates set up a simulation to generate an approximate sampling distribution. Then, they used their sampling distribution and either counted dots or used the empirical rule to estimate a margin of error with a 95% confidence level.

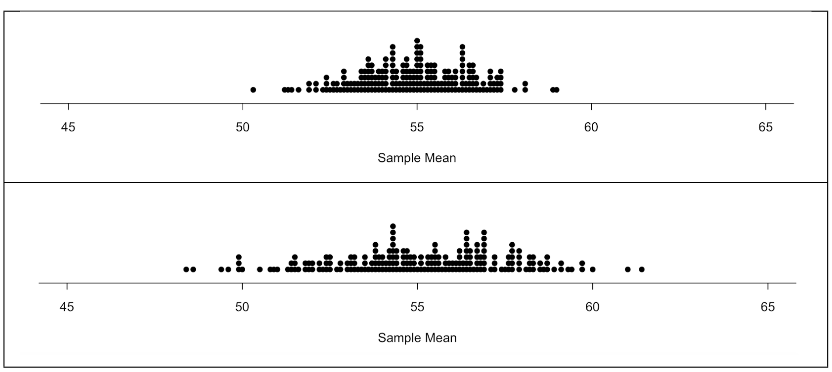
Homework Problems

At the end of the lesson, assign the following homework problems, and assign any additional homework problems at your discretion.

Problem 1 presents two sampling distributions of sample means, and asks undergraduates to explain what the population mean is and to estimate a margin of error for a confidence level of 95%. Because the standard error is not given, undergraduates need to rely on the "counting dots" method to construct a margin of error.

Homework Problem 1

1. Consider the two simulated sampling distributions of sample means.



(a) Assume the sample means came from a random sample. What's your best guess at where the population mean is? Explain.

Sample Response:

- A good estimate of the population mean is about 55 because both sampling distributions are centered around 55, and if we have used a random sample, then the center of a sampling distribution of sample means will be close to the true population mean.

(b) Based on the top sampling distribution, what is a reasonable estimate of a margin of error for a confidence level of 95%? Explain your reasoning. (Note that the sampling distribution was created using 200 samples of size 40.)

Sample Response:
I need to capture the middle 95% of the sample means which means that I need to exclude 2.5% of the sample means from each tail end. Since the sampling distribution was created with 200 random samples of size 40 and 2.5% of 200 is 5, I need to exclude 5 dots from each tail end. When I do this, the middle 95% of the sample means land between 49.9 and 59.4. The mean of the sampling distribution is about 55. Since $55 - 49.9 = 5.1$ and $59.4 - 55 = 4.4$, a reasonable estimate of the margin of error for a confidence level of 95% is about 4.8.

In Problem 2, undergraduates consider why a margin of error is reported and discuss what understanding students do and do not yet have when they select an incorrect answer. This problem is adapted from a problem created by the *LOCUS Project* (Levels of Conceptual Understanding in Statistics). You can view commentary and correct answers to the original problem at `https://locus.statisticseducation.org/professional-development/questions/interpret-results?type=prodev_multiple_choice_question&field_prodev_level_tid=8`.

The LOCUS Project is a useful resource for future teachers to be aware of as they can potentially use these kinds of questions in their future classrooms. The problems from the LOCUS Project have been developed to measure students' understanding across levels of development (elementary, middle school, high school) as identified in the Pre-K–12 Guidelines for Assessment and Instruction in Statistics Education II (GAISE II) Report (Bargagliotti et al., 2020), and they align with the Common Core State Standards for Mathematics.

Homework Problem 2

2. Saskia, Aaron, Gerlie, and Moses are working on the following problem.

> A survey of 625 randomly selected students was conducted to determine the average amount of time students sleep during a weekday. The survey reported an average of 6.5 hours. The survey estimate had a margin of error of half an hour. A margin of error is reported because
> A. Sample means vary from sample to sample.
> B. Students may intentionally respond incorrectly.
> C. Students may misunderstand the survey questions.
> D. The people doing the survey may have recorded results incorrectly.

Each student selects a different reason for why a margin of error is reported.

Choice A	Choice B	Choice C	Choice D
Gerlie	Saskia	Aaron	Moses

(a) Who selected the correct (and complete) answer and why?

Sample Undergraduate Responses:
- The correct answer is A because random samples vary which mean the sample means will also vary.
- Gerlie is the most correct because two simple random samples will most likely be different, even from the same population.
- Gerlie is correct because variability is inherent in sampling methods.

(b) For each of the three students who selected an incorrect choice, explain what conception they have of margin of error.

Sample Undergraduate Responses:

- Moses likely has some understanding but doesn't understand the point that sample variation is not necessarily a result of collection error.
- The three other students may be taking margin of error too literally. They may see the word error and assume that the data given were wrong or it was interpreted wrong, when it really means that the average that was determined for that sample could vary if the survey was replicated with a different 625 randomly selected students.
- Each response is similar in that they all have to do with the surveyed students themselves and not the data. They all sound like an error that could occur.
- Saskia, Aaron, and Moses may think that "error" means someone is at fault—whether it be the people giving the survey or the students taking the survey.

Problem 3 asks undergraduates to consider (and explain) what will happen to a margin of error when you increase the confidence level.

Homework Problem 3

3. Recall the context of test scores from the Class Activity. Suppose that you wanted to capture the middle 99% (instead of the middle 95%) of the mean test scores from a sample of size 10. How would your margin of error change? Explain your reasoning.

Sample Responses:

- Because we are capturing more of the sample mean test scores, we need to go out further from the mean of the sampling distribution which will make the margin of error larger.
- The ME will increase because we are getting a larger scope of values to see.

Problem 4 prompts undergraduates to describe how they can create a range of plausible values. Undergraduates will describe how they can use a random sample to estimate an unknown value of a population and also describe how to conduct a simulation to construct a sampling distribution and estimate a margin of error. This problem is modeled after the "Gettysburg Address Problem" (see Chance & Rossman, 2006).

Homework Problem 4

4. The proliferation of text generated by artificial intelligence has led to questions about how to distinguish passages that are written by humans compared to passages written by artificial intelligence, which leads to the need to examine characteristics of blocks of text. One characteristic of a block of text is the mean word length. Consider the excerpt of Martin Luther King Jr.'s "I Have a Dream" speech.

> And so even though we face the difficulties of today and tomorrow, I still have a dream. It is a dream deeply rooted in the American dream. I have a dream that one day this nation will rise up and live out the true meaning of its creed: We hold these truths to be self-evident, that all men are created equal.
>
> I have a dream that one day on the red hills of Georgia, the sons of former slaves and the sons of former slave owners will be able to sit down together at the table of brotherhood.
>
> I have a dream that one day even the state of Mississippi, a state sweltering with the heat of injustice, sweltering with the heat of oppression, will be transformed into an oasis of freedom and justice.
>
> I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream today.

Draw on your experiences from the Class Activity to estimate the mean word length of this passage with some margin of error:

(a) Without going through and calculating the length of every single word in the passage, describe how you would use statistics to construct a "best guess" at the true mean word length.

Sample Response:

I would randomly select 10 words, count how many letters are in each word, and compute the mean. This would give me a sample mean word length and since I used a random sample, it will be my best guess at the population mean word length.

(b) Describe how you would use simulation to construct a margin of error for the true mean word length?

Sample Response:

I would randomly select 10 words from the passage, record their length, and then compute the mean length from my sample. I would then repeat this process for a total of 500 times. Now I have 500 sample means. I can create a dotplot of these sample means and this would represent my approximate sampling distribution. I would use the mean of my first random sample as an estimate of the population mean word length. To create a plausible range of values for the mean word length, I would first count where the middle 95% of the dots in the sampling distribution are and see how far each is from the mean of the sampling distribution. This would give me a margin of error with a 95% confidence level. Finally, my range of plausible values would be constructed by computing "sample mean $\pm$ ME".

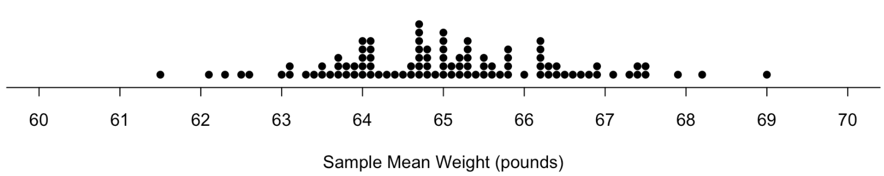
Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Problem 1 assess undergraduates' understanding of how a dotplot was created, how to develop a margin of error given a sampling distribution, and what happens to a margin of error when the associated confidence level is increased.

Assessment Problem 1

1. Below is a dotplot of the sample mean weight for 100 different random samples of size 10 from a population of adult Labrador retrievers where the mean weight is 65 pounds.



(a) Describe what one dot in the dotplot represents.

Sample Responses:

- The mean weight (in pounds) of a random sample of 10 retrievers.
- 10 adult Labrador retrievers were randomly selected and their body weight was recorded. One dot represents the mean of these 10 body weights.

(b) Fill in the blanks.

95% of the sample mean weights fall between ________________ *and* ________________.

Explain how you came up with these endpoints.

Sample Undergraduate Responses:

- 95% of the sample mean weights fall between 62.4 pounds and 67.8 pounds. To figure out where the middle 95% of the sample means were, I noticed that there are 100 samples shown in the sampling distribution. Therefore, I needed to capture all but 5 of the dots (or about 2.5 dots on each end).
- 62.3 and 67.7. I came up with these points by "chopping" off 2.5% of the dots on either side.
- 62.4 and 67.8. I found that 95% of 100 is 95, so I needed to count in 2.5 dots on each side to remove the other 5%, so the number of dots between my two endpoints marks 95% of the dots.

(c) Based on your answer in 1(b), estimate a margin of error with a confidence level of 95%. Explain your work.

Sample Responses:

Answers will vary depending on the endpoints undergraduates found in 1(b). The sample responses below correspond to the three sample responses in 1(b).

- The population mean is 65 pounds and the interval above should be approximately $65 \pm \text{ME}$. Since $65 - 62.4 = 2.6$ and $67.8 - 65 = 2.8$, the margin of error is about 2.7 pounds.
- $62.3 + 2.7 = 65$ and $67.7 - 2.7 = 65$. So the MoE $= 2.7$ pounds.
- I took the endpoints above, found the distance between the two values and divided by 2 to find the margin of error. $67.8 - 62.4 = 5.4$ and $5.4 \div 2 = 2.7$

(d) Would a margin of error with a confidence level of 99% be larger or smaller than the margin of error you estimated in 1(c)? Explain your reasoning.

Sample Responses:

- A margin of error with a confidence level of 99% would be larger than a margin of error with a confidence level of 95% because we are capturing more of the sample means, so the two endpoints would be farther from the mean.
- With a 99% confidence level we are capturing more sample means. This means the margin of error will be larger.

Problem 2 assesses undergraduates' understanding of the two methods they used in the Class Activity to estimate a margin of error, given a sampling distribution. In this problem, they will explain when one method is more efficient over the other, under the assumption that a sampling distribution is unimodal and symmetric.

Assessment Problem 2

2. Kyle's and Luis' teacher knows that her students can count dots and use the empirical rule on a unimodal and symmetric sampling distribution to estimate a margin of error. It's because the sampling distribution is symmetric and unimodal that both methods will give approximately the same answer for a margin of error. The teacher wants her students to understand when each method is (or is not) the most useful method for estimating a margin of error in this situation.

   (a) The teacher gives her students a unimodal and symmetric distribution and asks them to estimate a margin of error with a confidence level of 68%. Explain why the teacher uses this prompt to help her students understand when the empirical rule is more useful than the counting dots method to estimate a margin of error associated with a 68% confidence level.

Sample Response:
The teacher uses a confidence level of 68% because it's one of the three values in the empirical rule. All students would need to do is to compute 1 standard deviation of the sampling distribution. This is much quicker than counting the middle 68% of the dots in the sampling distribution.

(b) The teacher then asks her students to use the same sampling distribution from part (a) to now estimate a margin of error with a confidence level of 90%. Explain why this question is useful in helping students understand when the counting dots method is more useful than the empirical rule to estimate a margin of error associated with a 90% confidence level.

Sample Response:
Because 90% is not one of the three numbers in the empirical rule, it would not be a single calculation (like 3 times the standard deviation of the sampling distribution) to estimate a margin of error. Students would need to figure out how to get 90% from using 68%, 95%, and 99.7% to use the empirical rule. At this point, it would be more efficient to count the middle 90% of the sample means, or rather exclude 5% of the sample means from each tail in the sampling distribution. Then find the distance between those endpoints and the mean of the sampling distribution to construct a margin of error associated with a 90% confidence level.

5.6 References

[1] Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., Spangler, D. A. (2020). *Pre-K–12 guidelines for assessment and instruction in statistics education II (GAISE II): A framework for statistics and data science education.* American Statistical Association.

[2] Burrill, G. (2021). Simulation as a tool to make sense of the world. In R. Helenius & E. Falck (Eds.), *Statistics education in the era of data science: Proceedings of the satellite conference of the International Association for Statistical Education (IASE).*

[3] Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M, . . . , Wood, B. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report*. American Statistical Association.

[4] De Veaux, R. D., Velleman, P. F., & Bock, D. E. (2012). *Stats: Data and models* (3rd ed.). Pearson Education.

[5] *LOCUS Project* (Accessed April 7, 2023). `https://locus.statisticseducation.org/professional-development`.

[6] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from `http://www.corestandards.org/`

[7] Peck, R., Gould, R., Miller, S., Wilson, P., & Zbiek, R. (2013). *Developing essential understanding of statistics for teaching mathematics in grades 9–12*. National Council of Teachers of Mathematics.

[8] StatKey Simulation Environment. (Accessed April 7, 2023). `http://www.lock5stat.com/StatKey/`.

5.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** UNDERSTANDING MARGIN OF ERROR (page 1 of 1)

1. Consider the following sentence from a statistical report, as presented in De Veaux, Velleman, and Bock (2012).

Based on meteorological data for the past century, a local TV weather forecaster estimates that the region's average winter snowfall is 23 inches, with a margin of error of $\pm$ *2 inches.*

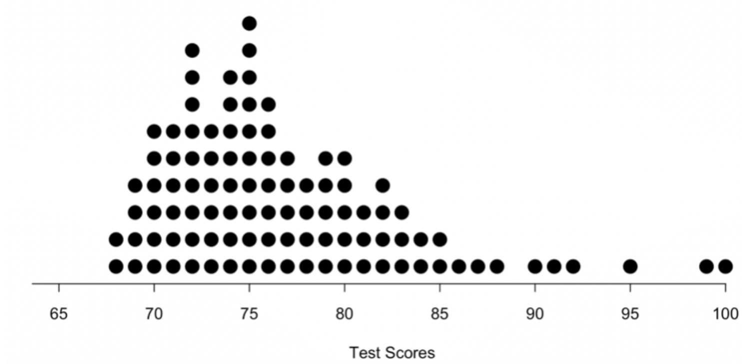
(a) If you lived in this region, would you want the margin of error to be large or small? Explain.

(b) Why do you think a margin of error is reported?

NAME: **CLASS ACTIVITY:** UNDERSTANDING MARGIN OF ERROR (page 1 of 5)

Investigating Test Scores for Seniors at Maplewood High School

Suppose that our entire population is 100 student test scores taken from the seniors at Maplewood High School (see the *population distribution* below), and we know the population mean test score (out of 100 points) is $\mu = 76.96$ points.



1. You have a set of 100 cards that represent the 100 test scores from the entire population of seniors at Maplewood High School. Randomly draw a sample of **size 10** from these cards.
   (a) Write down the 10 test scores you randomly selected and compute the mean test score of your sample.
   (b) Share your sample mean test score with classmates and compare. Why do you think your sample mean test score is different from others?

2. What happens if we repeat this process? From your set of 100 cards, continue to randomly draw a sample of size 10 and compute the mean test score of your sample. Repeat this process until you have a total of 10 sample means.
   (a) Write the 10 sample mean test scores below.
   (b) As a class, create a dotplot of everyone's sample means and sketch it below. This is the class's *approximate sampling distribution* of sample mean test scores. Describe what one dot in the dotplot represents.

We can continue to perform a simulation by hand with the cards, but technology provides us a more efficient way to do this! Our goal is to see what other sample means we might have gotten from different samples of size 10, and simulation is a tool we can use to see how variable the sample mean test scores might be.

3. Using the population of 100 test scores, use technology to conduct a simulation. Randomly draw a sample of **size 10** and compute the mean test score of your sample. Repeat this process for a total of 500 times.

   (a) Create a dotplot of your 500 sample mean test scores (i.e., another approximate sampling distribution) and sketch it below. Describe what one dot in the dotplot represents.

   (b) Describe the shape, center, and spread of your sampling distribution from Problem 3(a). What is the mean and standard deviation of your sampling distribution? Note that the standard deviation of a sampling distribution is referred to as a *standard error*.

   (c) Compare the shape, center, and spread of your sampling distribution from Problem 3(a) to the population distribution presented on page 1 of the Class Activity.

   (d) Based on your sampling distribution from Problem 3(a), is a sample mean (computed from a random sample of size 10) a good way to estimate the population mean test score? Explain your reasoning.

In Problem 3, we found that a sample mean (computed from a random sample) is usually a really good guess of the population mean. But, not all good guesses are created equally! We also saw in that sampling distribution that some of our good guesses were further from the population mean than others. In statistics, we care about how much variability is present among these sample means (estimated by the standard error you found in Problem 3), and we often report a range of plausible values for a population mean. Problems 4 and 5 guide you through two methods for estimating a margin of error from a unimodal and symmetric sampling distribution (use your sampling distribution from Problem 3 to answer Problems 4 and 5).

4. **Method 1: Counting Dots**

   (a) Based on counting dots in your sampling distribution, the **middle 95%** of the sample mean test scores land between ___________ points and ___________ points. Explain how you came up with these two values.

   (b) The two values you found above in Problem 4(a) are both approximately ___________ points from the mean of your sampling distribution. Explain how you came up with your answer, which is a margin of error with a confidence level of 95%.

5. **Method 2: Empirical Rule**
   A student, Kyle, used the empirical rule instead of counting dots to estimate a margin of error with a confidence level of 95%. They said that using the empirical rule was quicker than counting dots and that they got a similar answer to their friends who counted dots.

   (a) Use the empirical rule and your sampling distribution to estimate a margin of error with a confidence level of 95% (show your work). In other words, how many test score points do you need to go out from the mean of the sampling distribution to capture the **middle 95%** of the sample mean test scores? Compare your answer to your answer in Problem 4(b).

(b) Describe why Kyle might have thought to use the empirical rule to compute a margin of error in this situation.

(c) Explain why it's appropriate for Kyle to use the empirical rule to compute a margin of error in this situation.

6. Luis is another student in the same class as Kyle. He noticed what Kyle did and wonders if it will always work. He asks the teacher if they can always use the empirical rule in this situation to estimate a margin of error. The teacher recognizes that this is an opportunity to help her students understand when these methods can be used to estimate a margin of error and when one method is more convenient over the other. The teacher ask her students the following question:

   If you are given a sampling distribution that is approximately unimodal and symmetric, describe how you could estimate a margin of error with a confidence level of 68% and another margin of error with a confidence level of 90%.

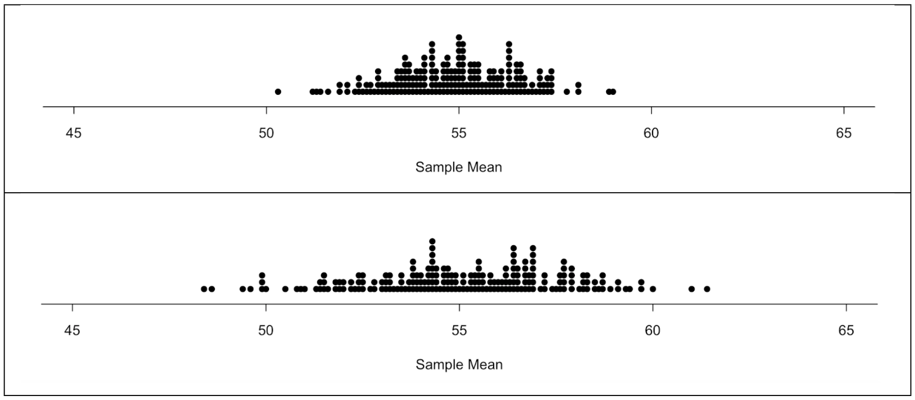
   (a) Explain how this question can help Luis to understand more about under what conditions the empirical rule is a convenient method for estimating a margin of error from a unimodal and symmetric sampling distribution.

   (b) What other confidence levels could the teacher ask about to help students understand under what conditions the empirical rule is (or is not) a convenient method for estimating a margin of error from a unimodal and symmetric sampling distribution? Explain.

7. Use the mean from your first random sample of 10 test scores (see Problem 1(a)) and the margin of error you found using your approximate sampling distribution (see Problem 4 or 5) to write a sentence that describes what you found about the typical test score for seniors at Maplewood High School. Your sentence can be modeled after the sentence describing snowfall in Problem 1 of the Pre-Activity.

NAME: **HOMEWORK PROBLEMS:** UNDERSTANDING MARGIN OF ERROR (page 1 of 2)

1. Consider the two simulated sampling distributions of sample means.



   (a) Assume the sample means came from a random sample. What's your best guess at where the population mean is? Explain.

   (b) Based on the top sampling distribution, what is a reasonable estimate of a margin of error for a confidence level of 95%? Explain your reasoning. (Note that the sampling distribution was created using 200 samples of size 40.)

2. Saskia, Aaron, Gerlie, and Moses are working on the following problem.

   > A survey of 625 randomly selected students was conducted to determine the average amount of time students sleep during a weekday. The survey reported an average of 6.5 hours. The survey estimate had a margin of error of half an hour. A margin of error is reported because
   >
   > A. Sample means vary from sample to sample.
   >
   > B. Students may intentionally respond incorrectly.
   >
   > C. Students may misunderstand the survey questions.
   >
   > D. The people doing the survey may have recorded results incorrectly.

   Each student selects a different reason for why a margin of error is reported.

Choice A	**Choice B**	**Choice C**	**Choice D**
Gerlie	Saskia	Aaron	Moses

   (a) Who selected the correct (and complete) answer and why?

   (b) For each of the three students who selected an incorrect choice, explain what conception they have of margin of error.

3. Recall the context of test scores from the Class Activity. Suppose that you wanted to capture the middle 99% (instead of the middle 95%) of the mean test scores from a sample of size 10. How would your margin of error change? Explain your reasoning.

HOMEWORK PROBLEMS: UNDERSTANDING MARGIN OF ERROR (page 2 of 2)

4. The proliferation of text generated by artificial intelligence has led to questions about how to distinguish passages that are written by humans compared to passages written by artificial intelligence, which leads to the need to examine characteristics of blocks of text. One characteristic of a block of text is the mean word length. Consider the excerpt of Martin Luther King Jr.'s "I Have a Dream" speech.

   > And so even though we face the difficulties of today and tomorrow, I still have a dream. It is a dream deeply rooted in the American dream. I have a dream that one day this nation will rise up and live out the true meaning of its creed: We hold these truths to be self-evident, that all men are created equal.
   >
   > I have a dream that one day on the red hills of Georgia, the sons of former slaves and the sons of former slave owners will be able to sit down together at the table of brotherhood.
   >
   > I have a dream that one day even the state of Mississippi, a state sweltering with the heat of injustice, sweltering with the heat of oppression, will be transformed into an oasis of freedom and justice.
   >
   > I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character. I have a dream today.

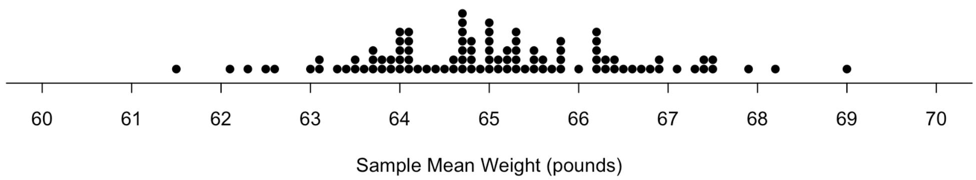
   Draw on your experiences from the Class Activity to estimate the mean word length of this passage with some margin of error:

   (a) Without going through and calculating the length of every single word in the passage, describe how you would use statistics to construct a "best guess" at the true mean word length.

   (b) Describe how you would use simulation to construct a margin of error for the true mean word length?

NAME: **ASSESSMENT PROBLEMS:** UNDERSTANDING MARGIN OF ERROR (page 1 of 2)

1. Below is a dotplot of the sample mean weight for 100 different random samples of size 10 from a population of adult Labrador retrievers where the mean weight is 65 pounds.



(a) Describe what one dot in the dotplot represents.

(b) Fill in the blanks.

95% of the sample mean weights fall between _________________ *and* _________________.

Explain how you came up with these endpoints.

(c) Based on your answer in 1(b), estimate a margin of error with a confidence level of 95%. Explain your work.

(d) Would a margin of error with a confidence level of 99% be larger or smaller than the margin of error you estimated in 1(c)? Explain your reasoning.

2. Kyle's and Luis' teacher knows that her students can count dots and use the empirical rule on a unimodal and symmetric sampling distribution to estimate a margin of error. It's because the sampling distribution is symmetric and unimodal that both methods will give approximately the same answer for a margin of error. The teacher wants her students to understand when each method is (or is not) the most useful method for estimating a margin of error in this situation.

   (a) The teacher gives her students a unimodal and symmetric distribution and asks them to estimate a margin of error with a confidence level of 68%. Explain why the teacher uses this prompt to help her students understand when the empirical rule is more useful than the counting dots method to estimate a margin of error associated with a 68% confidence level.

   (b) The teacher then asks her students to use the same sampling distribution from part (a) to now estimate a margin of error with a confidence level of 90%. Explain why this question is useful in helping students understand when the counting dots method is more useful than the empirical rule to estimate a margin of error associated with a 90% confidence level.

6

The Binomial Theorem

Discrete Mathematics or Introduction to Proof

Elizabeth A. Burroughs, *Montana State University*
Elizabeth G. Arnold, *Colorado State University*
Elizabeth W. Fulton, *Montana State University*
Douglas E. Ensley, *Shippensburg University*
Nancy Ann Neudauer, *Pacific University*

6.1 Overview and Outline of Lesson

The binomial theorem is studied in college-level Discrete Mathematics or Introduction to Proof courses, and its use is included in most high school mathematics standards. This lesson can fit into a course after students have learned that $\binom{n}{k}$ counts the number of ways to select k different objects from a set of n objects. This lesson examines binomial expansions, binomial coefficients, and the appearance of the binomial theorem in secondary mathematics. Undergraduates use combinatorial reasoning to expand the general expression $(x + y)^n$ and apply the binomial theorem in various problems. They also analyze hypothetical student work in order to develop skills in understanding school student thinking about expanding binomials and in creating questions to guide school students' understanding.

Throughout this lesson we use the term *arithmetic triangle* to refer to the triangular array of natural numbers that is also known as Pascal's triangle and Yang Hui's triangle, among other names.

Different courses in Discrete Mathematics or Introduction to Proof may organize the material contained in this lesson in different ways. This lesson proceeds by showing that the number of strings made of k x's and $n - k$ y's is $\binom{n}{k}$. When these strings are interpreted as multiplication of x's and y's, then each string composed of k x's and $n - k$ y's becomes $x^k y^{n-k}$, and the binomial theorem follows. Along the way, the lesson establishes that the entries in the arithmetic triangle can be computed by $\binom{n}{k}$ and that $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

1. Launch—Pre-Activity
   Prior to the lesson, undergraduates complete a Pre-Activity where they identify patterns in the arithmetic triangle and look for patterns among the expansions of $(x + y)^2$, $(x + y)^3$, and $(x + y)^4$. Instructors can launch the lesson by reviewing undergraduates' responses on the Pre-Activity.

2. Explore—Class Activity

   • *Problems 1–4*:
   Undergraduates use combinatorial reasoning to count the number of 4-character strings that use 3 of one

character and 1 of another. They use this reasoning to reconsider the three binomial expansions from the Pre-Activity to identify the total number of terms in each binomial expansion and the number of like terms in each binomial expansion. They reason through the $n = 4$ instance of $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$. Then, undergraduates use combinatorial reasoning to determine the coefficients of terms in the expansion of $(x + y)^5$, examining why the coefficients of terms in binomial expansions are called "binomial coefficients" (i.e., $\binom{n}{k}$).

- *Problems 5 & 6*:
  Undergraduates generalize the expansion of $(x + y)^n$. Problem 5 provides instructors an opportunity to formally state and prove the binomial theorem and to address how and when the binomial theorem appears in secondary mathematics. Undergraduates apply the binomial theorem in Problem 6.
- *Problems 7 & 8*:
  Undergraduates analyze hypothetical student work to make sense of the binomial theorem. Problem 7 highlights two different perspectives of the binomial coefficients in the binomial theorem, and undergraduates will develop questions to guide the hypothetical students' understanding. In Problem 8, undergraduates examine how a student incorrectly applied the binomial theorem to expand $(2x - y)^4$. Undergraduates are prompted to identify correct and incorrect thinking in the student's work and to consider how they would respond to the student to help guide the student's understanding of the binomial theorem.

3. Closure—Wrap-Up
   Conclude the lesson by summarizing the connections between the arithmetic triangle and the binomial theorem and how and when the binomial theorem appears in secondary mathematics. Additionally, discuss how undergraduates used combinatorial reasoning throughout the lesson.

6.2 Alignment with College Curriculum

The binomial theorem is a topic that fits naturally in a Discrete Mathematics or an Introduction to Proof course. The binomial theorem offers undergraduates an opportunity to learn combinatorial proof techniques and contrast them with the algebraic techniques they are often more familiar with.

6.3 Links to School Mathematics

The binomial theorem has direct connections to multiplying binomials, a skill students use throughout their mathematical studies. Prospective teachers should understand the process of multiplying binomials beyond rote procedure and memorization. By studying connections between the arithmetic triangle and the binomial theorem, prospective teachers will develop their abilities to use combinatorial reasoning.

This lesson highlights:

- Connections between the binomial theorem and the arithmetic triangle;
- Using patterns to develop a combinatorial understanding of the arithmetic triangle and the binomial theorem.

This lesson addresses several mathematical knowledge and practice expectations included in high school standards documents, such as the Common Core State Standards for Mathematics (CCSSM, 2010). High school students are expected to know and apply the binomial theorem for binomial expansions and be able to determine binomial coefficients using the arithmetic triangle (c.f. CCSS.MATH.CONTENT.HSA.APR.C.5). This lesson also provides opportunities for prospective teachers to think about the reasoning of others, construct sound mathematical arguments, and look for and make use of structure.

6.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- The multiplication and addition rules for solving counting problems;
- That $\binom{n}{k}$ counts the number of ways to select k different objects from a set of n objects.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Use combinatorial notation to write the expansion of $(x + y)^n$;
- Apply the binomial theorem to expand binomials and to determine specific coefficients of binomial expansions;
- Identify combinatorial patterns and reasoning in both the arithmetic triangle and the proof of binomial theorem;
- Identify connections between the binomial theorem and the mathematics of secondary school;
- Examine hypothetical school student work in order to identify what a school student does and does not yet understand about the binomial theorem and pose questions to help guide school students' understanding about the use of the binomial theorem.

Anticipated Length

Two 50-minute class sessions.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity (print Problems 1–4, 5–6, and 7–8 to pass out separately)
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

6.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to complete in preparation for the lesson, and ask undergraduates to bring their solutions to class on the day you start the Class Activity.

Pre-Activity Review (15 minutes)

As a class, review undergraduates' responses to the Pre-Activity.

For Problem 1, first ask undergraduates to share some of the patterns they observed in the arithmetic triangle. They will likely have different levels of familiarity with the triangle, and you can use this discussion to establish a common set of patterns to refer to throughout the lesson.

Pre-Activity Problem 1

1. Below are the first seven rows of the *arithmetic triangle*, also known as Yang Hui's triangle or Pascal's triangle, among other names. By custom, the rows and entries are numbered starting at 0.

$$\begin{array}{ccccccccccccc}
&&&&&&1&&&&&&\\
&&&&&1&&1&&&&&\\
&&&&1&&2&&1&&&&\\
&&&1&&3&&3&&1&&&\\
&&1&&4&&6&&4&&1&&\\
&1&&5&&10&&10&&5&&1&\\
1&&6&&15&&20&&15&&6&&1
\end{array}$$

(a) Write down at least three patterns that you observe.

Sample Responses:

- Each row starts and ends with the number 1.
- It's symmetric. The second and second to last number in a row are the same.
- Every number is the sum of the two above it.
- The second diagonal is the counting numbers.
- The sum of each row is a power of two.

Commentary:

Some undergraduates may already know that $\binom{n}{k}$ is entry k of row n of the arithmetic triangle, but others may not. Problem 3 of the Class Activity and Problem 1 of the Homework will lead undergraduates to establish this pattern in the triangle.

(b) Generate the next row of the triangle using some or all of these patterns.

Solution:

1 7 21 35 35 21 7 1

Commentary:

From our experience, undergraduates will generally use the "Every number is the sum of the two above it" pattern to generate this row.

We have found that many undergraduates are familiar with this triangle from high school and often remember it by the name "Pascal's triangle." Discuss the history of this triangle and how other mathematicians also discovered this triangular array of numbers. This discussion may include the following ideas (see also Ensley & Crawley (2006); Wilson & Watkins (2013)).

- **Mathematicians who have discovered the triangle**
  - The French mathematician/philosopher **Blaise Pascal** (1623–1662) wrote his Traité du triangle arithmétique (Treatise on the Arithmetic Triangle) in 1654, but this familiar triangular array of numbers was far from unknown to the world at this time. Even within Europe, Pascal's treatise was built up over several generations from correspondence and competition, just as mathematics is developed to this day.
  - The Indian mathematician **Halayudha** (ca 975) documents applications of the arithmetic triangle going back as far as the 2nd century in India.
  - In Iran the triangle is referred to as the **Khayyam triangle** after Persian poet Omar Khayyám (1048–1131), though there is evidence of earlier knowledge there, too.
  - The story is similar in the Chinese mathematics tradition, where the triangle is named for **Yang Hui** (1238–1298) even though it was known at least 200 years earlier in China.
- **Why it is also referred to as the "arithmetic triangle" instead of only by a person's name**
  The challenges of ancient communication and the scarcity of existing records guarantee there is no way to identify a single person who discovered this triangular array of numbers, so it is also referred to as "the arithmetic triangle," which is what Pascal called it in the treatise that he did not live to see published.

- **How it arose in different cultures**
  More interesting than conversations about intellectual provenance is the wide variety of reasons the arithmetic triangle arose in these cultures:
  - The early Indian and Persian work was focused on combinatorial aspects, some motivated by patterns in chants and poetry.
  - The Chinese motivation included methods for approximating roots of numbers, related to what we now call the binomial theorem.
  - And Pascal famously corresponded with Pierre de Fermat on problems involving gambling and probability, providing yet another application for the triangle entries.

To conclude the discussion of Problem 1, emphasize the following connection to teaching.

Discuss This Connection to Teaching

Secondary school teachers are called on by many mathematics content standards (e.g., CCSSM) to help students see how patterns in the arithmetic triangle can be used when computing coefficients in a binomial expansion.

Problems 2 and 3 focus undergraduates' attention on a hypothetical student, Anton, who is investigating why the patterns in the arithmetic triangle are the same as the patterns of coefficients in binomial expansions. Anton's work lays the foundation for the combinatorial reasoning on which the Class Activity is built. For the Pre-Activity review, it is sufficient for undergraduates to notice the patterns; they will return to developing explanations in the Class Activity. Give undergraduates a few minutes to review the context of Anton's work.

Context for Pre-Activity Problem 2

In high school, Anton learned that binomial expansions can be computed using the arithmetic triangle, and he wants to investigate why the numerical pattern in the arithmetic triangle emerges from the algebraic process of computing powers of a binomial. As part of his investigation, Anton expanded each of the following expressions using the distributive property of multiplication over addition, without simplifying the expressions using the commutative or associative properties of multiplication. Below are the expansions he computed, and all of his computations are correct.

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y)\\ &= xx + xy + yx + yy\end{aligned}$$

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)(x+y)\\ &= xxx + xxy + xyx + xyy + yxx + yxy + yyx + yyy\end{aligned}$$

$$\begin{aligned}(x+y)^4 &= (x+y)(x+y)(x+y)(x+y)\\ &= xxxx + xxxy + xxyx + xxyy + xyxx + xyxy + xyyx + xyyy + yxxx + yxxy\\ &\quad + yxyx + yxyy + yyxx + yyxy + yyyx + yyyy\end{aligned}$$

Next, Anton looked for patterns in the types of each term, making a table listing the terms with the following attributes.

$(x + y)^2$

Two x's and Zero y's	One x and One y	Zero x's and Two y's
xx	xy yx	yy

$(x + y)^3$

Three x's and Zero y's	Two x's and One y	One x and Two y's	Zero x's and Three y's
xxx	xxy xyx yxx	xyy yxy yyx	yyy

$(x + y)^4$

Four x's and Zero y's	Three x's and One y	Two x's and Two y's	One x and Three y's	Zero x's and Four y's
$xxxx$	$xxxy$ $xxyx$ $xyxx$ $yxxx$	$xxyy$ $xyxy$ $xyyx$ $yxxy$ $yxyx$ $yyxx$	$xyyy$ $yxyy$ $yyxy$ $yyyx$	$yyyy$

Before working on Problem 2, ask undergraduates what method they think Anton used to expand $(x + y)^2$ and if this method works for all polynomial multiplication and discuss the following connection to teaching. Undergraduates will likely say they used "FOIL," which is a commonly-used secondary school mnemonic to remind students that when multiplying binomials they multiply the First, Outside, Inside, and Last pairs of terms.

Discuss This Connection to Teaching

Secondary students typically use the FOIL method to expand binomial products such as $(x + y)^2$. However, they often view the FOIL method as a procedure to memorize and use it for **all** polynomial multiplication, which can lead to computational errors when there are more than four terms in the resulting product. This can interfere with sense-making, because students ought to make sense of the procedure, understand when it can and cannot be used, and understand that the FOIL method is a repeated application of the distributive property. Prospective teachers who also understand these nuances can help their future students.

Pre-Activity Problem 2

2. Anton first looked for patterns in the total number of terms in each of his expansions. He noticed that his expansion of $(x + y)^2$ has 4 terms, his expansion of $(x + y)^3$ has 8 terms, and his expansion of $(x + y)^4$ has 16 terms. Without multiplying everything out and counting, how can you use the pattern Anton noticed to predict the number of terms that $(x + y)^5$ would have?

Sample Responses:

- Anton's number of terms in each expansion is 2^n where n is the value of the exponent. So if Anton expanded $(x+y)^5$, then there would be $2^5 = 32$ terms.
- I notice the pattern is 2^n, where n is the row of the arithmetic triangle, so $(x+y)^5$ would have $2^5 = 32$ terms.

Pre-Activity Problem 3

3. In his tables, Anton observed that the number of expressions of each type correspond to entries of the arithmetic triangle. For example, in his first table for $(x+y)^2$ he counted the following number of terms in each category which corresponded to Row 2 of the arithmetic triangle.

Two x's and Zero y's	One x and One y	Zero x's and Two y's
xx	xy	yy
	yx	
(1)	(2)	(1)

Use the pattern he found to conjecture why it is customary that the rows and entries of the arithmetic triangle are numbered beginning with 0.

Sample Responses:

- Because $(x+y)^0 = 1$ and the only value in the row numbered 0 of the arithmetic triangle is 1.
- Because the pattern is easier to identify if the row of the triangle labeled, for example, "row 4" corresponds to $(x+y)^4$.

Class Activity: Problems 1–4 (20 minutes)

Pass out **Problems 1–4** of the Class Activity.

Instruct undergraduates to work in groups on Problems 1 and 2 and then discuss the solutions, asking undergraduates to share their work when appropriate. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion.

Class Activity Problem 1

1. Consider a string of four letters made up of only the letters x or y.

— — — —

(a) How many such strings are there? Explain your reasoning.

Sample Responses:

- $2^4 = 16$, because I have two choices (x or y) for each blank.
- Using the multiplication principle, there are $2 \cdot 2 \cdot 2 \cdot 2 = 16$ distinct strings.

(b) List the strings that contain exactly three x's. How many are there? Explain using combinatorial reasoning.

The strings are

$$xxxy \qquad xxyx \qquad xyxx \qquad yxxx$$

Sample Responses:

- 4, because $\binom{4}{3} = 4$. There are 4 blanks and we are choosing 3 of them to be an x, and the order of the x's doesn't matter.
- 4, because $\binom{4}{1} = 4$. There are 4 blanks and if we choose 1 of them to be a y, the other 3 are x, and the order of the x's doesn't matter.

(c) List the strings that contain exactly one y. How many are there? Explain using combinatorial reasoning.

The strings are

$$xxxy \qquad xxyx \qquad xyxx \qquad yxxx$$

Sample Responses:

- 4, because $\binom{4}{1} = 4$. There are 4 blanks and we are choosing 1 blank to be a y.
- 4, because placing 1 y and then 3 x's is the same as placing 3 x's and then 1 y, and we can choose exactly 3 blanks to place an x with $\binom{4}{3} = 4$.

(d) List the strings that contain exactly three y's. How many are there? Explain using combinatorial reasoning.

The strings are

$$xyyy \qquad yxyy \qquad yyxy \qquad yyyx$$

Sample Responses:

- 4, because $\binom{4}{3} = 4$. There are 4 blanks and we are choosing 3 of them to place a y, and the order of the y's doesn't matter
- $\binom{4}{1} = 4$ or $\binom{4}{3} = 4$, depending on whether I think of it as placing 3 y's and then 1 x or as placing 1 x and then 3 y's.

Commentary:

If undergraduates do not specifically identify the symmety in their answers, facilitate a discussion that highlights that if you choose 1 of the 4 blanks to place an x, then you are also specifying 3 of the 4 blanks to place a y, for example.

For Problems 2 & 3, refer to Anton's table from the Pre-Activity.

Class Activity Problem 2

2. Reexamine the table Anton created in his quest to understand binomial patterns. Notice the terms in his expansion of $(x+y)^4$ with three x's and one y are: $xxxy, xxyx, xyxx$, and $yxxx$. Also notice that using the context of Problem 1, Anton has listed all of the strings with three x's and one y.

(a) Explain to Anton why $\binom{4}{1}$ counts the number of such terms.

Sample Response:

These four are all of the possible strings that have exactly 1 y and 3 x's. This is what $\binom{4}{1}$ gives: how to place 1 character in a 4-character string and filling in the rest of the string with x's.

(b) Now, use associative and commutative properties of multiplication to generate like terms. Explain to Anton why the coefficient of x^3y in the expansion of $(x+y)^4$ is exactly the same as the number of terms that have three x's and one y.

Sample Response:

Each of these terms is equivalent to x^3y. Since there are four such terms, each occuring once, when you combine like terms, you get the coefficient of x^3y, which is 4, and logically, is $\binom{4}{1}$.

Problem 3 examines the $n = 4$ instance of $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$.

Class Activity Problem 3

3. Look again at Anton's table for the expansion of $(x+y)^4$. Note that the $(x+y)^4$ entries in the column labeled "Two x's and Two y's" are of two types: those that start with x and those that start with y. Explain how to generate each of these entries by starting with appropriate entries in the table for $(x+y)^3$.

Solution:

Those that start with x can be constructed by prepending an x on the $(x+y)^3$ entries in the column labeled "One x and Two y's," and those that start with y can be constructed by prepending a y on the $(x+y)^3$ entries in the column labeled "Two x's and One y."

Commentary:

Note that this explains the relation $\binom{4}{2} = \binom{3}{1} + \binom{3}{2}$ without the need to compute the values involved. This is the pattern used to construct the arithmetic triangle, providing evidence for the observation that the counting numbers $\binom{n}{k}$ are the values in the triangle. You can direct undergraduates to use similar reasoning to predict how many terms will be in a hypothetical $(x+y)^5$ table under the column heading, "Two x's and Three y's" by referencing Anton's table for $(x+y)^4$.

In Problem 4, undergraduates apply combinatorial reasoning to an expansion that is not yet in Anton's chart, moving toward the expansion of the general case $(x+y)^n$.

Class Activity Problem 4

4. Suppose Anton were to expand

$$(x+y)^5 = (x+y)(x+y)(x+y)(x+y)(x+y)$$

From his earlier work, he knows that he will have 5 terms in the expansion, one each corresponding to x^5, x^4y, x^3y^2, x^2y^3, xy^4, and y^5.

Determine the coefficients in the expansion, and explain to Anton how you determined the coefficients.

Sample Responses:

I used combinations. $\binom{5}{5} = 1$, $\binom{5}{4} = 5$, $\binom{5}{3} = 10$, $\binom{5}{2} = 10$, $\binom{5}{1} = 5$, and $\binom{5}{0} = 1$. We are choosing how to arrange strings of length five and specifying that all 5 are x's, then that 4 are x's, then 3, 2, 1, and 0.

Commentary:

Undergraduates will usually write $\binom{5}{5} = 1$ instead of $\binom{5}{0} = 1$ as the first coefficient. Asking them to explain what that combination represents in the context of the problem will highlight whether they are choosing the characters in the string to be x's or choosing them to be y's. The discussion of symmetry from Problem 1 is useful for moving from undergraduates' observations about the coefficients based on counting arguments to the statement of the binomial theorem as it is usually written.

Advice on Teaching the Lesson over Two Days

If you are teaching the lesson over two class periods, this can be an effective place to stop for the day. At this point in the lesson, you can demonstrate that the arithmetic triangle can be rewritten using combinatorics, as follows, emphasizing that the triangle can be thought of as displaying the results of computing combinations, and undergraduates can complete Problem 1 from the homework, which provides a generalization of the observation from Problem 3 in the Class Activity.

$$
\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}
$$

See Chapter 1 for guidance on using exit tickets to facilitate instruction in a two-day lesson.

Class Activity: Problems 5–6 (25 minutes)

If you assigned Problem 1 from the homework, you can discuss that result to initiate the discussion.

Distribute **Problems 5 and 6** of the Class Activity. We suggest you discuss the solution to Problem 5, then formally state the binomial theorem, and conclude by discussing the solutions to Problem 6.

Class Activity Problem 5

5. **Expanding Binomial Products.**
Use combinatorial notation to write the expansion of $(x+y)^n$. Explain how you determined this is the appropriate expression.

Sample Response:

$$(x+y)^n = \binom{n}{n}x^n y^0 + \binom{n}{n-1}x^{n-1}y^1 + \cdots + \binom{n}{1}x^1 y^{n-1} + \binom{n}{0}x^0 y^n$$

The number of strings made up of k x's and $n-k$ y's is $\binom{n}{k}$.

Commentary:
Undergraduates will usually write the formula as shown above, where the first coefficient is $\binom{n}{n}$ rather than $\binom{n}{0}$. Prompt undergraduates to notice the alternative

$$(x+y)^n = \binom{n}{0}x^n y^0 + \binom{n}{1}x^{n-1}y^1 + \cdots + \binom{n}{n-1}x^1 y^{n-1} + \binom{n}{n}x^0 y^n,$$

focusing on symmetry in the arithmetic triangle and the equivalence of the algebraic expressions for $\binom{n}{n}$ and $\binom{n}{0}$.

State the binomial theorem and point out that $\binom{n}{k}$ is called a "binomial coefficient." If you usually ask your students to prove the binomial theorem, assign them to groups to generate an outline of a proof of the binomial theorem based on the combinatorial arguments they have developed thus far. Alternatively, if you prefer to provide a proof in class, we have found that the Class Activity has laid the foundation for undergraduates to understand an instructor-led proof based on combinatorial arguments.

Theorem 1 (Binomial Theorem)**.** *The expansion of the binomial $(x+y)^n$ to an integer power $n \geq 1$ is given by*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

Outline of proof. Consider the expansion of $(x+y)^n$, which will consist of a sum of terms of the form $x^{n-k}y^k$, for each integer $0 \leq k \leq n$. For each k, the term $x^{n-k}y^k$ appears in the sum the same number of times as there are strings of length n with $(n-k)x$'s and k y's, which is $\binom{n}{k}$. Combining like terms yields

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

□

Discuss the following connection to teaching, which emphasizes how the binomial theorem is used in secondary mathematics.

Discuss This Connection to Teaching

Expanding binomial products is fundamental to school mathematics, and the binomial theorem is typically taught in Intermediate Algebra as a core content standard. School students often use this theorem to expand binomials to a power higher than 2. Many school students memorize the binomial theorem and rely on the arithmetic triangle to derive the necessary coefficients. Combinatorial reasoning helps undergraduates to understand and express the algebraic patterns and coefficients of the binomial expansions. All undergraduates should understand the combinatorial reasoning that underlies the binomial theorem in order to develop a thorough understanding of binomial expansion.

Problem 6 contains two parts that are representative of the kind of problems found in both high school and undergraduate textbooks that ask students to apply the binomial theorem. We have found that undergraduates will more often use an algebraic approach on Problem 6(a) (finding the coefficient of the term that contains x^7), unless they are prompted to use a combinatorial approach.

Class Activity Problem 6

6. **Applying the Binomial Theorem.**

   (a) What is the coefficient of the term that contains x^7 in the expansion of $(x+4y)^{10}$? Explain how you determined this.

   Sample Responses:

   - Using the binomial theorem, the term that contains x^7 in it is $\binom{10}{7}(x)^7(4y)^3$. Thus, the coefficient is $\binom{10}{7}(1)^7(4)^3 = (120)(1)(64) = 7,680$
   - We need exactly 7 of the 10 binomials to contribute an x and this can be done in $\binom{10}{7} = 120$ ways. Then because the coefficient of the x is 1, we'd have $(1)^7 = 1$. If the exponent of the x is 7, then the exponent of the y must be $10 - 7 = 3$ and since there is a 4 in front of the y, we'd also have $4^3 = 64$ as part of the coefficient. All together, the coefficient is $(120)(1)(64) = 7680$.

(b) Use the binomial theorem to expand $(3x-2y)^5$. Show all of your work.

Solution:

$$\begin{aligned}(3x-2y)^5 &= \binom{5}{5}(3x)^5(-2y)^0 + \binom{5}{4}(3x)^4(-2y)^1 + \binom{5}{3}(3x)^3(-2y)^2 + \binom{5}{2}(3x)^2(-2y)^3 \\ &\quad + \binom{5}{1}(3x)^1(-2y)^4 + \binom{5}{0}(3x)^0(-2y)^5 \\ &= (1)(243x^5)(1) + (5)(81x^4)(-2y) + (10)(27x^3)(4y^2) + (10)(9x^2)(-8y^3) \\ &\quad + (5)(3x)(16y^4) + (1)(1)(-32y^5) \\ &= 243x^5 - 810x^4y + 1080x^3y^2 - 720x^2y^3 + 240xy^4 - 32y^5\end{aligned}$$

Class Activity: Problems 7 & 8 (15 minutes)

Pass out **Problems 7 and 8** of the Class Activity. Before instructing undergraduates to work on these problems, discuss the following connection to teaching.

Discuss This Connection to Teaching

Problems 7 and 8 focus on analyzing hypothetical students' thinking in order to develop undergraduates' skills in understanding school student thinking and developing questions to guide school students' understanding. All undergraduates (especially prospective teachers) should explore how others use, reason with, and communicate mathematics. These problems also give prospective teachers (and tutors and future graduate students) an opportunity to think about how they would respond to student work in ways that nurture students' assets and understanding and in ways that help develop a students' mathematical understanding.

Problem 7 emphasizes two different perspectives that depend on a specific frame of reference—choosing blanks to place an x or choosing blanks to place a y. The goal of this problem is to help undergraduates see how multiple perspectives can arise in student work. (See Benson et al., 2005, p. 182, for more examples of problems about combinatorial thinking based on hypothetical student work.)

Class Activity Problem 7

7. Evelyn and Ivy were working on Problem 4 in the Class Activity where they determined the coefficients in the expansion of $(x+y)^5$. Evelyn says that the coefficient of x^3y^2 is $\binom{5}{3}=10$ because she was counting the ways to place 3 x's in a five-character string. Ivy claims that the coefficient is $\binom{5}{2}=10$ since she was counting ways to place 2 y's in a five-character string. Write two questions you could ask Evelyn and Ivy to help them find the common ground between their distinct approaches. Explain how your questions might help them.

   Sample Responses:

   - Is there a difference between counting the ways to place 3 x's versus counting the ways to place 2 y's? Why do both $\binom{5}{3}$ and $\binom{5}{2}$ equal 10? What is the coefficient of x^2y^3? These questions will help them see both are correct and you could view it as counting the ways to place x's or counting the ways to place y's, as long as we are specific in what we are counting. It also helps them to recognize the symmetry in the coefficients of their terms.
   - What variable do you want in the strings you are choosing? How could choosing 3 ways to place an x be similar to choosing 2 ways to place a y? The idea of these questions would be to get them to see that if you choose 3 ways to place an x, then the other blanks have to be y's.

- If you count the strings with three x's how many places in the string would we need to fill with y's? If you count the strings with two y's, how many places in the string would we need to fill with x's? These questions will help show that both answers are correct from different perspectives and all elements of the string are still filled.

Commentary:
Posing questions to a hypothetical student may be new and challenging for your undergraduates. If they are stuck, it may be helpful to ask them to first discuss what Evelyn and Ivy understand based on their work. You may also need to remind undergraduates to explain how their questions might help guide Evelyn's and Ivy's understanding.

Problem 8 contains a hypothetical student, Henry, who makes a mistake when applying the binomial theorem. Emphasize the purpose of Problem 8, as follows:

- We provide this example of student work because it illustrates a common error students make when applying the binomial theorem.
- As mathematicians, it is useful to examine someone else's mathematical thinking and to be able to explain any errors that occur. This process helps deepen our own mathematical understanding.
- Prospective teachers need to examine student thinking and reflect on how they can respectfully resolve mathematical disputes because these are skills they will use in their future careers.

Class Activity Problem 8 : Parts a & b

8. Henry, a high school student, expanded $(2x - y)^4$ using the binomial theorem and made some errors. Below is his work.

$$(2x - y)^4 = 2x^4 - 8x^3y - 12x^2y^2 - 8xy^3 - y^4$$

(a) What does Henry understand about the binomial theorem?

Sample Responses:

- He understands the pattern of the exponents—that they decrease/increase by 1.
- He has the correct exponents on the variables.
- He knows the coefficients from Row 4 of the arithmetic triangle are 1, 4, 6, 4, 1, and then he multiplies the coefficient by 2 when the term includes an x and also multiplies the coefficient by -1 when the term includes a y.

(b) What does Henry not yet understand about the binomial theorem?

Sample Responses:

- Henry wasn't consistent in including the 2 with the x and the -1 with the y when taking powers. He multiplied 2 by each of the coefficients in the arithmetic triangle. He forgets to apply the power to the 2 and -1.
- He doesn't yet understand that the coefficients in his final answer should alternate between positive and negative.
- Henry doesn't understand that when he raises something to an even power the resulting coefficients will always be positive, not negative.

Commentary:
Henry's work demonstrates two common errors many students make when using the binomial theorem (incorrectly applying powers to negative signs and applying powers only to the variables in a term and not to the coefficients of the term as well). The correct answer is $16x^4-32x^3y+24x^2y^2-8xy^3+y^4$, and we have found that most undergraduates will recognize that the errors revolve around the coefficients. Undergraduates may also use the term "coefficient" to mean the binomial coefficients $(1, 4, 6, 4, 1)$ as well as the coefficients in Henry's solution $(2, -8, -12, -8, -1)$, without distinguishing between the

two. They may say things like "He got the coefficients wrong." and "He understands how to get the coefficients from the arithmetic triangle." If this occurs, ask undergraduates to clarify which coefficients they are referring to.

Problem 8(c) asks undergraduates to evaluate a set of pre-written questions one might ask Henry about his work. Asking undergraduates to evaluate a set of questions is a way to scaffold their understanding of guiding school students' understanding and to help them develop their skills of writing questions on their own.

Class Activity Problem 8 : Part c

(c) Consider the following questions that someone might ask Henry about his work.

i. Explain how the following question could help Henry to advance in his understanding of the binomial theorem:

How is $(2x - y)^4$ *similar to* $(x + y)^4$ *and how is it different?*

Sample Responses:

- This will help Henry realize that $-y$ is really $+(-y)$.
- This can help Henry realize that **all** of $2x$ (both the 2 and x) need to be raised to powers in the binomial theorem, not just the x.

ii. Explain how the following question can help you assess what Henry understands about the binomial theorem:

Why doesn't $(-3y)^2 = -3y^2$*?*

Sample Responses:

- Henry's seems to be incorrectly applying the powers in the binomial theorem. It seems like he may have written $-2x^3$ instead of $(-2x)^3$ and that is how he got his error. Asking this question will help us know if he wrote out the binomial theorem expansion without the parentheses or if he doesn't know how to distribute the power to both the constant and variable.
- This question helps us understand if Henry's error is in using the binomial theorem or if it's a computation error.

iii. Explain why the following question would not help Henry:

Do the exponents and the coefficients look right?

Sample Responses:

- Henry probably thinks his answer is right, so asking this question isn't helpful.
- This would be like saying "your error is in the coefficients" and it is kind of telling Henry where the problem is instead of guiding him to figure out where his error is on his own.
- Henry's exponents in his final answer are correct so this doesn't help him understand that he is not applying the exponents correctly to the coefficients.

Wrap-Up (5 minutes)

Conclude the lesson by discussing the connections between the arithmetic triangle and the binomial theorem, emphasizing how undergraduates have used combinatorial reasoning throughout the lesson. This may include the following ideas:

- As high school students, many of the undergraduates were taught that the binomial theorem was related to the arithmetic triangle, but this relationship may not have been thoroughly explained.

- The purpose of this lesson is to use combinatorial reasoning to explain the connection between the arithmetic triangle and the binomial theorem.
- Using combinatorial reasoning is central to understanding the arithmetic triangle. For the prospective teachers in the class—high school students would not be asked to prove the binomial theorem, but prospective teachers should understand this connection and be able to show this.

You can ask undergraduates to complete an exit ticket, if you choose. See Chapter 1 for guidance on using exit tickets.

Homework Problems

At the end of the lesson, assign the following homework problems. Assign any additional homework problems at your discretion.

Homework Problem 1

1. Generalize the example from Problem 3 of the Class Activity to fill in the proof sketch for the following statement.

   Proposition. The counting numbers $\binom{n}{k}$ satisfy the arithmetic triangle pattern

   $$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

   for all $n \geq k \geq 1$.

   Proof sketch. Each of the $\binom{n}{k}$ arrangements of k x's and $n-k$ y's has one of the following forms:

   (i) An x followed by an arrangement of $k-1$ x's and $n-k$ y's; or
   (ii) A y followed by an arrangement of k x's and $n-k-1$ y's

   The number of arrangements of type (i) is ______ and the number of arrangements of type (ii) is ______, so . . .

   Solution:

   Proof. Each of the $\binom{n}{k}$ arrangements of k x's and $n-k$ y's has one of the following forms:

   (i) An x followed by an arrangement of $k-1$ x's and $n-k$ y's; or
   (ii) A y followed by an arrangement of k x's and $n-k-1$ y's

   The number of arrangements of type (i) is $\binom{n-1}{k-1}$ and the number of arrangements of type (ii) is $\binom{n-1}{k}$, so the total number of arrangements is

   $$\binom{n-1}{k} + \binom{n-1}{k-1},$$

Problem 2 addresses the commonly-held algebraic misconception that $(x+y)^2 = x^2 + y^2$. High school students will have likely seen area models that can be used to illustrate why this is not the case. Having undergraduates also connect their investigations with the binomial theorem to expanding binomials provides another approach to address the conception many school students have when they use this flawed equation.

Homework Problem 2

2. Students with a not-yet-complete understanding of high school algebra commonly argue that $(x+y)^2 = x^2 + y^2$. Below are three different approaches a teacher can use to help students see that this is not true.

   - One approach is to choose particular values for a and b, say $a = 1$ and $b = 2$, and ask students to plug those values into the left-hand side of the equation and simplify and then plug those same values into the right-hand side of the equation and simplify. Students can compare their answers to see that they are not equal. This approach can quickly show students that $(x + y)^2 \neq x^2 + y^2$ but doesn't give them insight into why the equality does not hold.
   - Another approach is to create an area model, such as the one below, which shows students that $(x + y)^2 \neq x^2 + y^2$. In this case the binomial $x + y$ is represented as a segment of length $x + y$, and the product $(x + y)^2$ is represented as the area of a rectangle whose side lengths are both $x + y$. This representation shows students which mathematical components they are missing in their answer, specifically the term $2xy$.

	x	y
x	x^2	xy
y	yx	y^2

   - A third approach is to apply the binomial theorem to $(x + y)^2$.

(a) Use the binomial theorem to explain to a high school student why $(x + y)^2 \neq x^2 + y^2$.

Solution:

The binomial theorem provides a formula that can be used to expand a binomial to the second power. The expansion will contain three terms, an x^2 term, a y^2 term, and an xy term, as follows. It is because of the xy term that $(x + y)^2$ is not equal to $x^2 + y^2$.

$$\begin{aligned}(x+y)^2 &= \binom{2}{2}x^2y^0 + \binom{2}{1}x^1y^1 + \binom{2}{0}x^0y^2\\ &= (1)(x^2)(1) + (2)(x)(y) + (1)(1)(y^2)\\ &= x^2 + 2xy + y^2\\ &\neq x^2 + y^2\end{aligned}$$

(b) Compare the "area model" approach with the "binomial theorem" approach. For what kinds of problems would each approach work? What insight does each approach highlight that will help students understand why $(x + y)^2 \neq x^2 + y^2$?

Sample Responses:

- Both the "area model" and the "binomial theorem" approach can be used when expanding binomial products. The area model works when multiplying two binomials, so for example computing $(x+y)^3$ would require two separate uses of the area model. The binomial theorem can be used to expand a binomial to any power.
- While the binomials being multiplied together need to be the same to use the binomial theorem, the area model can multiply two different binomials together, such as $(x + y)(2x - 1)$.

- Each approach highlights the $2xy$ term that is missing when students think $(x+y)^2 = x^2+y^2$. In the area model approach, students can see how each piece of the binomial contributes an x and a y and where the two xy pieces come from.

Problem 3 provides undergraduates an opportunity to examine hypothetical student work and write questions that will help guide the student's mathematical understanding.

Homework Problem 3

3. Charlie incorrectly expanded $(x+y)^2$ as follows.

$$\begin{aligned}(x+y)^2 &= 1x^2y^2 + 2x^1y^1 + 1x^0y^0 \\ &= x^2y^2 + 2xy + 1\end{aligned}$$

(a) What does Charlie understand about the binomial theorem?

Sample Responses:

- Charlie understands the binomial coefficients.
- They correctly understand the coefficients of the x's and that they decrease by 1 each time.

(b) What does Charlie not yet understand about the binomial theorem?

Sample Responses:

- They don't yet understand how the exponents in the binomial theorem work. They give both x and y the same power. However, the power of x plus the power of y should add to 2 for every term. Instead of having 2 x's and no y's for the first term they have 2 of each. The exponents on their second term are correct but their reasoning for it may not be correct. For their last term they have 0 x's and 0 y's but should have 0 x's and 2 y's.
- Charlie doesn't yet understand the pattern in the exponents. They think that exponents count down from n at the same time on both variables.

(c) Write at least two questions that will help Charlie revise their work and develop a deeper understanding of the binomial theorem. Explain how those questions will help Charlie.

Sample Responses:

- Use the distributive property to multiply $(x+y)(x+y)$. How does that answer compare to your original answer? Do we have terms where there are no x's and no y's? These questions will help Charlie recognize where the errors in their work are.
- Plug in $x = 3$ and $y = 4$ into both sides of your equation. What do you notice and what does this tell you? Hopefully Charlie will get $49 = 169$ and realize that their expansion is not correct.
- Can you describe how you got 0 as an exponent in your last term? This will help me understand how Charlie is applying the binomial theorem.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Problem 1 assesses undergraduates' use of combinatorial reasoning to explain how to determine a coefficient of a particular term in a binomial expansion.

Assessment Problem 1

1. Consider the expansion of $(x+y)^{10}$.

   (a) Beyond "because the binomial theorem says so," explain why the coefficient of the x^4y^6 term is $\binom{10}{6}$.

   Solution:

   There are two separate points to be made. Putting A and B together leads to a complete explanation.

   A. When there are 4 x's and 6 y's available, there are $\binom{10}{6}$ ways they can be arranged in a line. This is because of the 10 spots in the line available, you choose which 6 spots should get a y (and then the x's must go everywhere else without any more choices being made).

   B. The distributive property tells us that when multiplying out $(x+y)^{10} = (x+y)(x+y)(x+y)\ldots(x+y)$, we choose either an x or a y from each of the 10 binomial terms. Those choices that result in 4 x's and 6 y's will each contribute one to the coefficient of x^4y^6 in the expansion.

   (b) One student states that the coefficient of the x^4y^6 term is $\binom{10}{4}$, and a second student states it is $\binom{10}{6}$. Beyond "because of symmetry," explain why each of them correctly computes the coefficient of x^4y^6.

   Solution:

   - In part A of the explanation above, we could have taken the position that of the 10 spots in the line available, you choose which 4 spots should get an x (and then the y's must go everywhere else without any more choices being made).
   - If you have 10 things and each has either x or y written on it and you know that only 4 of them have x written on them, then the other 6 must have y written on them as those were the only two options. If you instead knew that only 6 had a y written on them, you would also know that 4 had an x written on them as it was the only other option. Either way you know the same thing: 4 have x and 6 have y.

Problem 2 assesses how undergraduates analyze hypothetical student work and how they can write questions to guide the student's understanding.

Assessment Problem 2

2. Tencha, a high school student, expanded $(x-3y)^4$ using the binomial theorem and made some errors. Below is their work.

$$(x-3y)^4 = x^4 - 12x^3y - 18x^2y^2 - 12xy^3 - 3y^4$$

   (a) What does Tencha understand about the binomial theorem?

   Sample Responses:

   - Tencha knows the statement of the theorem exactly right, and they will get the correct answer for any problem about $(x+y)^n$, like $(x+y)^3$.
   - That each term has exponents that add up to 4.

   (b) What does Tencha not yet understand about the binomial theorem?

   Sample Response:

   The correct work is

$$x^4 + 4x^3(-3y)^1 + 6x^2(-3y)^2 + 4x^1(-3y)^3 + (-3y)^4$$

We can see by comparison that Tencha doesn't yet understand that there need to be parentheses around the $-3y$ term when using the binomial theorem.

(c) Write two questions you can ask Tencha to help them revise their work. Explain how your questions could help guide their mathematical understanding.

Sample Responses:

- I would ask Tencha if $2x^2 = (2x)^2$. If Tencha can understand why $2x^2 \neq (2x)^2$, then they may be able to see why $-3y^4 \neq (-3y)^4$.
- Does the left hand side of your equation equal the right hand side of your equation when you plug a value in for x and y and simplify both sides of the equation? This will help them see that there is an error in their work somewhere.
- What happens when we multiply two negative numbers? How does this relate to your expansion? This can help Tencha remember that $(-)(-)$ becomes positive and that when we expand we are multiplying each term so some of the terms in their answer should be positive.

6.6 References

[1] Benson, S., Addingon, S., Arshavsky, N, Cuoco, A., Goldenberg, E. & Karnowski, A. (2005) *Ways to think about mathematics: Activities and investigations for grade 6–12 teachers.* Corwin Press.

[2] Ensley, D. E. & Crawley, J. W. (2006). *Introduction to discrete mathematics: Mathematical reasoning with puzzles, patterns, and games.* John Wiley and Sons.

[3] Epp, S. S. (2011). *Discrete mathematics: An introduction to mathematical reasoning* (Brief Edition). Brooks/Cole Publishing Co.

[4] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from `http://www.corestandards.org/`

[5] Wilson, R., & Watkins, J. J. (Eds.). (2013). *Combinatorics: Ancient and modern*. (Especially Chapter 7, "The arithmetical triangle" by A. W. F. Edwards). Oxford University Press.

6.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** BINOMIAL THEOREM (page 1 of 3)

1. Below are the first seven rows of the *arithmetic triangle*, also known as Yang Hui's triangle or Pascal's triangle, among other names. By custom, the rows and entries are numbered starting at 0.

$$\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}$$

(a) Write down at least three patterns that you observe.

(b) Generate the next row of the triangle using some or all of these patterns.

PRE-ACTIVITY: BINOMIAL THEOREM (page 2 of 3)

In high school, Anton learned that binomial expansions can be computed using the arithmetic triangle, and he wants to investigate why the numerical pattern in the arithmetic triangle emerges from the algebraic process of computing powers of a binomial. As part of his investigation, Anton expanded each of the following expressions using the distributive property of multiplication over addition, without simplifying the expressions using the commutative or associative properties of multiplication. Below are the expansion he computed, and all of his computations are correct.

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y)\\ &= xx+xy+yx+yy\end{aligned}$$

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)(x+y)\\ &= xxx+xxy+xyx+xyy+yxx+yxy+yyx+yyy\end{aligned}$$

$$\begin{aligned}(x+y)^4 &= (x+y)(x+y)(x+y)(x+y)\\ &= xxxx+xxxy+xxyx+xxyy+xyxx+xyxy+xyyx+xyyy+yxxx+yxxy\\ &\quad +yxyx+yxyy+yyxx+yyxy+yyyx+yyyy\end{aligned}$$

Next, Anton looked for patterns in the types of each term, making a table listing the terms with the following attributes.

$(x+y)^2$

Two x's and Zero y's	One x and One y	Zero x's and Two y's
xx	xy	yy
	yx	

$(x+y)^3$

Three x's and Zero y's	Two x's and One y	One x and Two y's	Zero x's and Three y's
xxx	xxy	xyy	yyy
	xyx	yxy	
	yxx	yyx	

$(x+y)^4$

Four x's and Zero y's	Three x's and One y	Two x's and Two y's	One x and Three y's	Zero x's and Four y's
$xxxx$	$xxxy$	$xxyy$	$xyyy$	$yyyy$
	$xxyx$	$xyxy$	$yxyy$	
	$xyxx$	$xyyx$	$yyxy$	
	$yxxx$	$yxxy$	$yyyx$	
		$yxyx$		
		$yyxx$		

2. Anton first looked for patterns in the total number of terms in each of his expansions. He noticed that his expansion of $(x+y)^2$ has 4 terms, his expansion of $(x+y)^3$ has 8 terms, and his expansion of $(x+y)^4$ has 16 terms. Without multiplying everything out and counting, how can you use the pattern Anton noticed to predict the number of terms that $(x+y)^5$ would have?

3. In his tables, Anton observed that the number of expressions of each type correspond to entries of the arithmetic triangle. For example, in his first table for $(x+y)^2$ he counted the following number of terms in each category which corresponded to Row 2 of the arithmetic triangle.

Two x's and Zero y's	One x and One y	Zero x's and Two y's
xx	xy	yy
	yx	
(1)	(2)	(1)

Use the pattern he found to conjecture why it is customary that the rows and entries of the arithmetic triangle are numbered beginning with 0.

NAME: **CLASS ACTIVITY:** BINOMIAL THEOREM (page 1 of 5)

1. Consider a string of four letters made up of only the letters x or y.

   — — — —

   (a) How many such strings are there? Explain your reasoning.

   (b) List the strings that contain exactly three x's. How many are there? Explain using combinatorial reasoning.

   (c) List the strings that contain exactly one y. How many are there? Explain using combinatorial reasoning.

   (d) List the strings that contain exactly three y's. How many are there? Explain using combinatorial reasoning.

2. Reexamine the table Anton created in his quest to understand binomial patterns. Notice the terms in his expansion of $(x+y)^4$ with three x's and one y are: $xxxy, xxyx, xyxx$, and $yxxx$. Also notice that using the context of Problem 1, Anton has listed all of the strings with three x's and one y.

   (a) Explain to Anton why $\binom{4}{1}$ counts the number of such terms.

   (b) Now, use associative and commutative properties of multiplication to generate like terms. Explain to Anton why the coefficient of x^3y in the expansion of $(x+y)^4$ is exactly the same as the number of terms that have three x's and one y.

3. Look again at Anton's table for the expansion of $(x+y)^4$. Note that the $(x+y)^4$ entries in the column labeled "Two x's and Two y's" are of two types: those that start with x and those that start with y. Explain how to generate each of these entries by starting with appropriate entries in the table for $(x+y)^3$.

4. Suppose Anton were to expand

$$(x+y)^5 = (x+y)(x+y)(x+y)(x+y)(x+y)$$

   From his earlier work, he knows that he will have 5 terms in the expansion, one each corresponding to x^5, x^4y, x^3y^2, x^2y^3, xy^4, and y^5.

   Determine the coefficients in the expansion, and explain to Anton how you determined the coefficients.

5. **Expanding Binomial Products.**
   Use combinatorial notation to write the expansion of $(x + y)^n$. Explain how you determined this is the appropriate expression.

6. **Applying the Binomial Theorem.**
   (a) What is the coefficient of the term that contains x^7 in the expansion of $(x + 4y)^{10}$? Explain how you determined this.

   (b) Use the binomial theorem to expand $(3x - 2y)^5$. Show all of your work.

7. Evelyn and Ivy were working on Problem 4 in the Class Activity where they determined the coefficients in the expansion of $(x+y)^5$. Evelyn says that the coefficient of x^3y^2 is $\binom{5}{3} = 10$ because she was counting the ways to place 3 x's in a five-character string. Ivy claims that the coefficient is $\binom{5}{2} = 10$ since she was counting ways to place 2 y's in a five-character string. Write two questions you could ask Evelyn and Ivy to help them find the common ground between their distinct approaches. Explain how your questions might help them.

8. Henry, a high school student, expanded $(2x-y)^4$ using the binomial theorem and made some errors. Below is his work.

$$(2x-y)^4 = 2x^4 - 8x^3y - 12x^2y^2 - 8xy^3 - y^4$$

   (a) What does Henry understand about the binomial theorem?

   (b) What does Henry not yet understand about binomial theorem?

(c) Consider the following questions that someone might ask Henry about his work.

i. Explain how the following question could help Henry to advance in his understanding of the binomial theorem:

How is $(2x - y)^4$ *similar to* $(x + y)^4$ *and how is it different?*

ii. Explain how the following question can help you assess what Henry understands about the binomial theorem:

Why doesn't $(-3y)^2 = -3y^2$*?*

iii. Explain why the following question would not help Henry:

Do the exponents and the coefficients look right?

NAME: **HOMEWORK PROBLEMS:** BINOMIAL THEOREM (page 1 of 1)

1. Generalize the example from Problem 3 of the Class Activity to fill in the proof sketch for the following statement.

   Proposition. The counting numbers $\binom{n}{k}$ satisfy the arithmetic triangle pattern

   $$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

   for all $n \geq k \geq 1$.

   Proof sketch. Each of the $\binom{n}{k}$ arrangements of k x's and $n-k$ y's has one of the following forms:

   (i) An x followed by an arrangement of $k-1$ x's and $n-k$ y's; or

   (ii) A y followed by an arrangement of k x's and $n-k-1$ y's

   The number of arrangements of type (i) is ______ and the number of arrangements of type (ii) is ______, so . . .

2. Students with a not-yet-complete understanding of high school algebra commonly argue that $(x+y)^2 = x^2 + y^2$. Below are three different approaches a teacher can use to help students see that this is not true.

   - One approach is to choose particular values for a and b, say $a = 1$ and $b = 2$, and ask students to plug those values into the left-hand side of the equation and simplify and then plug those same values into the right-hand side of the equation and simplify. Students can compare their answers to see that they are not equal. This approach can quickly show students that $(x+y)^2 \neq x^2 + y^2$ but doesn't give them insight into why the equality does not hold.
   - Another approach is to create an area model, such as the one below, which shows students that $(x+y)^2 \neq x^2 + y^2$. In this case the binomial $x+y$ is represented as a segment of length $x+y$, and the product $(x+y)^2$ is represented as the area of a rectangle whose side lengths are both $x+y$. This representation shows students which mathematical components they are missing in their answer, specifically the term $2xy$.

	x	y
x	x^2	xy
y	yx	y^2

   - A third approach is to apply the binomial theorem to $(x+y)^2$.

   (a) Use the binomial theorem to explain to a high school student why $(x+y)^2 \neq x^2 + y^2$.

   (b) Compare the "area model" approach with the "binomial theorem" approach. For what kinds of problems would each approach work? What insight does each approach highlight that will help students understand why $(x+y)^2 \neq x^2 + y^2$?

3. Charlie incorrectly expanded $(x+y)^2$ as follows.

   $$\begin{aligned}(x+y)^2 &= 1x^2y^2 + 2x^1y^1 + 1x^0y^0 \\ &= x^2y^2 + 2xy + 1\end{aligned}$$

   (a) What does Charlie understand about the binomial theorem?

   (b) What does Charlie not yet understand about the binomial theorem?

   (c) Write at least two questions that will help Charlie revise their work and develop a deeper understanding of the binomial theorem. Explain how those questions will help Charlie.

NAME: **ASSESSMENT PROBLEMS:** BINOMIAL THEOREM (page 1 of 2)

1. Consider the expansion of $(x+y)^{10}$.

   (a) Beyond "because the binomial theorem says so," explain why the coefficient of the x^4y^6 term is $\binom{10}{6}$.

   (b) One student states that the coefficient of the x^4y^6 term is $\binom{10}{4}$, and a second student states it is $\binom{10}{6}$. Beyond "because of symmetry," explain why each of them correctly computes the coefficient of x^4y^6.

Assessment Problems: Binomial Theorem (page 2 of 2)

2. Tencha, a high school student, expanded $(x-3y)^4$ using the binomial theorem and made some errors. Below is their work.

$$(x-3y)^4 = x^4 - 12x^3y - 18x^2y^2 - 12xy^3 - 3y^4$$

   (a) What does Tencha understand about the binomial theorem?

   (b) What does Tencha not yet understand about the binomial theorem?

   (c) Write two questions you can ask Tencha to help them revise their work. Explain how your questions could help guide their mathematical understanding.

7

Foundations of Divisibility

Discrete Mathematics or Introduction to Proof

Elizabeth A. Burroughs, *Montana State University*
Elizabeth G. Arnold, *Colorado State University*
Elizabeth W. Fulton, *Montana State University*
Douglas E. Ensley, *Shippensburg University*
Nancy Ann Neudauer, *Pacific University*

7.1 Overview and Outline of Lesson

The divisibility of integers is a typical early focus in courses that introduce undergraduate students to mathematical proof, primarily because integers are familiar compared to more abstract constructs, but also because school students begin their formal study of divisibility of integers as early as elementary school. This lesson guides undergraduates to provide explanations about divisibility properties of integers using the definition of divisibility, the quotient-remainder theorem, and knowledge about the base-ten representation of integers. The lesson provides a rich opportunity for "looking back and looking forward," as undergraduates examine methods and reasoning about divisibility that are used in elementary grades and make sense of these methods using number theoretic techniques by applying theorems common in undergraduate study.

The focus of the Class Activity in this lesson is to identify the distinction between **what it means** for an integer m to be divisible by an integer n and **how one can determine** by examining an integer m if it is divisible by a certain integer n (see Beckmann, 2018, Chapter 8).

A note about the terminology we use in this lesson: School students often refer to "numbers" instead of the more precise "natural numbers" or "integers." This lesson uses hypothetical school students as characters in various problems, and we use colloquial language when conveying their ideas. We use formal mathematical terminology when posing problems for undergraduates.

1. Launch—Pre-Activity
   Prior to the lesson, undergraduates complete a Pre-Activity where they consider how a fifth-grade student might approach justifying why an integer is or is not divisible by 5 by using the structure of the base-ten system. Undergraduates use integer division with remainders and the representation of integers in base ten to establish the well-used test for divisibility by 5.

2. Explore—Class Activity

   - *Problem 1—Divisibility by 3:*
     The proof for determining if an integer is divisible by 3 has a different structure than the proof for determining

if an integer is divisible by 5. Undergraduates will develop a pictorial and numerical argument about why a particular three-digit integer is not divisible by 3 in order to establish a test for divisibility by 3.

- *Problem 2—Divisibility by 4:*
  Undergraduates consider the advantages of two different tests for divisibility by 4.
- *Problem 3—Student Reasoning About Multiples:*
  Undergraduates practice explaining to a hypothetical student why some multiples of 3 are even and others are odd.

3. Closure—Wrap-Up
   Conclude the lesson by discussing how "looking back" at school students' reasoning about divisibility has enabled undergraduates to "look forward" to some tools of number theory and find important mathematical reasoning in school students' ideas. If you want undergraduates to prove any of the tests for divisibility, these can be assigned as homework.

7.2 Alignment with College Curriculum

The divisibility of integers is often taught in a discrete mathematics or introduction to proof course. A study of tests for divisibility offers an opportunity for undergraduates to apply direct proof techniques that rely on properties of the integers. This lesson may be appropriate as part of a lesson on the quotient-remainder theorem and comes after undergraduates have learned and applied the definition of divisibility of integers.

7.3 Links to School Mathematics

This lesson addresses several mathematical knowledge and practice expectations included in school standards documents, such as the Common Core State Standards for Mathematics (CCSSM, 2010). The place value system is central in K–12 mathematics, and understanding divisibility of the integers is fundamental to understanding prime factorization, greatest common factors, and least common multiples. This lesson guides undergraduates to provide explanations about divisibility properties of integers using the definition of divisibility, the quotient-remainder theorem, and knowledge about the base-ten representation of integers.

This lesson highlights:

- Proofs of various tests for divisibility;
- "Looking for and making use of structure," a K–12 mathematical practice (CCSSM, 2010), as undergraduates provide explanations for tests for divisibility that rely on the structure of the base-ten representation of integers.

7.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- Basic proof-writing techniques;
- The structure of an if-and-only-if proof;
- The definition of divisibility of integers;
- The quotient-remainder theorem.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Reason about mathematical arguments for tests for divisibility;
- Prove some tests for divisibility and explain why they work;

- Attend to the base-ten representation of integers in their proofs;
- Apply the quotient-remainder theorem in their proofs;
- Analyze hypothetical school student work that investigates different tests for divisibility and evaluate questions one might ask hypothetical students to help guide their understanding about divisibility.

Anticipated Length

One 50-minute class session.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity (print Problems 1 and 2–3 to pass out separately)
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on a quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

7.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to complete in preparation for the lesson, and ask undergraduates to bring their solutions to class on the day you start the Class Activity. The problem on the Pre-Activity sets the stage for the rest of the lesson, with a focus on making sense of school student reasoning using the tools of discrete mathematics.

Pre-Activity Review (10 minutes)

Before reviewing the solutions to the Pre-Activity as a class, ask undergraduates when they have used tests of divisibility and then discuss the following connection to teaching.

Discuss This Connection to Teaching

- Middle school and high school students might use tests of divisibility when working on problems such as factoring, finding the greatest common divisor, and finding the least common multiple, for example when operating with fractions.
- Students in middle school and high school often memorize tests for divisibility but do not understand why they work. Prospective teachers should understand how these tests for divisibility are the result of structured mathematical reasoning.

As needed, engage in a whole class discussion to examine Rickie's argument and propose a generalization of it. This discussion can focus on the distinction between using the definition of divisibility to prove whether or not a particular integer is divisible by 5, which is what Rickie has developed for the integer 243, and generating a proof of the test for divisibility by 5, which examines the ones digit of any integer written in base ten.

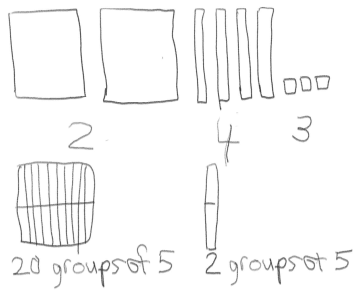
Pre-Activity

A fifth-grade class is discussing divisibility by 5. One student says, "My sister said you can tell if a number is divisible by 5 if it ends in a 0 or a 5." Another student adds, "Yeah, you can tell from skip counting by 5s: $5, 10, 15, 20, 25, \ldots$"

The teacher recognizes an opportunity to engage students in reasoning about properties of positive integers. The students have been using base-ten blocks to demonstrate whether various numbers are even or odd, so the teacher asks students to use their blocks to demonstrate whether the three numbers $243, 240$, and 245 are divisible by 5.

Notice that there is a difference between **how to recognize** if an integer is divisible by 5 or not, which involves inspecting the digit in the ones place, and **proving** that an integer is divisible by 5, which involves determining whether there exists an integer m such that the integer can be written as $5m$.

Below is what one student, Rickie, wrote to explain that 243 is not divisible by 5, but 240 and 245 are.



> *I represented 243 as 2 groups of 100, 4 groups of 10, and then 3 ones. Each group of 10 forms 2 groups of 5, and each group of 100 forms 20 groups of 5. Then you have to look at the leftover blocks.*
> *If the number ends in 0, then you haven't added any blocks and you can keep the groups you have. That's what happens with 240.*
> *If the number ends in 1, 2, 3, or 4, then you have leftover blocks. That's what happens with 243.*
> *If the number ends in 5, then you have 5 leftover blocks and you can put those into 1 group of 5. That's what happens with 245.*
> *I guess you could also have 6, 7, 8 or 9 leftover blocks, but that's it.*

State Rickie's test for divisibility by 5 as a biconditional statement. Explain how to generalize Rickie's argument about why the divisibility rule for 5 will always work.

Solution:
An integer is divisible by 5 if and only if the last digit of the integer is a 5 or a 0.
As a consequence of the quotient-remainder theorem, any integer n can be written as $n = 10q + r$, where q and r are integers and $0 \leq r \leq 9$. Because $10q = 5(2q)$ is always divisible by 5, divisibility of n by 5 is completely determined by r. If r is 0 or 5, r (and therefore n) is divisible by 5. Otherwise, r (and therefore n) is not divisible by 5.

Commentary:
Undergraduates' explanations may be much less formal than what we indicate here, and you can use their responses to this introductory problem to gauge how to focus the class discussions in this lesson. Rickie's assertion that you can list all of the possibilities for the single blocks in a number represented by base-ten blocks is a consequence of the quotient-remainder theorem (or integer division with remainders) upon which a formal proof relies. If undergraduates don't name this, you can address it with a brief reminder.

Theorem 1 (Quotient-Remainder Theorem)**.** *Given any integer n and a positive integer d, there exist unique integers q and r such that $n = dq + r$ and $0 \leq r < d$.*

After discussing Rickie's argument, emphasize that the divisibility test for 5 consists of a biconditional statement, so there are two statements that undergraduates must justify: "If the last digit of an integer is 0 or 5, then the integer is divisible by 5." and "If an integer is divisible by 5, then the last digit of the integer is either 0 or 5." Undergraduates should recognize that a formal proof requires that they establish the certainty that "divisibility of n by 5 is completely determined by r." They will have the chance to examine this further in the Class Activity and Homework Problems.

Emphasize that in the proof for divisibility by 5, undergraduates make use of the base-ten representation of integers and the ideas about integer division with remainders that are encapsulated in the quotient-remainder theorem. Discuss the following connection to teaching.

Discuss This Connection to Teaching

Examining the work from this fifth-grader is a way to introduce the representation of integers in the base-ten system that is both useful to undergraduates in writing proofs but that is also accessible to younger students. Younger students will explain tests for divisibility by considering base-ten representations of integers and by considering what happens when tens, hundreds, thousands, and so on are divided by the integer in question (Beckmann, 2018). With base-ten blocks, students will commonly say or draw, for example, "2 groups of 100, 4 groups of 10, and 3 groups of 1" to represent 243. This emphasis on base-ten representation, both physically and numerically, prepares prospective teachers to be aware of what their students learn prior to entering the secondary grades.

Class Activity: Problem 1 (10 minutes)

As you pass out **Problem 1** of the Class Activity, let undergraduates know they will continue to examine hypothetical student work to explore various tests of divisibility, and then discuss the following connection to teaching.

Discuss This Connection to Teaching

- In order to develop the skill of understanding how others use and reason with mathematics, undergraduates will consider what aspects of school student work make it mathematically valid.
- Exploring why divisibility rules work before writing formal proofs will help undergraduates understand the underlying mathematical ideas before engaging in the details of rigorous proof. It will also help prospective teachers explain mathematical ideas that are accessible to their future students.

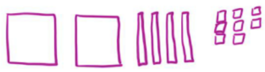
Instruct undergraduates to work in small groups on Problem 1 and then discuss the solutions. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion. From our experience, undergraduates may struggle with how to arrange 10 or 100 so that it is useful in recognizing groups of 3. Decide how long to let undergraduates wrestle with this idea as you circulate the class, helping them to realize that $10 = 9 + 1$, $100 = 99 + 1$, etc. It may be useful to have students focus on the tens digit in order to recognize groups of 3.

Class Activity Problem 1 : Part a

1. **Test for Divisibility by 3:**
   An integer is divisible by 3 if and only if the sum of its digits is divisible by 3.

   (a) Use a sketch to demonstrate how base-ten blocks could show whether or not the integer 248 is divisible by 3.

Sample Response:



248

each $\square = 99+1$, and $99 = 3 \cdot 33$

each $\square = 9+1$, and $9 = 3 \cdot 3$

To arrange the blocks in groups of 3, there will be 2 ones left from the hundreds, 4 ones left from the tens, and the 8 ones. $2+4+8 = 14$, which is not divisible by 3.

For Problem 1(b), some undergraduates may not understand how to numerically express an argument, and it may be helpful to start this solution together (e.g., writing $248 = 2(100) + 4(10) + 8(1)$). As you circulate the class, help undergraduates to use their reasoning from part (a).

Class Activity Problem 1 : Part b

(b) Write a sequence of equivalent equations that numerically express the argument you made above with base-ten blocks.

Sample Response:

$$\begin{aligned}
248 &= 2(100) + 4(10) + 8(1) \\
&= 2(99+1) + 4(9+1) + 8 \\
&= (2 \times 99 + 2) + (4 \times 9 + 4) + 8 \\
&= (2 \times 99 + 4 \times 9) + (2 + 4 + 8) \\
&= (2 \times 3 \times 33 + 4 \times 3 \times 3) + (2 + 4 + 8)
\end{aligned}$$

The first quantity, $(2 \times 3 \times 33 + 4 \times 3 \times 3)$, is a multiple of 3. The second quantity, $(2+4+8)$, is the sum of the digits in 248, but $(2+4+8) = 14$ is not a multiple of 3, so 248 is not divisible by 3.

After discussing the solution to Problem 1(b), tell undergraduates that they wrote 248 in expanded form, and emphasize the following connection to teaching.

Discuss This Connection to Teaching

The base-ten and place value systems are central in K–12 mathematics. Writing integers in expanded form is a part of the elementary school mathematics curriculum, so all prospective teachers should be familiar with writing integers in expanded form.

In generalizing this argument of divisibility by 3 in Problem 1(c), undergraduates may need a hint from you in order to recognize that they can use variables for digits and write a generic three-digit integer in base-ten representation as $abc = a(100) + b(10) + c(1)$.

Class Activity Problem 1 : Part c

(c) Outline an algebraic argument that shows that the test for divisibility by 3 holds for any three-digit integer. Ensure you consider both directions of the biconditional statement.

Solution:
Consider a three-digit integer n with digits a, b, and c. That is, a, b, and c are integers where $0 < a \leq 9$, $0 \leq b \leq 9$, and $0 \leq c \leq 9$. Thus,

$$\begin{aligned} n &= a(100) + b(10) + c(1) \\ &= a(99+1) + b(9+1) + c \\ &= (a \times 99 + b \times 9) + (a+b+c) \\ &= (a \times 3 \times 33 + b \times 3 \times 3) + (a+b+c) \\ &= 3t + (a+b+c) \end{aligned}$$

where $t = 33a + 3b$ is an integer because integers are closed under addition and multiplication.

If $a+b+c$ is divisible by 3, then the integer n is divisible by 3 (proving the direction, "If the sum of the digits is divisible by 3, then the integer is divisible by 3."). If the integer n is divisible by 3, then $3t + (a+b+c) = 3m$ for some integer m, and $a+b+c$ must also be divisible by 3 (proving the direction, "If the integer is divisible by 3, then the sum of the digits is divisible by 3.").

Commentary:
To generalize this argument for an integer with any number of digits, consider an integer with digits a_i in the 10^i place, and rewrite 10^i as $(10^i - 1) + 1$. You may want to give undergraduates this hint if you assign the homework problem where they prove that the test for divisibility by 3 holds for any integer.

Class Activity: Problems 2 & 3 (20 Minutes)

Distribute **Problems 2 and 3** of the Class Activity. Instruct undergraduates to work on the problems in small groups and discuss their solutions. These problems focus undergraduates' attention on responding to others' mathematical conjectures, which you can emphasize by discussing the following connection to teaching.

Discuss This Connection to Teaching

Problems 2 and 3 focus on analyzing other students' thinking in order to develop undergraduates' skills in understanding school student thinking and guiding school students' understanding. All undergraduates (especially prospective teachers) should explore how others use, reason with, and communicate mathematics. These problems also give prospective teachers (and tutors and future graduate students) an opportunity to think about how they would respond to student work in ways that nurture students' assets and understanding and in ways that help develop students' mathematical understanding.

Problem 2 distinguishes between how **one can recognize** whether an integer is divisible by 4 (Colby's method) and how to **prove** that an integer is divisible by 4 (Quinn's method).

Class Activity Problem 2

2. **Test for Divisibility by 4:**
   An integer is divisible by 4 if and only if the integer formed by the last two digits (that is, the integer between 0 and 99) is divisible by 4.

   Colby provided the following argument for explaining why the divisibility test for 4 always works.

   > *The* 100 *block is always divisible by* 4*, because* $4 \times 25 = 100$*. Any number of* 100 *blocks will be divisible by* 4*. So I just have to check whether the number formed by the tens and ones is divisible by* 4*.*

Quinn prefers to use divisibility by 2 to check for divisibility by 4.

> *First I check if the number is even, and if it is, then I know it is divisible by* 2. *So I divide the number by* 2, *and if that answer is even, then I know the number is divisible by* 4.

In Problems 2(a) and 2(b), undergraduates compare the relative advantages of Colby's method and Quinn's method for determining which integers are divisible by 4.

Class Activity Problem 2 : Parts a & b

(a) Explain when Colby's method might have an advantage over Quinn's when trying to recognize if a four-digit integer is divisible by 4.

Sample Responses:

- Colby's rule might be faster than Quinn's if you can recognize which integers between 0 and 99 are divisible by 4.
- Colby's rule works well for students who know all of the two-digit multiples of 4.
- Quinn's method can take longer because you have to divide twice, although dividing by 2 isn't too hard.
- Colby's rule explains why it is true in a physical sense, because it relies on building the integer with base-ten blocks and then forming groups of 4. It seems like young students would understand it this way.
- Quinn's method could be more difficult because you have to think about all of the digits in an integer. With Colby's method, you only have to look at the last two digits of an integer.

(b) Explain why Quinn might have suggested this method of checking for divisibility by 4 while studying prime factorization.

Sample Responses:

- When finding prime factors, Quinn would recognize that if there are two 2's in the prime factorization of an integer, then 4 is a factor of the integer.
- Factorization and divisibility are connected ideas, because if 4 is a factor of an integer n, then $n = 4m$ for some integer m, which is equivalent to saying n is divisible by 4. Quinn might not say it that way though, depending on how old they are.
- Quinn hasn't really found a test for divisibility by 4, but Quinn is applying the definition of divisibility accurately.

In Problem 2(c), undergraduates have the opportunity to critically analyze the relative advantages of tests for divisibility. This can help them see that though the test for divisibility by 4 has a straightforward proof that relies on the definition of divisibility and integer division with remainders, it isn't as useful as the test for divisibility by 5 because multiples of 4 aren't as easy to recognize as multiples of 5.

Class Activity Problem 2 : Part c

(c) Explain why students might have an easier time determining whether an integer is divisible by 5 than determining whether an integer is divisible by 4.

Sample Responses:

- It's not as easy to recognize which integers between 0 and 99 are divisible by 4, but it is really easy to recognize which integers between 0 and 9 are divisible by 5.

- Students only need to look at the last digit and see if it's a 0 or 5 to determine whether the integer is divisible by 5. The divisibility rule for 4 is harder because you have to look at the last two digits of an integer and then think about whether it's divisible by 4.

Commentary:
The proof for a test of divisibility by 4 proceeds similarly to the proof for divisibility by 5, that is, writing an integer n as $n = 100q + r$, where r is an integer between 0 and 99. Though the proof is no more difficult than the proof for divisibility by 5, the resulting test is more difficult to apply. It can be useful for undergraduates to recognize the distinction.

Problem 3 prompts undergraduates to examine another hypothetical student's work. Parts (a) and (b) of this problem give them an opportunity to construct proofs using the definitions of even and odd.

Class Activity Problem 3 : Parts a & b

3. **Student Reasoning About Multiples:**
A student, Malik, tells his teacher that he has noticed that "when you skip count by 3s, you get the pattern odd-even-odd-even-odd-even, but when you skip count by 2s or 4s, you only get evens." Malik asks why that happens.

(a) Explain why skip counting by 2s (as in, "$2, 4, 6, 8, 10, \ldots$") or 4s (as in, "$4, 8, 12, 16, 20, \ldots$") yields only even integers. Use the definition of even integers in your explanation.
Sample Response:
Skip counting by 2s results in integers of the form $2i$, where i is an integer. These integers are even by the definition of even integer. Skip counting by 4s results in integers of the form $4j$, where j is an integer. $4j = 2(2j)$, which is even by the definition of even integer (since $2j$ must also be an integer).

(b) Explain why skip counting by 3s (as in, "$3, 6, 9, 12, 15, 18, \ldots$") yields both odd and even integers.
Sample Response:
Skip counting by 3s results in integers of the form $3i$, where i is an integer. If i is odd, then there exists an integer n such that $i = 2n + 1$. Then

$$3i = 3(2n + 1) = 6n + 3 = 2(3n) + 2 + 1 = 2(3n + 1) + 1,$$

which is odd. If i is even, then there exists an integer n such that $i = 2n$ and

$$3i = 3(2n) = 2(3n),$$

which is even.

Problems 3(c) and 3(d) introduce undergraduates to guiding the mathematical thinking of others by evaluating mathematical questions that can be posed to students. The prompts posed here encourage the use of drawings as part of building mathematical reasoning.

Class Activity Problem 3 : Parts c & d

(c) Explain how the two sketches below, representing 12 and 15, might help Malik to understand why the pattern holds.

Sample Response:
By arranging the blocks with the groups of 6 made obvious, a student could see that a number that can be arranged by both groups of 3 and groups of 2 will be even. This is the case for the number 12. If the arrangements of blocks has a group of three "left over" (like 15) then the number is a multiple of 3 and is odd.

(d) Explain how the following question might help Malik to understand why the pattern holds.

If I skip count by 3*s and the last number I said was an even number, will the next multiple of* 3 *be even or odd? Draw a picture to show why.*

Sample Responses:

- By having a student draw a picture, the student will see that if they have an integer already grouped by 2s, adding another group of 3 will show that the integer is odd.
- By having the student focus on what happens when skipping from a multiple of three that is even to the next multiple of three, it simplifies the reasoning by focusing on one of the two possible cases.

After discussing the solutions to Problems 3(c) and 3(d), tell undergraduates that Malik's reasoning is an example of how looking for structure one way (i.e., the numerical structure that results when dividing by 2) can lead to finding structure in other ways—in Malik's case, he doesn't have a name for it, but he's recognized that multiples of 3 can be classified by whether they are even or odd. This leads directly to a notion of divisibility by 6, and undergraduates will further consider divisibility by 6 in the Homework Problems.

Emphasize the following connection to teaching.

Discuss This Connection to Teaching

School students often make mathematical observations. It is valuable for teachers to emphasize that mathematical reasoning and proof are what is used to establish the truth of a conjecture, and it is in the student's power to establish that a statement is either true or false. By using questioning to guide Malik's understanding, teachers are developing his mathematical reasoning abilities. Teachers help students by recognizing what the students understand and identifying the important mathematical ideas the students are trying to communicate. Students learn that the mathematics, rather than the teacher, is the authority of mathematical truth.

Wrap-Up (up to 10 minutes)

Conclude the lesson by briefly discussing the similarities and differences in the proofs of the divisibility tests for $5, 3$, and 4. This discussion may include the following ideas:

- Divisibility tests for 5 and 4 are similar, because 5 is a factor of any number of 10s and 4 is a factor of any number of 100s. Thus, these tests can proceed by inspecting the digits in certain place values.

- Three is not a factor of any power of 10, so the structure of the divisibility test for 3 is different than for 5 and 4. Writing the integer in expanded form is useful, because each power of 10 can be written as $(10^k - 1) + 1$, and $10^k - 1$ is divisible by 3.

Summarize a few connections to teaching that were present throughout the lesson, focusing on how undergraduates engaged in problems that focused on analyzing student thinking and guiding student understanding.

Discuss This Connection to Teaching

- Undergraduates made use of the structure of integers in the base-ten system, relying on place value representation and closure of integers under various addition and multiplication.
- Looking for and making use of structure is a mathematical practice exercised by students and by mathematicians.
- The problems in this lesson demonstrate how formal proofs underlie the mathematical reasoning of younger students.

We have found it useful to ask undergraduates to complete an exit ticket at the end of the lesson. (See Chapter 1 for guidance on using exit tickets.)

Homework Problems

At the end of the lesson, assign the following homework problems, and include any additional homework problems at your discretion. The first two problems highlight connections to teaching and the third problem gives undergraduates an opportunity to write formal proofs of the divisibility rules they encountered throughout the lesson.

Problem 1 builds from the "Malik" problem from the Class Activity and prompts undergraduates to consider the validity of two hypothetical students' tests for divisibility by 6. High school students may find or develop divisibility rules and prospective teachers will need to be able to consider ideas that students have and assess whether their ideas are sometimes or always true. This problem relies on the mathematical practice of constructing viable arguments and critiquing the reasoning of others.

Homework Problem 1

1. Adam and Charlotte learn that a number is divisible by 6 if it is divisible by both 2 and 3. Each attempts to apply similar reasoning to state a divisibility rule for 20. Adam says that "because $20 = 2 \times 10$, if a number is divisible by both 2 and 10, then the number is divisible by 20." Charlotte states that "because 20 is divisible by 4 and 5, if a number is divisible by both 4 and 5, then the number is divisible by 20."

   (a) Why doesn't Adam's rule work, but Charlotte's rule does? What is the key difference between their two rules?

   Sample Responses:

   - Adam's rule does not work because if a number is divisible by 10, then you know it is divisible by 2; checking to see if it is divisible by 2 after knowing it is divisible by 10 does not give us any new information.
   - Charlotte's rule works because 4 and 5 are relatively prime.
   - 4 and 5 share no common factors but 2 and 10 share a factor of 2. As long as the factors are relatively prime, this method will work.
   - Adam might be confusing the statement with its converse. It is true that if a number is divisible by 20, then it is divisible by 2 and 10, but that's the converse of the rule he stated.

(b) Adam's proof to their conjecture is shown below. Identify the error in their proof and explain why it is an error.

> Let n be any integer that is divisible by 2 and 10. By the definition of divisibility, since n is divisible by 2, there is an integer k where $n = 2k$. Since n is also divisible by 10, that means that k must be divisible by 10. By the definition of divisibility, there is an integer l where $k = 10l$. Using substitution, $n = 2k = 2(10l) = 20l$. Since $n = 20l$ for some integer l, n is divisible by 20.

Solution:
The error occurs in the following sentence: "Since n is also divisible by 10, that means that k must be divisible by 10." Because n is divisible by 10, there must be an integer m such that $n = 10m$. Adam could conclude that $2k = 10m$ so k is divisible by 5, but not by 10.

(c) Their teacher recognizes that Adam was probably testing the number 20 or 40 when they conjectured their rule. What are some other integers the teacher could encourage Adam and Charlotte to experiment with that would help them to understand divisibility by 20?

Sample Response:
Any number that is a multiple of 10 but not a multiple of 4 can show what Adam's mistake is. Numbers like 30 and 50 fit the hypothesis of Adam's rule but not the conclusion and are good numbers for them to investigate.

Problem 2 highlights another hypothetical student's conjecture about divisibility tests, this time about a divisibility test for 7.

Homework Problem 2

2. A student, Isla, tells you that she has created a "test" for divisibility by 7. She claims that an integer n is divisible by 7 if and only if the rightmost two digits of n form an integer that is a multiple of 7. Provide a counterexample showing that Isla's test for divisibility by 7 doesn't work, and explain why Isla may believe her test works.

Sample Responses:

- 114 has rightmost two digits that are divisible by 7, but 114 is not divisible by 7.
- 105 is a multiple of 7 (i.e., $7 \times 15 = 105$) but its rightmost two digits form an integer (5) that is not divisible by 7.
- Isla might believe this rule is true because she is copying the format of the test for divisibility by 4. She still needs to learn why the rule is true for testing divisibility by 4.

Problem 3 prompts undergraduates to write formal proofs of the divisibility rules for 5, 3, 4, and 6. Proofs for these statements are generalizations of the reasoning presented in class. Writing formal proofs of these statements can give undergraduates practice with proving biconditional statements, and they may need prompting to write these tests as biconditional statements.You might choose to assign only some of these four, depending on your course goals. The style and rigor of these proofs should follow your expectations for other proofs in your course.

Homework Problem 3

3. Proving Tests for Divisibility

(a) Prove that an integer is divisible by 5 if and only if its last digit is 0 or 5.

Commentary:
This proof was outlined during the Class Activity. We want undergraduates to recognize that any integer n can be written $n = 10q + r$, where q is an integer and r is an integer such that $0 \le r \le 9$, and that the base-ten representation of integers is such that r is the ones digit of the integer n.

(b) Prove that an integer is divisible by 3 if and only if the sum of its digits is divisible by 3.

Commentary:
The proof of this statement for a 3-digit integer was outlined during the Class Activity. To generalize an argument for an integer with any number of digits, consider an integer with digits a_i in the 10^i place.

(c) Prove that an integer is divisible by 4 if and only if the integer formed by its last two digits is divisible by 4.

Commentary:
This proof was outlined during the Class Activity. We want undergraduates to recognize that any integer n can be written $n = 100q + r$, where q is an integer and r is an integer such that $0 \le r \le 99$, and that the base-ten representation of integers is such that r is the integer formed by the tens and ones digits of the integer n. The structure of this argument otherwise mimics that of part (a).

(d) Prove that an integer is divisible by 6 if and only if the integer is divisible by both 2 and 3.

Commentary:
This problem provides undergraduates an opportunity to prove that even multiples of three, which they examined using Malik's skip counting observation, are divisible by 6. The proof of the statement "If an integer is divisible by 6, then it is divisible by both 2 and 3." is straightforward. To prove the other direction, there are a few approaches that naturally arise after undergraduates have engaged in the reasoning from this lesson. One approach recognizes that using 6 as the divisor in the quotient-remainder theorem is useful: From the quotient remainder theorem, $n = 6q + r$ for some integer q and some integer r such that $0 \le r \le 5$. Also, if an integer is divisible by both 2 and 3, then it satisfies $n = 2k$ and $n = 3l$ for some integers k and l. Since $n = 2k$, it follows that r is a multiple of 2. Since $n = 3l$, it follows that r is a multiple of 3. The only value of r in $\{0, 1, 2, 3, 4, 5\}$ that is both a multiple of 2 and a multiple of 3 is 0, so $n = 6q$ and n is divisible by 6. A second approach relies on the definition of divisibility and the closure of the integers: $n = 2k$ and $n = 3l$ for some integers k and l. It follows that $3n = 6k$ and $2n = 6l$, so subtracting these equations from the left and right sides, respectively, yields $n = 6(k - l)$, where $k - l$ is an integer and thus, n is divisible by 6.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Problem 1 assesses undergraduates' proof-writing techniques and how they relied on properties of integers in their proof.

Assessment Problem 1

1. An integer n is divisible by 10 if and only if the final digit of the integer is 0.

   (a) Prove that this is true.

   Sample Response:

   Since this is a biconditional statement, the proof involves two statements.

   Claim 1. *If n is divisible by 10, then the final digit of n is 0.*

   Proof. Let n be an integer that is divisible by 10. This means that $n = 10k$ for some integer k. By the quotient-remainder theorem, we also know that $n = 10q + r$ for some integer q and some integer r, where $0 \leq r \leq 9$. By the base-ten representation of integers, r is the final digit of n. Thus,

$$\begin{aligned} 10k &= 10q + r \\ 10k - 10q &= r \\ 10(k - q) &= r \end{aligned}$$

   Because $k - q$ is an integer (since integers are closed under addition), r is a multiple of 10. The only multiple of 10 such that $0 \leq r \leq 9$ is 0, so r must be 0. □

   Claim 2. *If the final digit of n is 0, then n is divisible by 10.*

   Proof. Let n be an integer whose final digit is 0. By the quotient-remainder theorem, $n = 10q + r$ for some integer q and some integer r, where $0 \leq r \leq 9$. By the base-ten representation of integers, r is the final digit of n, so $r = 0$. Thus $n = 10q$, and n is divisible by 10. □

   (b) Describe how you relied on properties of integers in your proof.

   Sample Responses:

   - Combining integers via multiplication, subtraction, and addition always leads to other integers.
   - In the base-ten system of integers, the remainder from dividing n by 10 is the last digit of the numeral for n.

Problem 2 assesses undergraduates' ability to analyze school student thinking and to guide school student understanding. It also focuses their attention on the biconditional statement that underlies each test for divisibility.

Assessment Problem 2

2. In class, Olivia learned that a number is divisible by 6 if it is divisible by both 2 and 3 and used that to conjecture a divisibility rule for 60. Olivia says that because 60 is divisible by 6 and 10, you can tell which numbers are divisible by 60 by checking if the number is divisible by both 6 and 10.

   (a) Rewrite Olivia's conjecture as a biconditional statement.

   Sample Responses:

   - An integer is divisible by 60 if and only if it is divisible by both 6 and 10.
   - "If an integer is divisible by 60, then it is divisible by both 6 and 10." and "If an integer is divisible by both 6 and 10, then it is divisible by 60.".

(b) One direction of the biconditional statement is true and the other is false. State which direction is false and find a counterexample showing that it is false.

Solution:

This statement is false: If an integer is divisible by both 6 and 10, then it is divisible by 60. Many counterexamples exist. 30 is the smallest counterexample.

(c) Explain how the following question might help Olivia to understand that her rule doesn't always work.

What is the least common multiple of 6 *and* 10*?*

Sample Response:

It would help Olivia see that 30 is a number that is divisible by both 6 and 10 but isn't divisible by 60.

(d) Provide two reasons why you think Olivia made this conjecture.

Sample Responses:

- This sort of statement is true with other numbers. For example, it is true that if a number is divisible by both 2 and 5, then the number is divisible by 10. The distinction is that 2 and 5 don't share any common prime factors, but 6 and 10 share 2 as a common prime factor
- Olivia is only thinking about one direction of this biconditional statement. It is true that if a number is divisible by 60, then the number is divisible by 6 and 10.

7.6 References

[1] Beckmann, S. (2018). *Mathematics for elementary teachers with activities*. Pearson.

[2] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from http://www.corestandards.org/

7.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

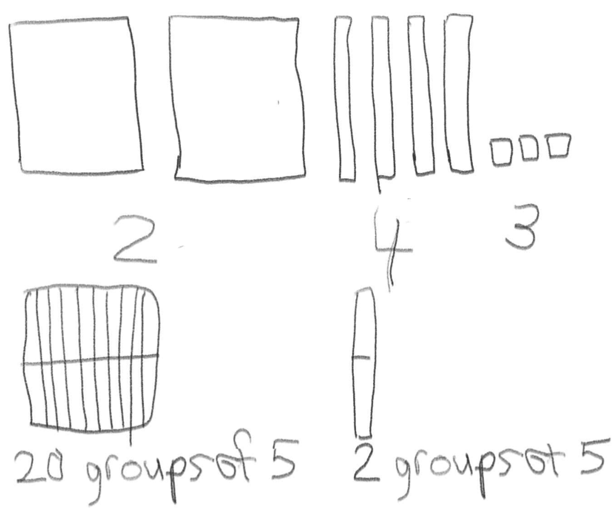
NAME: **PRE-ACTIVITY:** FOUNDATIONS OF DIVISIBILITY (page 1 of 1)

A fifth-grade class is discussing divisibility by 5. One student says, "My sister said you can tell if a number is divisible by 5 if it ends in a 0 or a 5." Another student adds, "Yeah, you can tell from skip counting by 5s: $5, 10, 15, 20, 25, \ldots$"

The teacher recognizes an opportunity to engage students in reasoning about properties of positive integers. The students have been using base-ten blocks to demonstrate whether various numbers are even or odd, so the teacher asks students to use their blocks to demonstrate whether the three numbers 243, 240, and 245 are divisible by 5.

Notice that there is a difference between **how to recognize** if an integer is divisible by 5 or not, which involves inspecting the digit in the ones place, and **proving** that an integer is divisible by 5, which involves determining whether there exists an integer m such that the integer can be written as $5m$.

Below is what one student, Rickie, wrote to explain that 243 is not divisible by 5, but 240 and 245 are.



I represented 243 as 2 groups of 100, 4 groups of 10, and then 3 ones. Each group of 10 forms 2 groups of 5, and each group of 100 forms 20 groups of 5. Then you have to look at the leftover blocks.
If the number ends in 0, then you haven't added any blocks and you can keep the groups you have. That's what happens with 240.
If the number ends in 1, 2, 3, or 4, then you have leftover blocks. That's what happens with 243.
If the number ends in 5, then you have 5 leftover blocks and you can put those into 1 group of 5. That's what happens with 245.
I guess you could also have 6, 7, 8 or 9 leftover blocks, but that's it.

State Rickie's test for divisibility by 5 as a biconditional statement. Explain how to generalize Rickie's argument about why the divisibility rule for 5 will always work.

NAME: **CLASS ACTIVITY:** FOUNDATIONS OF DIVISIBILITY (page 1 of 3)

1. **Test for Divisibility by 3:**
   An integer is divisible by 3 if and only if the sum of its digits is divisible by 3.

   (a) Use a sketch to demonstrate how base-ten blocks could show whether or not the integer 248 is divisible by 3.

   (b) Write a sequence of equivalent equations that numerically express the argument you made above with base-ten blocks.

   (c) Outline an algebraic argument that shows that the test for divisibility by 3 holds for any three-digit integer. Ensure you consider both directions of the biconditional statement.

CLASS ACTIVITY: FOUNDATIONS OF DIVISIBILITY (page 2 of 3)

2. **Test for Divisibility by 4:**
An integer is divisible by 4 if and only if the integer formed by the last two digits (that is, the integer between 0 and 99) is divisible by 4.

Colby provided the following argument for explaining why the divisibility test for 4 always works.

> *The* 100 *block is always divisible by* 4*, because* $4 \times 25 = 100$*. Any number of* 100 *blocks will be divisible by* 4*. So I just have to check whether the number formed by the tens and ones is divisible by* 4*.*

Quinn prefers to use divisibility by 2 to check for divisibility by 4.

> *First I check if the number is even, and if it is, then I know it is divisible by* 2*. So I divide the number by* 2*, and if that answer is even, then I know the number is divisible by* 4*.*

(a) Explain when Colby's method might have an advantage over Quinn's when trying to recognize if a four-digit integer is divisible by 4.

(b) Explain why Quinn might have suggested this method of checking for divisibility by 4 while studying prime factorization.

(c) Explain why students might have an easier time determining whether an integer is divisible by 5 than determining whether an integer is divisible by 4.

3. **Student Reasoning About Multiples:**
A student, Malik, tells his teacher that he has noticed that "when you skip count by 3s, you get the pattern odd-even-odd-even-odd-even, but when you skip count by 2s or 4s, you only get evens." Malik asks why that happens.

   (a) Explain why skip counting by 2s (as in, "$2, 4, 6, 8, 10, \ldots$") or 4s (as in, "$4, 8, 12, 16, 20, \ldots$") yields only even integers. Use the definition of even integers in your explanation.

   (b) Explain why skip counting by 3s (as in, "$3, 6, 9, 12, 15, 18, \ldots$") yields both odd and even integers.

   (c) Explain how the two sketches below, representing 12 and 15, might help Malik to understand why the pattern holds.

   (d) Explain how the following question might help Malik to understand why the pattern holds.
If I skip count by 3s and the last number I said was an even number, will the next multiple of 3 be even or odd? Draw a picture to show why.

NAME: **HOMEWORK PROBLEMS:** FOUNDATIONS OF DIVISIBILITY (page 1 of 1)

1. Adam and Charlotte learn that a number is divisible by 6 if it is divisible by both 2 and 3. Each attempts to apply similar reasoning to state a divisibility rule for 20. Adam says that "because $20 = 2 \times 10$, if a number is divisible by both 2 and 10, then the number is divisible by 20." Charlotte states that "because 20 is divisible by 4 and 5, if a number is divisible by both 4 and 5, then the number is divisible by 20."

   (a) Why doesn't Adam's rule work, but Charlotte's rule does? What is the key difference between their two rules?

   (b) Adam's proof to their conjecture is shown below. Identify the error in their proof and explain why it is an error.

   > Let n be any integer that is divisible by 2 and 10. By the definition of divisibility, since n is divisible by 2, there is an integer k where $n = 2k$. Since n is also divisible by 10, that means that k must be divisible by 10. By the definition of divisibility, there is an integer l where $k = 10l$. Using substitution, $n = 2k = 2(10l) = 20l$. Since $n = 20l$ for some integer l, n is divisible by 20.

   (c) Their teacher recognizes that Adam was probably testing the number 20 or 40 when they conjectured their rule. What are some other integers the teacher could encourage Adam and Charlotte to experiment with that would help them to understand divisibility by 20?

2. A student, Isla, tells you that she has created a "test" for divisibility by 7. She claims that an integer n is divisible by 7 if and only if the rightmost two digits of n form an integer that is a multiple of 7. Provide a counterexample showing that Isla's test for divisibility by 7 doesn't work, and explain why Isla may believe her test works.

3. Proving Tests for Divisibility

   (a) Prove that an integer is divisible by 5 if and only if its last digit is 0 or 5.

   (b) Prove that an integer is divisible by 3 if and only if the sum of its digits is divisible by 3.

   (c) Prove that an integer is divisible by 4 if and only if the integer formed by its last two digits is divisible by 4.

   (d) Prove that an integer is divisible by 6 if and only if the integer is divisible by both 2 and 3.

NAME: **ASSESSMENT PROBLEMS:** FOUNDATIONS OF DIVISIBILITY (page 1 of 2)

1. An integer n is divisible by 10 if and only if the final digit of the integer is 0.

   (a) Prove that this is true.

   (b) Describe how you relied on properties of integers in your proof.

ASSESSMENT PROBLEMS: FOUNDATIONS OF DIVISIBILITY (page 2 of 2)

2. In class, Olivia learned that a number is divisible by 6 if it is divisible by both 2 and 3 and used that to conjecture a divisibility rule for 60. Olivia says that because 60 is divisible by 6 and 10, you can tell which numbers are divisible by 60 by checking if the number is divisible by both 6 and 10.

   (a) Rewrite Olivia's conjecture as a biconditional statement.

   (b) One direction of the biconditional statement is true and the other is false. State which direction is false and find a counterexample showing that it is false.

   (c) Explain how the following question might help Olivia to understand that her rule doesn't always work.

   What is the least common multiple of 6 *and* 10*?*

   (d) Provide two reasons why you think Olivia made this conjecture.

8

Solving Equations in Alternative Number Systems

Abstract (Modern) Algebra I

Andrew Kercher, *Simon Fraser University*
James A. M. Álvarez, *The University of Texas at Arlington*
Kyle Turner, *The University of Texas at Arlington*

8.1 Overview and Outline of Lesson

Many undergraduate courses in abstract algebra include a learning goal of developing the capacity for students to reason abstractly about mathematical structures. As such, this reasoning is often applied in the context of highlighting the mathematical structures that make familiar high school algebra techniques possible. However, attempting to apply these familiar techniques to solving polynomial equations over finite integer rings leads to outcomes that undergraduates might not expect. For example, a linear equation with finitely many solutions might still have more than one solution; similarly, quadratic equations may have more than two real roots. Ultimately, these observations can be leveraged to motivate undergraduates to think about the ring structure for which their familiar techniques work and eventually to determine for which n the ring $\mathbb{Z}_n$ is an integral domain. This lesson focuses on building intuition about the ring $\mathbb{Z}_n$ while also embedding targeted opportunities for reflecting on the reasoning behind various methods used to solve polynomial equations from high school mathematics.

1. Launch—Pre-Activity
   Prior to the lesson, undergraduates complete the Pre-Activity. First, undergraduates identify which elements of $\mathbb{Z}_{10}$ have the same algebraic behavior as certain rational numbers. Then, they graph the solution set of a linear equation in two variables in both $\mathbb{Z}_5 \times \mathbb{Z}_5$ and $\mathbb{Z}_6 \times \mathbb{Z}_6$. Finally, undergraduates describe the precise algebraic steps used to solve a linear equation of one variable in $\mathbb{R}$ and attempt to transfer these steps to $\mathbb{Z}_5$ and $\mathbb{Z}_6$, where they encounter difficulty if that ring is not also a field. Instructors can launch this lesson by reviewing the answers to the Pre-Activity.

2. Explore—Class Activity

   • *Problems 1–3*:
   Using their graphical representations from Problem 2 of the Pre-Activity, undergraduates are informally introduced to the idea of a zero divisor in $\mathbb{Z}_6$; this lays the groundwork for undergraduates' capacity to later deduce that no zero divisor can be a unit. In Problem 3, a hypothetical student makes a conjecture which undergraduates then evaluate. In doing so, they consider which elements of $\mathbb{Z}_n$ are zero divisors and which

are units. By comparing these lists, undergraduates deduce that any element of $\mathbb{Z}_n$ represented by an integer that is relatively prime with n is a unit but is not a zero divisor.

- *Discussion—The Structure of* $\mathbb{Z}_n$:
  The instructor leverages undergraduates' work on Problems 1–3 to motivate the definitions of zero divisors and integral domains. Then, they prove (or discuss) a number of results that culminate in the conclusion that $\mathbb{Z}_p$ is a field and integral domain when p is prime. Central to this claim is the fact that any element of $\mathbb{Z}_n$ which is relatively prime with n is a unit and not a zero divisor.
- *Problems 4–6*:
  Undergraduates examine two techniques for solving quadratic equations: factoring and the application of the standard quadratic formula. They find that both techniques can be problematic in $\mathbb{Z}_{10}$.

3. Closure—Wrap-Up
   Instructor wraps up the lesson by reiterating that $\mathbb{Z}_n$ is both a field and an integral domain when n is prime. In this case, many familiar algebraic techniques from secondary school can still be used to solve equations. When n is composite, the application of these techniques proves to be problematic.

8.2 Alignment with College Curriculum

A study of $\mathbb{Z}_n$ appears in undergraduate abstract algebra courses and in introductory number theory courses. It is the classic example of a finite commutative ring, and attempting to solve polynomial equations in $\mathbb{Z}_n$ leads naturally to definitions of zero divisors and units. This lesson builds a foundation from which lessons on rings of polynomials, factorization, and algebraic closure could be built.

8.3 Links to School Mathematics

Solving equations is a significant component of high school algebra. The majority of the experiences in equation-solving occur over $\mathbb{R}$, but the field properties of $\mathbb{R}$ that underlie the process of equation-solving are typically not emphasized. This lesson addresses the Common Core State Standards for Mathematics (CCSSM, 2010) related to solving linear equations and quadratic equations and attends to the properties of $\mathbb{R}$ "taken for granted" by exploring the same material in alternative settings. In particular, these properties include the zero product property and the existence of a multiplicative inverse for every nonzero element.

This lesson highlights:

- Solving equations in $\mathbb{Z}_n$ can be both similar to and different from solving equations in $\mathbb{R}$;
- Visualizing solutions to equations in $\mathbb{Z}_n$ can be both similar to and different from visualizing solutions in $\mathbb{R}$.

This lesson addresses several mathematical knowledge and practice expectations included in high school standards documents, such as the CCSSM. For example, high school students are regularly taught how to solve both linear and quadratic equations of one variable (c.f. CCSS.MATH.CONTENT.HSA.REI.B.3, CCSS.MATH.CONTENT.HSA.REI.B.4). In the case of a quadratic equation, there are multiple associated algebraic techniques of which high school students are expected to develop a comprehensive understanding (see CCSS.MATH.CONTENT.HSA.REI.B.4.B). One such technique is factoring, which exemplifies the types of algebraic manipulations that high school students are expected to master when working with equations—namely, those manipulations that do not change the expected solution set of the equation (see CCSS.MATH.CONTENT.HSA.SSE.B.3). Finally, this lesson and its activities aim to provide prospective teachers with opportunities to construct mathematical arguments, analyze and respond to the arguments of others, and to critique the underlying reasoning of such an argument.

8.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- The definition of a ring (i.e., the ring axioms);

- The definitions of unit and field;
- The definition of $\mathbb{Z}_n$ and how addition and multiplication are conducted in this ring.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Justify when an element of $\mathbb{Z}_n$ is a zero divisor or a unit;
- Contrast the process of solving quadratic equations in $\mathbb{Z}_n$ with the process of solving in $\mathbb{R}$ by:
  - Explaining how factoring outside of an integral domain might fail to find the entire set of roots;
  - Explaining how the quadratic formula is more difficult to apply in general, but especially outside of a field.
- Analyze hypothetical student work or conjectures to explore student thinking about solving equations in $\mathbb{Z}_n$;
- Pose guiding questions to help a hypothetical student determine when a ring is an integral domain.

Anticipated Length

One 75-minute class session.

Materials

The following materials are required for this lesson.

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

8.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to be completed in preparation for this lesson.

> *We recommend that you collect this Pre-Activity the day before the lesson so that you can review undergraduates' responses before you begin the Class Activity. This will help you determine if you need to spend additional time reviewing the solutions to the Pre-Activity with your undergraduates.*

Pre-Activity Review (10 minutes)

Discuss the solutions to the Pre-Activity as needed; if you see that most undergraduates completed each problem correctly, you may not need to spend much time reviewing or discussing the solutions.

> *The Pre-Activity is designed to re-familiarize undergraduates with modular arithmetic in $\mathbb{Z}_n$ by asking them to make basic computations, visualize lines, and solve equations of one variable in several different finite integer rings. These are all skills that will be used directly in the Class Activity.*

Probe undergraduates' understanding of Problem 1(b) as it can be referenced later to facilitate Problem 3(c) of the Class Activity, if needed. Use questions similar to the following to generate additional discussion:

- What element of $\mathbb{Z}_{10}$ do you think might represent "$3/7$"? Is this the only such element? Explain.
- What element(s) of $\mathbb{Z}_{10}$ do you think might represent "$\sqrt{5}$"? Is this the only such element? Explain.
- Based on your answers to Problem 1(b), which are the units of $\mathbb{Z}_{10}$?

Pre-Activity Problem 1

1. Recall that $\mathbb{Z}_n$ is the set of *equivalence classes* on the integers, where two integers are in the same equivalence class if and only if they both have the same (smallest, non-negative) remainder when divided by n. The set $\mathbb{Z}_n$ contains n such equivalence classes which, canonically, are represented by the possible remainders when an integer is divided by n: $\{0, 1, \ldots, n-2, n-1\}$.

   (a) Fill in the following chart with the representative of each integer's equivalence class in $\mathbb{Z}_{10}$.

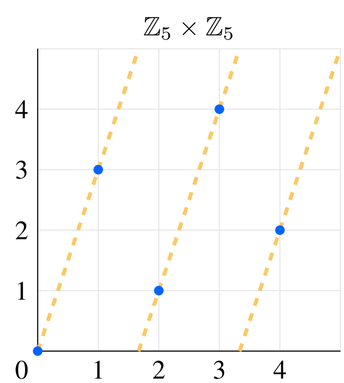
Integer	36	17	-4	-17
Representative in $\mathbb{Z}_{10}$	6	7	6	3

   If we are careful, we can also (sometimes) represent non-integers as elements of $\mathbb{Z}_n$. For example, if we interpret the notation "$1/3$" as "the element you multiply by 3 to get 1," we would then consider 7 in $\mathbb{Z}_{10}$ a representative of "$1/3$", since $3 \cdot 7 = 21 = 1$ (where $21 = 1$ because 21 has remainder 1 when divided by 10). Furthermore, no other element of $\mathbb{Z}_{10}$ has this property.

   (b) Fill in the following chart with the representative in $\mathbb{Z}_{10}$, if it exists.

"Non-integer"	"$1/1$"	"$1/2$"	"$1/3$"	"$1/4$"	"$1/5$"	"$1/6$"	"$1/7$"	"$1/8$"	"$1/9$"
Representative in $\mathbb{Z}_{10}$	1	X	7	X	X	X	3	X	9

In Problem 2, we found it useful to ask undergraduates to visualize the graph as a continuous geometric line with slope 3 which "wraps around" the finite space. To see this, we "extend" the grid to include the space "between" $n-1$ and n as shown here.



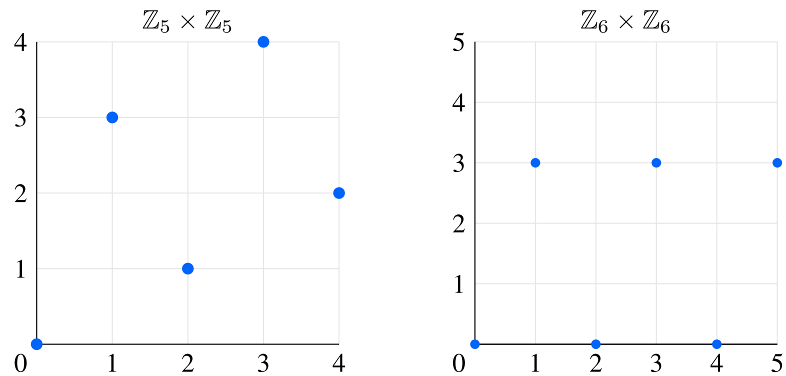
Undergraduates will be directed to refer back to and use their graphs when they answer Problems 1 and 2 of the Class Activity, so make sure that it is clear what is and is not part of the graph if discussing this representation.

Pre-Activity Problem 2

2. Let A be a set of elements (numbers) with a well-defined notion of addition and multiplication. We define a *line over A* as the solution set to an equation of the form $ax + by = c$ for some fixed values of $a, b, c \in$ A. That is, a line is the set of all ordered pairs $(x, y) \in$ A $\times$ A that make $ax + by = c$ a true statement in A. The *graph of a line* is a scatter plot of the solution set on a coordinate plane, usually one with perpendicular axes marked by the elements of A.

   For example, in the system of real numbers, the set of all ordered pairs that make $y = 3x$ a true statement in $\mathbb{R}$ has a graph which is a continuous straight geometric line of slope 3 through the point $(0,0)$ in our usual Cartesian coordinate system.

Graph the line $y = 3x$ in the following spaces on the provided axes.



Facilitate a class discussion focused on why it would be useful to create a graph of an equation. Prompt undergraduates to consider both the pros and cons of visualizing abstract constructs—for example, you might touch on the fact that using a dotted line to produce the graph is useful but could also be misleading if it is incorrectly taken to be part of the graph itself. Discuss the following connection to teaching:

Discuss This Connection to Teaching

High school teachers are expected to give their students opportunities to work with both symbolic and visual representations of mathematical concepts. This context focuses prospective teachers' attention on the meaning of the linear equation that may be overlooked when considering equations over $\mathbb{R}$. By asking them to transition between symbolic and visual representations of a line in unconventional spaces, we model pedagogical practices that they will need to replicate in their future classrooms and look back to high school topics from a different point of view.

The questions in Problem 3 are intended to elicit undergraduate thinking about assertions of equality and whether usual algebraic manipulations are applicable and reliably produce the entire solution set. Encourage undergraduates to think in terms of ring structure by using questions such as the following:

- The ring axioms are defined using addition and multiplication; what do we mean if we talk about "subtraction" and "division" in a ring? Do these concepts always exist?
- How might we justify that "adding the same quantity to each side" preserves equality in rings other than $\mathbb{R}$?

Pre-Activity Problem 3

3. In solving the equation $x + 4 = 1 + 4x$ in $\mathbb{R}$, a student makes the following algebraic manipulations:

$$\begin{aligned} x + 4 &= 1 + 4x \\ 4 &= 1 + 3x \\ 3 &= 3x \\ 1 &= x \end{aligned}$$

The student then concludes that $x = 1$ is the solution set to $x + 4 = 1 + 4x$ in $\mathbb{R}$.

(a) Describe the mathematical justification for each step in the student's solution.

Solution:

The student is using additive and multiplicative inverses to simplify the equation until it's in a form where the solution is evident. First, the student adds $-x$ to each side; then, they add -1 to each side; finally, they multiply everything by $\frac{1}{3}$. The last line shows that the entire solution set is $\{1\}$.

(b) To solve $x + 4 = 1 + 4x$ for x in $\mathbb{Z}_5$, are we allowed to repeat the process the student used (in $\mathbb{R}$) as presented above? Does this process yield the entire solution set to the equation? Explain.

Solution:

We must make some adjustments, but the basic idea of each step is still sound in $\mathbb{Z}_5$:

- Because $-x = 4x$ in $\mathbb{Z}_5$, we first add $4x$ to each side: $x+4 = 1+4x \Rightarrow 5x+4 = 1+8x \Rightarrow 4 = 1 + 3x$.
- Because $-1 = 4$ in $\mathbb{Z}_5$, we next add 4 to each side: $4 = 1 + 3x \Rightarrow 8 = 5 + 3x \Rightarrow 3 = 3x$.
- Finally, because $3 \cdot 7 = 21 = 1$, $1/3 = 7$ in $\mathbb{Z}_5$. So we multiply both sides by 7, and: $3 = 3x \Rightarrow 21 = 21x \Rightarrow 1 = x$.

Just like in the real numbers, we see from here that the solution set is $\{1\}$.

(c) To solve $x + 4 = 1 + 4x$ for x in $\mathbb{Z}_6$, are we allowed to repeat the process the student used (in $\mathbb{R}$) as presented above? Does this process yield the entire solution set to the equation? Explain.

Solution:

If we attempt to make the same adjustments as in $\mathbb{Z}_5$:

- Now $-x = 5x$ in $\mathbb{Z}_6$, so we add $5x$ to each side: $x + 4 = 1 + 4x \Rightarrow 6x + 4 = 1 + 9x \Rightarrow 4 = 1 + 3x$.
- Similarly, because $-1 = 5$ in $\mathbb{Z}_6$, we next add 5 to each side: $4 = 1 + 3x \Rightarrow 9 = 6 + 3x \Rightarrow 3 = 3x$.
- At this point, however, we can check by exhaustion that none of the elements of $\mathbb{Z}_6$ function like $1/3$: that is, $\forall\, x \in \mathbb{Z}_6, 3x \neq 1$.

Because we cannot simplify any further, we must check by inspection to find that the solution set is $\{1, 3, 5\}$.

Commentary:

Make sure undergraduates take away the following points from this problem.

- "Division" may not be well-defined in $\mathbb{Z}_n$ if not all of its elements are units (i.e., if it is not a field).
- In reasonably small finite rings such as $\mathbb{Z}_5$ or $\mathbb{Z}_6$, it is easy to check for solutions to equations by exhaustion.

Wrap up the Pre-Activity by discussing and summarizing ways that finite integer rings do not have the familiar structure of $\mathbb{R}$ and, as a result, certain ideas we have about solving linear equations are not preserved. Even when written as simply as $ax = b$, the solution to a linear equation may not be as straightforward as we might have hoped.

Class Activity: Problems 1–3 (25 minutes)

To introduce the lesson, discuss the following idea with your class:

Solving an equation of one variable entails listing or representing the numbers which, when written in place of the variable, yield a true mathematical sentence. This means we may think of an equation as a way of stating a property, and the solution set of that equation as the collection of numbers that have that property. Typically, we identify the solutions by translating the equation into equivalent but simpler forms until the set of numbers that make the equation true is evident. In making these translations, we appeal to the properties of the number system within which the equation is being solved. As we saw in the Pre-Activity, we need to understand what properties the finite integer rings possess so that we can reliably use equations in this way.

Distribute the Class Activity. Instruct undergraduates to work on the Problems 1 and 2 in small groups. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion. It may be helpful to illustrate (or offer a hint about) how Problems 1 and 2 can be solved visually by drawing the graph of the line $y = 3x$ in $\mathbb{R} \times \mathbb{R}$. Ask undergraduates to think about how the change in domain affects the usual way we find all of the solutions to $0 = 3x$ in $\mathbb{R} \times \mathbb{R}$ by locating where the graph of the line crosses the x-axis (that is, where $y = 0$). Encourage them to use the graphs of the lines in $\mathbb{Z}_5 \times \mathbb{Z}_5$ and $\mathbb{Z}_6 \times \mathbb{Z}_6$ from the Pre-Activity for these problems.

Some of the following questions can be used to motivate discussion:

- How could we have solved these problems if we didn't already have the relevant graphs on hand?
- Since no solution exists to the equation $1 = 3x$ in $\mathbb{Z}_6$, what does that tell us about the element $3 \in \mathbb{Z}_6$?
- Why do you think $\mathbb{Z}_5$ appears to behave more like $\mathbb{R}$ than $\mathbb{Z}_6$ does?

Class Activity Problems 1 & 2

Consider the linear equation $y = 3x$ (and your corresponding graphs) from the Pre-Activity.

1. How many solutions to $0 = 3x$ exist in the following domains? What are they?

Domain	$\mathbb{R}$	$\mathbb{Z}_5$	$\mathbb{Z}_6$
# of Sol.	1	1	3
Sol. Set	$\{0\}$	$\{0\}$	$\{0, 2, 4\}$

2. How many solutions to $1 = 3x$ exist in the following domains? What are they?

Domain	$\mathbb{R}$	$\mathbb{Z}_5$	$\mathbb{Z}_6$
# of Sol.	1	1	0
Sol. Set	$\{\frac{1}{3}\}$	$\{2\}$	$\emptyset$

Before allowing undergraduates to move on to Problem 3, hold a brief whole-class discussion to discuss key answers to Problems 1 and 2 and also to verify that the class is well-positioned to engage in Problem 3. Give undergraduates time to first discuss Problem 3(a) in small groups.

Class Activity Problem 3 : Part a

3. Based on his work in Problem 1, Omar guesses that, in $\mathbb{Z}_{10}$, the equation $0 = 3x$ will have multiple solutions.

   (a) Why do you think Omar might have made this hypothesis?

   Sample Responses:

   - Omar might think that equations in $\mathbb{Z}_{10}$ behave more like equations in $\mathbb{Z}_6$ because both 6 and 10 are even numbers, unlike 5.
   - Omar sees that both 10 and 6 are composite numbers, unlike 5, and assumes that equations

will behave similarly in both $\mathbb{Z}_{10}$ and $\mathbb{Z}_6$.
- Omar might think that any modulus larger than 5 will have multiple solutions.

Discuss the following connection to teaching:

Discuss This Connection to Teaching

Problem 3(a) requires undergraduates to consider a hypothesis offered by a hypothetical student and to attempt to determine the mathematical reasoning that the student may have used to reach their hypothesis. Building capacity for considering the mathematical thinking of others strengthens undergraduates' own mathematical understanding and develops skills that they can apply in many settings, especially in the work of teaching, that require analyzing and valuing the thinking of others. Prospective teachers should be able to respond to student thinking when considering a reasonable, but incorrect, student answer by determining and addressing plausible reasoning trajectories, drawing out their conceptions and building on their understandings.

Ask undergraduates to work on Problems 3(b) and 3(c) in their groups. Think about the conclusions undergraduates offer in their groups and decide on the order in which you will have groups report out in a whole-class discussion. During whole-class discussion, ensure consensus is reached on the answers to Problems 3(b) and 3(c) before moving on to Problem 3(d).

Class Activity Problem 3 : Parts b & c

(b) In $\mathbb{Z}_{10}$, for which non-zero value(s) of a does the equation $0 = ax$ have a unique solution? Was Omar's hypothesis correct?

Solution:
For $a \in \{1, 3, 7, 9\}$ the solution is unique. This means Omar's hypothesis was not correct.

(c) In $\mathbb{Z}_{10}$, for which non-zero value(s) of a does the equation $1 = ax$ have a solution?

Solution:
For $a \in \{1, 3, 7, 9\}$ a solution exists.

Commentary:
Point out that Pre-Activity Problem 1(b) will help significantly with Problem 3(c) here. As you circulate the classroom, consider using the following prompts to generate discussion:

- How might you use a graph to help answer these questions?
- For Problem 3(c), will there ever be more than one solution for a given value of a? Why or why not?

Problems 3(b) and 3(c) are essentially asking undergraduates to find all the zero divisors and units in $\mathbb{Z}_{10}$, respectively. However, by framing these problems as solving equations, we reinforce the idea that an equation can be thought of as a property that applies precisely to elements of the solution set.

Problem 3(d) can be posed to the class at large and discussed without requiring undergraduates to work in groups first.

Class Activity Problem 3 : Part d

(d) Look back at your answers to Problems 3(b) and 3(c). What relationship do these integers have with 10, the modulus of $\mathbb{Z}_{10}$?

Solution:
The elements of $\mathbb{Z}_{10}$ that correctly answer 3(b) and 3(c) are those elements $x \in \mathbb{Z}_{10}$ for which $\gcd(x, 10) = 1$.

Discussion: The Structure of the Ring of Integers Modulo *n* (15 minutes)

To tie together all the parts of Problem 3, define the following vocabulary using the notation in your classroom textbook:

- Zero divisors.
- Integral domains.

Ask undergraduates for a conjecture about which elements of $\mathbb{Z}_{10}$ are zero divisors and which are units based on their work in Problem 3. To formalize this conjecture, introduce the following three theorems. As appropriate for your class (and considering time constraints), you may choose to prove some or all of the theorems in class.

Theorem 1. *Every nonzero element of $\mathbb{Z}_n$ is either a unit or a zero divisor.*

Proof. Let a be some nonzero element of $\mathbb{Z}_n$ and consider the set S of elements $\{0a, 1a, 2a \dots, (n-1)a\}$. If S has n distinct elements, then each one must match a unique element from $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$. So, $ba = 1$ for some b in $\mathbb{Z}_n$; since $\mathbb{Z}_n$ is also a commutative ring, $ab = 1$ as well and so a is a unit. If, on the other hand, S fails to contain n distinct elements, we must have $ba = ca$ for two distinct elements b and c in $\mathbb{Z}_n$. Consequently, $(b-c)a = 0$ and a is a zero divisor. $\square$

Theorem 2. *An element a of $\mathbb{Z}_n$ is a unit if and only if it is relatively prime with n.*

Proof. Let $a \in \mathbb{Z}_n$ be a unit. Then, for some $b \in \mathbb{Z}_n$, $ab = 1$. If we interpret this equation as a statement about integers (not equivalence classes of integers), this means that ab and 1 differ by some multiple of n; that is, we have that $ab + kn = 1$ for some integer k. Consequently, if a and n have any common factor, then it must divide 1. Thus, $\gcd(a, n) = 1$.

Conversely, let $a \in \mathbb{Z}_n$ be nonzero and $\gcd(a, n) = 1$. Assume, for the sake of contradiction, that a is not a unit of $\mathbb{Z}_n$. Then, by the previous result, it is a zero divisor of $\mathbb{Z}_n$ and so there is a nonzero element b such that $ab = 0$. If we interpret this equation as a statement about integers (not equivalence classes of integers), this means that ab is some multiple of n; that is, we have that $ab = kn$. This implies that $n|ab$, but because $\gcd(a, n) = 1$, it must be that $n|b$. This forces b to be zero, which is a contradiction. So our assumption cannot hold—and thus a is a unit of $\mathbb{Z}_n$. $\square$

Theorem 3. *$\mathbb{Z}_n$ is both a field and an integral domain when n is prime, but neither when n is composite.*

Proof. If n is prime, for every nonzero element $a \in \mathbb{Z}_n$, $\gcd(a, n) = 1$. By Theorem 2, this implies that all the nonzero elements of the ring are units. Then, by Theorem 1, this implies that no nonzero elements of the ring are zero divisors. Hence, $\mathbb{Z}_n$ is both a field and an integral domain. On the other hand, if n is composite, it is clear that any of its factors are zero divisors. These elements cannot be units, so $\mathbb{Z}_n$ is neither a field nor an integral domain. $\square$

To proceed, ensure that undergraduates understand the flow of ideas across these three theorems. The following explanation may help facilitate this understanding: If a and n are relatively prime, then a is a unit in $\mathbb{Z}_n$. This means it has a unique multiplicative inverse a^{-1}. Because the inverse is unique, $ax = b$ must have a unique solution in $\mathbb{Z}_n$: namely, $x = a^{-1}b$. When $b = 0$, this shows that a is also not a zero divisor.

You might also wish to show that no element of any ring can ever be both a unit and a zero divisor: Assume $a \neq 0$ is both a unit and a zero divisor. Then, $\exists\, b \neq 0$ such that $ab = 0$ and $b = 1b = (a^{-1}a)b = a^{-1}(ab) = 0$, which is a contradiction.

Class Activity: Problems 4–6 (20 minutes)

Before asking your class to begin Problem 4, initiate a brief discussion by asking undergraduates what familiar techniques we use in $\mathbb{R}$ to solve for the roots of quadratic expressions and record their answers (on the board, at the document camera, or post their answers). Then, ask undergraduates to work in small groups on Problem 4. While they work, ask them to think about which of the recorded techniques may no longer work as expected. Also, have them take note of new methods that can be used that would not apply when working in $\mathbb{R}$.

Consider asking any of the following prompts to promote discussion as you circulate the classroom:

- How do you expect finding the roots of this equation in $\mathbb{Z}_{10}$ to be similar to or different from working in $\mathbb{R}$?
- Which of these methods seems the most difficult to use in $\mathbb{Z}_{10}$? Easiest? Why?
- Make a prediction about the number of solutions to this equation.

Ask the groups to share out and adjust the list accordingly to reflect what methods may or may not work in $\mathbb{Z}_{10}$. If no one brought it up, point out that for small enough finite rings it is completely valid to check every element individually. Add this method to the previously compiled list.

Class Activity Problem 4

For Problems 4–6, consider the quadratic equation $x^2 - 5x + 6 = 0$ in $\mathbb{Z}_{10}$.

4. What are some ways that you might attempt to solve this equation for x?

   Sample Responses:

   - Factoring or applying the quadratic formula.
   - Completing the square.
   - Graphing the equation and looking for its roots.
   - Checking all the elements of $\mathbb{Z}_{10}$ to see which of them solve the equation.

Encourage undergraduates to find the proposed solution and others in Problem 5(a) using whatever method they like from the new list. Allow them to work in small groups for both parts of Problem 5.

Class Activity Problem 5

5. Notice that we can factor the left-hand side of this equation to obtain $(x-2)(x-3) = 0$, from which we find that $x - 2 = 0$ or $x - 3 = 0$. This yields the solutions $x = 2$ and $x = 3$.

   (a) Verify that $x = 7$ is also a solution. Are there any more? Why do you think factoring did not yield *all* the solutions?

   Solution:
   Because $7^2 - 5 \cdot 7 + 6 = 49 - 35 + 6 = 9 - 5 + 6 = 10 = 0$, 7 is a solution. We can also see that 8 is a solution: $8^2 - 5 \cdot 8 + 6 = 64 - 40 + 6 = 4 - 0 + 6 = 10 = 0$. Factoring did not yield all solutions because, as we saw in Problem 1, sometimes rings contain zero divisors. If so, we can't claim that $a \cdot b = 0 \Rightarrow a = 0$ or $b = 0$.

   (b) What important property of $\mathbb{R}$ are we using when we find the roots of a factored expression and claim those roots constitute the entire solution set?

   Solution:
   The zero product property.

For Problem 5(a), allow for some informality in undergraduates' responses since Problem 5(a) aims to elicit student thinking that scaffolds their more formal responses to Problem 5(b). When discussing Problem 5(b), discuss the following connection to teaching:

Discuss This Connection to Teaching

Solving quadratic equations in one variable is an ubiquitous standard in high school algebra. By attempting to reproduce familiar root-finding techniques outside of $\mathbb{R}$, prospective teachers are able to identify aspects of solving polynomial expressions that high school students might take for granted. That is, in high school, $\mathbb{R}$ has the *zero product property* only because it is an integral domain; similarly, every element of $\mathbb{R}$ has a multiplicative inverse only because $\mathbb{R}$ is also a field. Considering equation-solving in these other settings enables prospective teachers an anchor for applying reasoning that follows from $\mathbb{R}$ being a field and provides a foundation for looking forward to the introduction of other algebraic structures that may not have these familiar properties. This look back to equation-solving in high school underscores the important ideas that support the processes taught for solving equations.

There are ways to possibly determine additional solutions to a quadratic equation via the factoring technique, such as rewriting the coefficients of a given quadratic expression via equivalent elements of the ring. In this problem, $x^2 - 5x + 6 = x^2 + 5x + 6$ in $\mathbb{Z}_{10}$. So, we can also say $x^2 + 5x + 6 = 0 \Rightarrow (x+2)(x+3) = 0$, and so $x = -2 = 8$ or $x = -3 = 7$.

Another way to use factoring to find the entire solution set is, instead of setting each factor equal to zero, setting each factor equal to a pair of zero divisors. Then, if both factors give the same x value, that is a solution. For example, $x - 2 = 5$ and $x - 3 = 4$ both give $x = 7$, so this is a solution. On the other hand, $x - 2 = 5$ and $x - 3 = 8$ give $x = 7$ and $x = 1$ respectively, so we can make no claims about whether 7 or 1 are additional solutions.

Finally, give undergraduates a few minutes to discuss Problem 6 in small groups.

Class Activity Problem 6

6. Attempt to apply the quadratic formula to the above equation. What difficulties do you encounter?

   Sample Responses:

   - Modular arithmetic is more difficult than the usual arithmetic.
   - The quadratic formula requires us to "divide by 2," which does not make sense in rings where 2 is not a unit.
   - Taking square roots is difficult in general—even in a field, there is no guarantee that a particular element will even have a square root. In non-fields, certain elements may have a different number of square roots than we might expect.

Call the class back together for a whole-class discussion. Give groups a chance to report out first on their successes and obstacles encountered while attempting Problem 6. The most prominent point to raise is that the quadratic formula relies on multiplication by $1/2a$, which may not be well-defined if $2a$ is not a unit.

Like factoring, the quadratic formula is not a lost cause. In this case, we can write $2x = 5 + \sqrt{25 - 24} \Rightarrow 2x = 5 + \sqrt{5 + 6} \Rightarrow 2x = 5 + \sqrt{1}$, which yields two equations: $2x = 5 + 1 = 6$ and $2x = 5 + 9 = 4$. It can be seen that both equations have two solutions, resulting in the complete solution set of $\{2, 3, 7, 8\}$.

Wrap-Up (5 Minutes)

Recap the lesson briefly for the class:

- Some rings have zero divisors; those that don't are called integral domains and have the zero product property, which we rely upon when applying familiar processes for solving for the roots of a quadratic equation.

- $\mathbb{Z}_p$ is an integral domain and a field for p prime.

When working outside of a field or an integral domain to solve linear or quadratic equations, certain techniques that we learned in high school (such as factoring or the quadratic formula) become unreliable or much more difficult to apply.

We end this lesson with the use of an exit ticket. See Chapter 1 for advice about concluding mathematics lessons using exit tickets.

Homework Problems

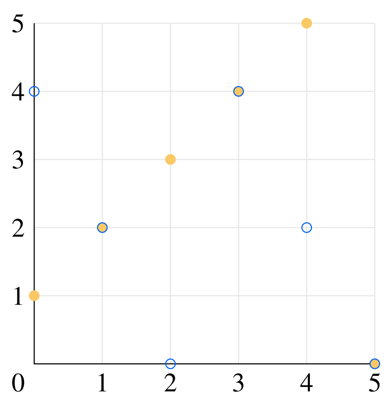
At the end of the lesson, assign the following homework problems.

Besides finding solutions to linear equations, another common task in high school algebra is the converse: given two points which lie on a line, find the corresponding equation to which they are both solutions. Problem 1 illustrates again that working outside of a field can have unexpected consequences.

Homework Problem 1

1. In $\mathbb{R} \times \mathbb{R}$, for any two distinct points A and B, there exists a unique line containing them. Show this statement is not true in $\mathbb{Z}_6 \times \mathbb{Z}_6$ by finding the equations of two distinct lines that both contain the points $(1, 2)$ and $(3, 4)$. [Recall that for a set A of elements (numbers) with a well-defined notion of addition and multiplication, we define a *line over A* as the solution set to an equation of the form $ax + by = c$ for some fixed values of $a, b, c \in \mathrm{A}$. That is, a line is the set of all ordered pairs $(x, y) \in \mathrm{A} \times \mathrm{A}$ that make $ax + by = c$ a true statement in A. The *graph of a line* is a scatter plot of the solution set on a coordinate plane, usually one with perpendicular axes marked by the elements of A.]

   Solution:
   The lines $y = x + 1$ and $y = 4x + 4$ are two examples, illustrated below.



   More generally, a line $ax + by = c$ in $\mathbb{Z}_6 \times \mathbb{Z}_6$ containing the points $(1, 2)$ and $(3, 4)$ must satisfy $a + 2b = c$ and $3a + 4b = c$, so $a + 2b = 3a + 4b \Rightarrow 4a + 4b = 0 \Rightarrow 4(a + b) = 0 \Rightarrow a + b \in \{0, 3\}$. If $a + b = 0$, then $a + 2b = c \Rightarrow (a + b) + b = c \Rightarrow b = c$, so $a + c = 0$ and $c = 5a$. Similarly, if $a + b = 3$, then $a + 2b = c \Rightarrow (a + b) + b = c \Rightarrow 3 + b = c \Rightarrow b = c + 3$, so $a + 3 + c = 3 \Rightarrow c = 5a$. That is, any line $ax + by = 5a$ where $a + b \in \{0, 3\}$ will work.

Problem 2 prompts undergraduates to examine and guide another student's thinking which helps them grow their own understanding of the topic. Here, the undergraduates learn that the structure of a ring does not necessarily transfer to a direct product of that ring with itself.

Homework Problem 2

2. Artyom says that since $\mathbb{R}$ is an integral domain, then the set of ordered pairs $\mathbb{R} \times \mathbb{R}$ must also be an integral domain under the operations given below:

$$(a, b) \oplus (c, d) = (a + c, b + d)$$
$$(a, b) \otimes (c, d) = (a \cdot c, b \cdot d)$$

(a) Why is Artyom incorrect?

Solution:
While $\mathbb{R} \times \mathbb{R}$ is a ring under the operations described above, it is not an integral domain. For $a, d \neq 0$, it is clear that neither $(a, 0)$ nor $(0, d)$ are "equal to zero". That is, neither is the additive identity. Then, $(a, 0) \otimes (0, d) = (0, 0)$ implies that $(a, 0)$ and $(0, d)$ are zero divisors.

(b) What question would you ask Artyom to help him understand his error? Why would your question be helpful?

Sample Responses:
- Can you find a pair of elements in $\mathbb{R} \times \mathbb{R}$ whose product has at least one component that is zero? Hopefully, this question will lead Artyom to consider if he could simultaneously make the other component equal to zero, leading to zero divisors.
- Can a nonzero element of $\mathbb{R} \times \mathbb{R}$ still contain a zero as one of its components? With this question, I want to lead Artyom towards possible counterexamples.

A useful alternative definition for integral domains (a ring is an integral domain if and only if cancellation law holds) is introduced and proven in Problem 3.

Homework Problem 3

3. Let R be a commutative ring in which the multiplicative identity and additive identity are distinct elements.

(a) Prove that if R is an integral domain, then for $a, b, c \in$ R and $a \neq 0$, $a \cdot b = a \cdot c \Rightarrow b = c$.

Solution:
Let $a \neq 0$. Then, $a \cdot b = a \cdot c \Rightarrow a \cdot b - a \cdot c = 0 \Rightarrow a \cdot (b - c) = 0$. R is an integral domain and $a \neq 0$, so by the fact that R contains no zero divisors we conclude that $b - c = 0 \Rightarrow b = c$.

(b) Prove that if $\forall\, a, b, c \in$ R with $a \neq 0$ we have that $a \cdot b = a \cdot c \Rightarrow b = c$, then R is an integral domain.

Solution:
Let $a \neq 0$. Let $a \cdot b = 0$. If we show that $b = 0$, then we have established that R has no zero divisors and is an integral domain. But $a \cdot b = 0 \Rightarrow a \cdot b = a \cdot 0$. By hypothesis, we now conclude $b = 0$.

After dealing with linear and quadratic functions in the Class Activity, a natural next step is to look for ways to solve simple cubic functions, as presented in Problem 4.

Homework Problem 4

4. When looking for solutions to the equation $x^3 = 1$ for $x \in \mathbb{Z}_{13}$, we see that $x = 1$ clearly works. To find other solutions, it might be useful to observe that every element in $\mathbb{Z}_{13}$ corresponds to a value 2^n

for some n by completing the following table of values in $\mathbb{Z}_{13}$. [Hint: Double the values in the table from left to right, remembering to convert to modulo 13 when appropriate]

2^0	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}
1	2	4	8	3	6	12	11	9	5	10	7	1

Now, to find other solutions, we might use the table above to help; for example, $x^3 = 1 = 2^{12} = (2^4)^3 = 3^3$. Thus, 3 is also a solution. It turns out there is only one more solution to this equation. Find it and justify your answer by using powers of 2.

Solution:
The remaining answer is 9: $x^3 = 1 = 1^2 = (2^{12})^2 = 2^{24} = (2^8)^3 = 9^3$.

In fact, we can show that there are no other solutions. Given that every element of $\mathbb{Z}_{13}$ corresponds to a value of 2^n for some n. Then, if $x = 2^n$ is a solution to $x^3 = 1$, we have that $2^{3n} = 1$. From the table, this means that $3n$ must be a multiple of 12 and so n is a multiple of 4. Then: $x = 2^n = 2^{4k} = 16^k = 3^k$, where k is some non-negative integer. Plugging in values of k, we see that the only solutions are the three we have found.

In the Class Activity, undergraduates learned when the equation $ax = 0$ has a unique solution in $\mathbb{Z}_n$. Problem 5 broadens that understanding, such that undergraduates are able to describe the number of solutions to $ax = 0$ in $\mathbb{Z}_n$ when it is not unique.

Homework Problem 5

5. How many solutions does the equation $ax = 0$ have in $\mathbb{Z}_{12}$ for each nonzero a? Use your answer to make a hypothesis about the number of solutions to the equation $ax = 0$ in $\mathbb{Z}_n$ when a is nonzero.

Sample Response:

a	0	1	2	3	4	5	6	7	8	9	10	11
#	X	1	2	3	4	1	6	1	4	3	2	1

I notice that the number of solutions to $ax = 0$ in $\mathbb{Z}_{12}$ is $\gcd(a, 12)$. I assume that this also holds in $\mathbb{Z}_n$.

In addition to solving linear and quadratic equations, high school students are often tasked with solving systems of equations. Problem 6 extends the connections to teaching made during the Class Activity to include systems of equations; that is, prospective teachers investigate how certain familiar techniques (in this case, the substitution or elimination methods) may or may not translate outside the field of real numbers.

Homework Problem 6

6. Solve the system of linear equations given below in the following rings, if possible.

$$2x + y = 4$$
$$x + 2y = 0$$

(a) $\mathbb{Z}_5 \times \mathbb{Z}_5$

Solution:

First, $x + 2y = 0 \Rightarrow x = 3y$. Substituting into the other equation, we have $2x + y = 4 \Rightarrow 2(3y) + y = 4 \Rightarrow 7y = 4 \Rightarrow 2y = 4$. Multiplying both sides by 3 yields $y = 12 = 2$. Then, $x = 3y \Rightarrow x = 6 = 1$, so $(1, 2)$ is a unique solution.

(b) $\mathbb{Z}_6 \times \mathbb{Z}_6$

Solution:

First, $x + 2y = 0 \Rightarrow x = 4y$. Substituting into the other equation, we have $2x + y = 4 \Rightarrow 2(4y) + y = 4 \Rightarrow 9y = 4 \Rightarrow 3y = 4$. But $\gcd(3, 6)$ is not 1, so this equation does not have a unique solution. In fact, since 3 has no multiplicative inverse in $\mathbb{Z}_6$, $3y = 4$ has no solution and neither does the system of equations.

(c) Was your process for solving parts (a) and (b) the same? Why or why not? What difficulties arose in parts (a) and (b)?

Sample Responses:

- I tried to answer both part (a) and part (b) algebraically, but was only successful in part (a). In part (b), the fact that 3 has no multiplicative inverse in $\mathbb{Z}_6$ prevented me from solving the equation.
- Graphing both lines on a coordinate axis reveals a point of intersection in $\mathbb{Z}_5 \times \mathbb{Z}_5$, but the scatter plots don't overlap in $\mathbb{Z}_6 \times \mathbb{Z}_6$.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Assessment Problems 1 & 2

1. List all the nonzero values of a which give the equation $ax = 0$ a unique solution in the following rings:

   (a) $\mathbb{Z}_{13}$

   Solution:

   Since 13 is prime, $\gcd(a, 13) = 1 \; \forall\, a \in \mathbb{Z}_{13}$. This means that $ax = 0$ has a unique solution $\forall\, a \in \mathbb{Z}_{13}$, so $a \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

   (b) $\mathbb{Z}_{14}$

   Solution:

   We need only identify the elements of $\mathbb{Z}_{14}$ which are relatively prime with 14. These are the values of a for which $ax = 0$ will have a unique solution. So, $a \in \{1, 3, 5, 9, 11, 13\}$.

2. Thuy's work for finding solutions to $x^2 - x = 0$ in $\mathbb{Z}_4$ is shown below.

   $x^2 - x = 0$

   $x(x-1) = 0$

   Therefore, either

   $x = 0$ or $x - 1 = 0$

   The solution set is $\{0, 1\}$

(a) From her work, what assumption does Thuy seem to be making about $\mathbb{Z}_4$? Is this assumption correct?

Sample Response:

Thuy assumes that the zero product property holds in $\mathbb{Z}_4$. The zero property does not hold in $\mathbb{Z}_4$. Take $2 \cdot 2 = 4 = 0 \in \mathbb{Z}_4$, but $2 = 0 \in \mathbb{Z}_4$. However, there does not exist an $x \neq 0,1 : (x)(x-1) = 0 \in \mathbb{Z}_4$, so Thuy thinks this is okay to assume.

(b) Thuy checks each element of $\mathbb{Z}_4$ and verifies that her solution set is correct. Her teacher asks her to attempt to solve the same equation, this time in $\mathbb{Z}_6$. What is the teacher hoping Thuy will understand about her approach by working in $\mathbb{Z}_6$?

Sample Response:

Thuy happened to find all the solutions in $\mathbb{Z}_4$ despite the fact that $\mathbb{Z}_4$ is not an integral domain, but she might not be so lucky with a different quadratic equation. If Thuy works the same problem the same way in $\mathbb{Z}_6$, she will not find all the solutions since 3 and 4 also solve the equation. This will help her see that she cannot (reliably) apply the zero product property in $\mathbb{Z}_n$ when n is composite.

8.6 References

[1] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from http://www.corestandards.org/

8.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** SOLVING EQUATIONS (page 1 of 2)

1. Recall that $\mathbb{Z}_n$ is the set of *equivalence classes* on the integers, where two integers are in the same equivalence class if and only if they both have the same (smallest, non-negative) remainder when divided by n. The set $\mathbb{Z}_n$ contains n such equivalence classes which, canonically, are represented by the possible remainders when an integer is divided by n: $\{0, 1, \ldots, n-2, n-1\}$.

   (a) Fill in the following chart with the representative of each integer's equivalence class in $\mathbb{Z}_{10}$.

Integer	36	17	-4	-17
Representative in $\mathbb{Z}_{10}$				

   If we are careful, we can also (sometimes) represent non-integers as elements of $\mathbb{Z}_n$. For example, if we interpret the notation "$1/3$" as "the element you multiply by 3 to get 1," we would then consider 7 in $\mathbb{Z}_{10}$ a representative of "$1/3$", since $3 \cdot 7 = 21 = 1$ (where $21 = 1$ because 21 has remainder 1 when divided by 10). Furthermore, no other element of $\mathbb{Z}_{10}$ has this property.

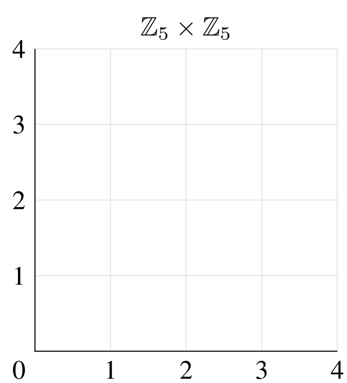
   (b) Fill in the following chart with the representative in $\mathbb{Z}_{10}$, if it exists.

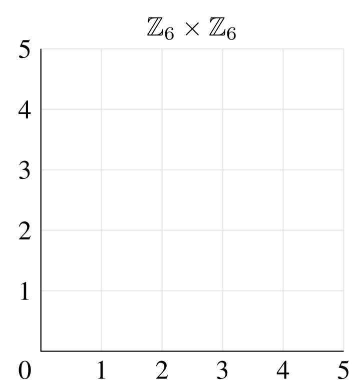
"Non-integer"	"$1/1$"	"$1/2$"	"$1/3$"	"$1/4$"	"$1/5$"	"$1/6$"	"$1/7$"	"$1/8$"	"$1/9$"
Representative in $\mathbb{Z}_{10}$	1		7						

2. Let A be a set of elements (numbers) with a well-defined notion of addition and multiplication. We define a *line over A* as the solution set to an equation of the form $ax + by = c$ for some fixed values of $a, b, c \in$ A. That is, a line is the set of all ordered pairs $(x, y) \in$ A $\times$ A that make $ax + by = c$ a true statement in A. The *graph of a line* is a scatter plot of the solution set on a coordinate plane, usually one with perpendicular axes marked by the elements of A.

   For example, in the system of real numbers, the set of all ordered pairs that make $y = 3x$ a true statement in $\mathbb{R}$ has a graph which is a continuous straight geometric line of slope 3 through the point $(0, 0)$ in our usual Cartesian coordinate system.

   Graph the line $y = 3x$ in the following spaces on the provided axes.





PRE-ACTIVITY: SOLVING EQUATIONS (page 2 of 2)

3. In solving the equation $x + 4 = 1 + 4x$ in $\mathbb{R}$, a student makes the following algebraic manipulations:

$$\begin{aligned} x + 4 &= 1 + 4x \\ 4 &= 1 + 3x \\ 3 &= 3x \\ 1 &= x \end{aligned}$$

The student then concludes that $x = 1$ is the solution set to $x + 4 = 1 + 4x$ in $\mathbb{R}$.

(a) Describe the mathematical justification for each step in the student's solution.

(b) To solve $x + 4 = 1 + 4x$ for x in $\mathbb{Z}_5$, are we allowed to repeat the process the student used (in $\mathbb{R}$) as presented above? Does this process yield the entire solution set to the equation? Explain.

(c) To solve $x + 4 = 1 + 4x$ for x in $\mathbb{Z}_6$, are we allowed to repeat the process the student used (in $\mathbb{R}$) as presented above? Does this process yield the entire solution set to the equation? Explain.

NAME: **CLASS ACTIVITY:** SOLVING EQUATIONS (page 1 of 2)

Consider the linear equation $y = 3x$ (and your corresponding graphs) from the Pre-Activity.

1. How many solutions to $0 = 3x$ exist in the following domains? What are they?

Domain	$\mathbb{R}$	$\mathbb{Z}_5$	$\mathbb{Z}_6$
# of Sol.			
Sol. Set			

2. How many solutions to $1 = 3x$ exist in the following domains? What are they?

Domain	$\mathbb{R}$	$\mathbb{Z}_5$	$\mathbb{Z}_6$
# of Sol.			
Sol. Set			

3. Based on his work in Problem 1, Omar guesses that, in $\mathbb{Z}_{10}$, the equation $0 = 3x$ will have multiple solutions.

   (a) Why do you think Omar might have made this hypothesis?

   (b) In $\mathbb{Z}_{10}$, for which non-zero value(s) of a does the equation $0 = ax$ have a unique solution? Was Omar's hypothesis correct?

   (c) In $\mathbb{Z}_{10}$, for which non-zero value(s) of a does the equation $1 = ax$ have a solution?

   (d) Look back at your answers to Problems 3(b) and 3(c). What relationship do these integers have with 10, the modulus of $\mathbb{Z}_{10}$?

For problems 4–6, consider the quadratic equation $x^2 - 5x + 6 = 0$ in $\mathbb{Z}_{10}$.

4. What are some ways that you might attempt to solve this equation for x?

5. Notice that we can factor the left-hand side of this equation to obtain $(x-2)(x-3) = 0$, from which we find that $x - 2 = 0$ or $x - 3 = 0$. This yields the solutions $x = 2$ and $x = 3$.

   (a) Verify that $x = 7$ is also a solution. Are there any more? Why do you think factoring did not yield *all* the solutions?

   (b) What important property of $\mathbb{R}$ are we using when we find the roots of a factored expression and claim those roots constitute the entire solution set?

6. Attempt to apply the quadratic formula to the above equation. What difficulties do you encounter?

NAME: **HOMEWORK PROBLEMS:** SOLVING EQUATIONS (page 1 of 1)

1. In $\mathbb{R} \times \mathbb{R}$, for any two distinct points A and B, there exists a unique line containing them. Show this statement is not true in $\mathbb{Z}_6 \times \mathbb{Z}_6$ by finding the equations of two distinct lines that both contain the points $(1, 2)$ and $(3, 4)$. [Recall that for a set A of elements (numbers) with a well-defined notion of addition and multiplication, we define a *line over A* as the solution set to an equation of the form $ax + by = c$ for some fixed values of $a, b, c \in \mathrm{A}$. That is, a line is the set of all ordered pairs $(x, y) \in \mathrm{A} \times \mathrm{A}$ that make $ax + by = c$ a true statement in A. The *graph of a line* is a scatter plot of the solution set on a coordinate plane, usually one with perpendicular axes marked by the elements of A.]

2. Artyom says that since $\mathbb{R}$ is an integral domain, then the set of ordered pairs $\mathbb{R} \times \mathbb{R}$ must also be an integral domain under the operations given below:

$$(a, b) \oplus (c, d) = (a + c, b + d)$$
$$(a, b) \otimes (c, d) = (a \cdot c, b \cdot d)$$

   (a) Why is Artyom incorrect?
   (b) What question would you ask Artyom to help him understand his error? Why would your question be helpful?

3. Let R be a commutative ring in which the multiplicative identity and additive identity are distinct elements.

   (a) Prove that if R is an integral domain, then for $a, b, c \in \mathrm{R}$ and $a \neq 0$, $a \cdot b = a \cdot c \Rightarrow b = c$.
   (b) Prove that if $\forall\, a, b, c \in \mathrm{R}$ with $a \neq 0$ we have that $a \cdot b = a \cdot c \Rightarrow b = c$, then R is an integral domain.

4. When looking for solutions to the equation $x^3 = 1$ for $x \in \mathbb{Z}_{13}$, we see that $x = 1$ clearly works. To find other solutions, it might be useful to observe that every element in $\mathbb{Z}_{13}$ corresponds to a value 2^n for some n by completing the following table of values in $\mathbb{Z}_{13}$. [Hint: Double the values in the table from left to right, remembering to convert to modulo 13 when appropriate]

2^0	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	2^{11}	2^{12}
				3								1

   Now, to find other solutions, we might use the table above to help; for example, $x^3 = 1 = 2^{12} = (2^4)^3 = 3^3$. Thus, 3 is also a solution. It turns out there is only one more solution to this equation. Find it and justify your answer by using powers of 2.

5. How many solutions does the equation $ax = 0$ have in $\mathbb{Z}_{12}$ for each nonzero a? Use your answer to make a hypothesis about the number of solutions to the equation $ax = 0$ in $\mathbb{Z}_n$ when a is nonzero.

6. Solve the system of linear equations given below in the following rings, if possible.

$$2x + y = 4$$
$$x + 2y = 0$$

   (a) $\mathbb{Z}_5 \times \mathbb{Z}_5$
   (b) $\mathbb{Z}_6 \times \mathbb{Z}_6$
   (c) Was your process for solving parts (a) and (b) the same? Why or why not? What difficulties arose in parts (a) and (b)?

NAME: **ASSESSMENT PROBLEMS:** SOLVING EQUATIONS (page 1 of 1)

1. List all the nonzero values of a which give the equation $ax = 0$ a unique solution in the following rings:

   (a) $\mathbb{Z}_{13}$

   (b) $\mathbb{Z}_{14}$

2. Thuy's work for finding solutions to $x^2 - x = 0$ in $\mathbb{Z}_4$ is shown below.

   $x^2 - x = 0$
   $x(x-1) = 0$
   Therefore, either
   $x = 0$ or $x - 1 = 0$
   The solution set is $\{0, 1\}$

   (a) From her work, what assumption does Thuy seem to be making about $\mathbb{Z}_4$? Is this assumption correct?

   (b) Thuy checks each element of $\mathbb{Z}_4$ and verifies that her solution set is correct. Her teacher asks her to attempt to solve the same equation, this time in $\mathbb{Z}_6$. What is the teacher hoping Thuy will understand about her approach by working in $\mathbb{Z}_6$?

9

Groups of Transformations

Abstract (Modern) Algebra I

Andrew Kercher, *Simon Fraser University*
James A. M. Álvarez, *The University of Texas at Arlington*
Kyle Turner, *The University of Texas at Arlington*

9.1 Overview and Outline of Lesson

This lesson follows a typical introduction to the group axioms. Undergraduates investigate sets of matrices that represent familiar geometric transformations: rotations about the origin and reflections about lines through the origin. Working in $\mathbb{R}^2$, these sets are examined for group structure and commutativity under matrix multiplication by building on undergraduates' existing knowledge of the geometry of these types of transformations. This supports prospective teachers' mathematical knowledge for considering ways to justify (or redirect) certain intuitions high school geometry students may have about these transformations, in particular whether "order matters" when applying a sequence of transformations to an object.

1. Launch—Pre-Activity
   Undergraduates complete this assignment prior to the lesson. In it, undergraduates engage in ascribing mathematical formality to the idea of when "order matters" by considering the closure, associativity, and commutativity of the composition of rotations about the origin. Then, they explore two sets of 2×2 matrices by varying a parameter and observing how the corresponding matrices differently affect vectors in $\mathbb{R}^2$. These sets of matrices are characterized as rotations and reflections, setting the stage for the Class Activity.

2. Explore—Class Activity

   - *Problems 1 & 2*:
     The group axioms (and commutativity) are explored with respect to the set of rotations (about the origin) by making geometric interpretations of matrix equations. Undergraduates conclude that this set is an abelian group under matrix multiplication.
   - *Problems 3 & 4*:
     The group axioms (and commutativity) are explored with respect to the set of reflections (about lines through the origin) by reasoning geometrically about the existence of inverses and an identity element. Undergraduates conclude that this set is neither a group nor commutative under matrix multiplication.
   - *Problems 5 & 6*:
     The commutativity of elements in the set of rotations and reflections is explored by helping a hypothetical geometry student reason about the congruency of figures. Undergraduates conclude that the set of rotations and reflections are not commutative under matrix multiplication.

3. Closure—Wrap-Up
   The instructor wraps up the lesson by reviewing the two sets of matrices, which rigid motions they represent, and whether these sets are (abelian) groups under matrix multiplication. This information supports prospective teachers in discussions with their future students while considering whether "order matters" when applying rigid motions to geometric shapes.

9.2 Alignment with College Curriculum

Undergraduates explore sets of matrices that represent familiar geometric transformations for group structure, thereby adding an accessible, easily-visualized example of a (non-) group to their understanding of the topic. They are then asked to consider whether "order matters" when applying each type of transformation, laying the groundwork for discussions about the associativity and commutativity of binary operations.

9.3 Links to School Mathematics

Matrices are often presented in high school mathematics classrooms as arbitrary objects that calculators use to solve unwieldy systems of equations; on the other hand, rigid motions are used to establish the congruence of shapes but are not treated as especially formal mathematical objects. These shortcomings are addressed simultaneously by demonstrating that two rigid motions (reflections and rotations) can be represented as linear transformations on $\mathbb{R}^2$ via matrices. The sets of these matrix transformations, once explored for group structure, are used to interpret a high school geometry problem through the lens of abstract algebra.

This lesson highlights:

- Using group structure to formalize ideas from school mathematics, such as "undoing" an operation and establishing congruency.
- Connecting matrix transformations to familiar concepts from high school geometry.

This lesson addresses several mathematical knowledge and practice expectations included in high school standards documents, such as the Common Core State Standards for Mathematics (CCSSM, 2010). For example, it is expected that high school students learn to use matrix arithmetic and the underlying properties of those arithmetic operations. This includes the fact that matrix multiplication is not commutative, but is associative; that the identity matrix and the zero matrix are the matrix analogues of 1 and 0 in the real numbers, respectively; and that matrices can act as transformations on vectors under matrix multiplication (see CCSS.MATH.CONTENT.HSM.VM.C for a complete list of properties). High school students also work with rigid motions (although not usually in the form of matrices) in order to establish the congruency of triangles on a plane (c.f. CCSS.MATH.CONTENT.HSG.CO.B.7). Finally, this lesson emphasizes the need for viable mathematical arguments, encourages undergraduates to look for and make use of structural similarities, and provides opportunities to both critique the reasoning of others and to practice the appropriate transference of reasoning from one setting to another.

9.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- Matrix multiplication, including how a 2×2 matrix acts on a vector from $\mathbb{R}^2$;
- The definitions of a group and an abelian group.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Decide whether sets of matrices are (abelian) groups under a particular binary operation;

- Explain whether "order matters" when considering a binary operation on a set;
- Translate the above conclusions into geometric language by describing the effect of (products of) matrices on a vector;
- Analyze hypothetical student work to evaluate reasoning about transformations.
- Pose guiding questions to help a hypothetical student connect geometric understandings about transformations to their assertions about whether corresponding sets of matrices are groups under a given binary operation.

Anticipated Length

One 75-minute class session.

Materials

The following materials are required for this lesson:

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

9.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to be completed in preparation for this lesson.

> *We recommend that you collect this Pre-Activity the day before the lesson so that you can review undergraduates' responses before you begin the Class Activity. This will help you determine if you need to spend additional time reviewing the solutions to the Pre-Activity with your undergraduates.*

Pre-Activity Review (10 minutes)

Briefly discuss your undergraduates' responses to each part of Problem 1. If appropriate for your class, you may wish to ask for some other examples of operations for which order does or does not matter. Make sure that undergraduates are specific about the set of objects on which the operation is acting. For example, "order doesn't matter" when multiplying integers, but it does when multiplying matrices.

> *If your class has already defined "order of a group" or "order of an element," you may wish to clarify here that the word "order" in the phrase "order matters" does not refer to either of these precise, mathematical definitions. Also, it is not exactly related to the "order of operations" either, since we are only considering one operation at a time. It is precisely this ambiguity that we hope to alleviate by defining, rigorously and mathematically, what we mean when we say "order matters."*
>
> *Undergraduates may be unclear exactly what is meant by "composition" in Problem 1. You may need to clarify that when we ask if rotations are closed under composition we really mean: can the combined act of rotating a vector twice about the origin be represented, in total, as one rotation?*

Pre-Activity Problem 1

Consider the sum $2 + 6 + 4 + 7$. When people say that "order does not matter" when computing such a sum they actually mean two things: the order of the individual terms of the sum can be rearranged without affecting the final result (for instance, $7 + 4 + 6 + 2$ and the original sum are sure to each give the same answer, 19) and, moreover, the order in which one chooses to compute the individual addition operations is unimportant (for instance, $((2+6)+4)+7$ and $2+((6+4)+7)$ both yield the same final result of 19). This conclusion relies on the three fundamental beliefs of integer arithmetic:

- Integer addition is **closed**; that is, $a+b$ is itself an integer for all integers a and b.
- Integer addition is **commutative**; that is, $a+b=b+a$ for all integers a and b.
- Integer addition is **associative**; that is, $(a+b)+c=a+(b+c)$ for all integers a, b, and c.

1. Consider the set of all rotations about the origin of the plane.
   [Recall that transformations (e.g., rotations) are functions. As such, for rotations r_α and r_β on $\mathbb{R}^2$, the *composition* of r_α followed by r_β, $r_\beta \circ r_\alpha$ is defined by $r_\beta \circ r_\alpha(P) = r_\beta(r_\alpha(P))$ where $P \in \mathbb{R}^2$.]

   (a) Is this set closed under composition? Explain.

   Sample Response:
   Yes. Rotating about the origin by some number of radians, followed by then rotating again by a different amount, is equivalent to rotating by the sum of the two angles of rotation. So the composition of two rotations is again a rotation.

   (b) Do rotations commute with each other under composition? Explain.

   Sample Response:
   Yes. I can rotate a vector by two different amounts in either order and the resultant vector will be the same amount either way.

   (c) Do rotations about the origin satisfy the associative law under composition? Explain.

   Sample Responses:
   - When rotating a vector by three different angles, the way that I group the rotations before doing them does not affect the overall rotation.
   - Yes. If you associate a rotation with the corresponding angle measure of that rotation (for example, you treat a rotation of $\frac{\pi}{6}$ as just $\frac{\pi}{6}$), then the composition of rotations about the origin is the same as real number addition. Real number addition is clearly associative, so composition of angles must also be.

   (d) Does "order matter" when performing a series of rotations about the origin in the plane? Explain.

   Sample Response:
   No. Rotations appear to be closed under composition and both commutative and associative with respect to composition.

Undergraduates may struggle to justify Problem 1(c), even imprecisely. One approach is to appeal to the natural isomorphism between $\mathbb{R}$ mod 2π under addition and the set of rotations about the origin under composition. Use a degree of formality when handling this relationship that is appropriate for your class—for example, you might choose to wait until the introduction of the set Σ in the next problem to make this isomorphism more explicit.

Pre-Activity Problem 2

2. Consider the set Σ, given below.

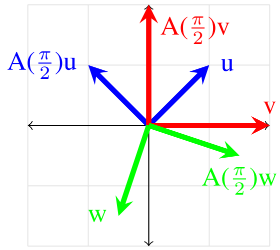
$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

   (a) Calculate $A(\frac{\pi}{2})$. Choose three nonzero vectors v_1, v_2, and v_3 in $\mathbb{R}^2$ that are not all scalar multiples of one another. Compute $A(\frac{\pi}{2})v_1$, $A(\frac{\pi}{2})v_2$, and $A(\frac{\pi}{2})v_3$. Sketch all six vectors on the same coordinate plane.

Sample Response:

First, note that $A(\frac{\pi}{2}) = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. Then, when using $u = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $v = \left[\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right]$, and $w = \left[\begin{smallmatrix} -1/2 \\ -3/2 \end{smallmatrix}\right]$, we have:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} -1/2 \\ -3/2 \end{bmatrix} = \begin{bmatrix} 3/2 \\ -1/2 \end{bmatrix}$$

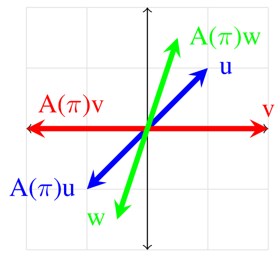


(b) Repeat the process of Problem 2(a) with $A(\theta)$ for a different nonzero value of θ and the same vectors.

Sample Response:

Choosing $A(\pi) = \left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$, with the same vectors as above:

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$
$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$
$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} -1/2 \\ -3/2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 3/2 \end{bmatrix}$$



(c) Write a geometric description of how an arbitrary matrix from Σ acts on vectors in $\mathbb{R}^2$ based on your sketches in Problems 2(a) and 2(b).

Solution:

Matrices of this type represent rotations about the origin. If θ is positive, then this appears as a counterclockwise rotation of θ radians; if θ is negative, then this appears as a clockwise rotation of $|\theta|$ radians.

If undergraduates are unsure about whether the vector's magnitude is preserved under these operations, encourage them to consider Problem 3 geometrically rather than through computations.

Pre-Activity Problem 3

3. Consider the set Φ, given below.

$$\Phi = \left\{ \mathrm{B}(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

(a) Repeat the process of Problem 2(a) with matrix $B(\frac{\pi}{2})$.

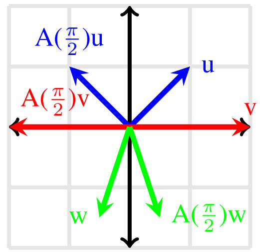
Sample Response:

This time, note that $B(\frac{\pi}{2}) = \left[\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. Then, with the same vectors as above:

$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$

$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1/2 \\ -3/2 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -3/2 \end{bmatrix}$



(b) Repeat the process of Problem 3(a) with $B(\theta)$ for a different value of θ and the same vectors.

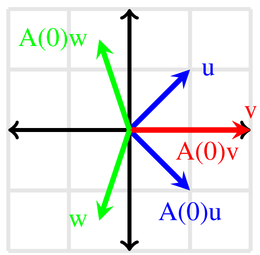
Sample Response:

First, we choose $B(0) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then, with the same vectors as above:

$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} -1/2 \\ -3/2 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 3/2 \end{bmatrix}$



(c) Write a geometric description of how an arbitrary matrix from Φ acts on vectors in $\mathbb{R}^2$ based on your sketches in Problems 3(a) and 3(b).

Solution:
Matrices of this type represent reflections about lines through the origin. The line of reflection forms an angle of θ radians with the positive x-axis.

Give undergraduates a few minutes to compare their answers for Problems 2(c) and 3(c) in small groups. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion. Ask groups to report out and reconcile any significant differences in responses. You may also want to facilitate discussion regarding the following:

- Based upon their analysis of Σ in Problem 2, you might ask undergraduates to specifically discuss $\mathrm{A}(\theta + 2k\pi)$ or to compare $\mathrm{A}(\theta)$ with $\mathrm{A}(-\theta)$.
- Based upon their analysis of Φ in Problem 3, you might ask undergraduates to compare $\mathrm{B}(\theta)$ with $\mathrm{B}(\theta + \pi)$.

Next, write the refined (correct) geometric descriptions of the sets on the board for them to reference throughout the Class Activity. It may help to include a geometric interpretation of the parameter θ for each set. So:

- Σ represents the set of rotations about the origin by θ radians.
- Φ represents the set of reflections about lines through the origin found by traveling θ radians from the positive x-axis.

After writing this list on the board, discuss the following connection to teaching:

Discuss This Connection to Teaching

Rigid motions (translations, reflections, and rotations) of geometric figures are used in high school geometry to assess the congruence of shapes. Translating this work into the language of matrix operations allows prospective teachers the chance to see how intuitive geometric ideas can be represented algebraically to provide an additional method of inquiry for supporting conjectures and conclusions.

Based upon recommendations of the CCSSM (2010), there has been a shift to a transformation approach to the teaching of geometric concepts of congruence, similarity, and symmetry. Mathematics educators have been calling for this shift in approach since the 1900s (Lai & Donsig, 2018). The emphasis on defining similarity based upon geometric transformations may allow students to apply this mathematical concept in meaningful ways (Seago, et al., 2013) and possibly address Ada and Kurtuluş' (2010) finding that students often seemed to know the algebraic definitions of translation and rotation, but not able to attend to the conceptual or associated geometric meanings. Nonetheless, Seago, et al. (2013) explain that attending to the recommendations of the CCSSM poses serious challenges for supporting both teachers and students in shifting from a traditional, static approach. Thus, providing opportunities for encountering ways to revisit the transformation approach to geometry targets supporting undergraduates' conceptual progress in this regard. Further, Lai and Donsig (2018) propose that "teaching geometry from a transformational approach provides an opportunity to showcase abstract algebraic ideas in ways that are accessible and relevant to secondary mathematics" (p. 63).

The purpose of Problem 4 is to prepare or enable undergraduates to quickly address the associativity of the particular 2×2 matrix subsets in the Class Activity. It is helpful to demonstrate that often, a set will inherit associativity under a particular operation from a more familiar set in which it is contained.

Pre-Activity Problem 4

4. Recall that 2×2 matrix multiplication is associative. On the other hand, 2×2 matrices do not always commute under matrix multiplication. Give an example of a pair of 2×2 matrices that do commute and a pair of 2×2 matrices that do NOT commute under matrix multiplication.

   Sample Response:
   Any matrix commutes with the identity matrix. On the other hand,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}$$

   Commentary:
   If you collected the Pre-Activity in advance, you might choose to showcase a relevant example or two from undergraduate submissions (such as a pair of diagonal matrices, which would be helpful for Class Activity Problem 3). Otherwise, only spend time reviewing this problem if their work reveals that review is needed.

Transition to the Class Activity by telling undergraduates that, now that we have identified these sets of matrices as particular geometric transformations, it is useful to explore the structure of these sets: we want to verify whether they are groups under multiplication so that we can use group theory as a tool to study them.

Class Activity: Problems 1 & 2 (20 minutes)

Distribute the Class Activity. Before undergraduates begin working on Problem 1 in small groups, remind them that they should not be computing the entries of any matrices or doing any matrix arithmetic—they should be "translating" the equations into an English sentence or a diagram, like the example provided in Problem 1(a). Allow the class to work on Problems 1(b) through 1(e) in small groups.

As you circulate the classroom and monitor undergraduates' responses, consider using any of the following questions to prompt further discussion:

- Besides $\theta = 2\pi$, what other values of θ give the identity matrix?
- Besides $-\theta$, what other angle measure could we use to "undo" a rotation by θ radians?
- Describe how the matrix product $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2)$ will transform a vector v in $\mathbb{R}^2$. Which matrix acts first on v? Does this matter?

Class Activity Problem 1

1. Consider the set Σ given below.

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

Restate the following equations in terms of the geometric effect matrices in Σ have on vectors in $\mathbb{R}^2$. Do not make any calculations—your answers should be written sentences or labeled sketches. A sample solution to part (a) and a partial solution to part (b) are provided.

(a) $\mathrm{A}(\theta_1)[\mathrm{A}(\theta_2)\mathrm{A}(\theta_3)] = [\mathrm{A}(\theta_1)\mathrm{A}(\theta_2)]\mathrm{A}(\theta_3)$

When rotating a vector by three different angles, the way that I group the rotations before doing them does not affect the overall rotation.

(b) $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_1 + \theta_2)$

Solution:

Rotating a vector by θ_2 radians and then by θ_1 radians is equivalent to rotating it by $\theta_1 + \theta_2$ radians.

(c) $\mathrm{A}(\theta)^{-1} = \mathrm{A}(-\theta)$

Solution:

The inverse of the matrix representing a rotation of θ radians is the matrix corresponding to a rotation of -θ radians. That is, the matrix that rotates the same amount in the opposite direction.

(d) $\mathrm{A}(2\pi) = \mathrm{I}$

Solution:

The matrix representing a rotation of 2π radians is just the identity matrix.

(e) $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_2)\mathrm{A}(\theta_1)$

Solution:

When composing two rotations, either one can be applied first. That is, rotating by θ_1 and then θ_2 radians is no different than rotating by θ_2 and then θ_1 radians.

After groups share their answers with the class, askl them to work on Problem 2 and emphasize that they should not be doing any matrix calculations for this problem. As you monitor the groups, you may point out, when needed, that they have already addressed most of this problem on the Pre-Activity. Consider making the following points as needed:

- To justify that Σ is closed and commutative under matrix multiplication, encourage undergraduates to appeal to their geometric intuition from Pre-Activity Problems 1(a) and 1(b).
- For associativity, refer undergraduates to Pre-Activity Problem 4.
- For identity, undergraduates will be tempted to compute $\mathrm{A}(2\pi)$ to see that it is the 2×2 identity matrix. Make sure they also consider why $\mathrm{A}(2k\pi) = \mathrm{I}$ makes geometric sense.
- To find an inverse without calculations, encourage undergraduates to think about how they might "undo" a rotation. That is, how they could return a vector to its original position after it has been rotated by θ radians.

Class Activity Problem 2

2. Explain whether Σ is an abelian group under matrix multiplication by analyzing your responses to Problem 1. Does "order matter" when multiplying together a string of matrices from Σ?

Solution:

- 1(a) corresponds with associativity. As we saw in Problem 4 of the Pre-Activity, 2×2 matrix multiplication is associative, and because Σ is a subset of 2×2 matrices, it is associative with respect to matrix multiplication.
- 1(b) corresponds with closure, and appears true by inspection.
- 1(c) corresponds with the existence of inverses, and appears true by inspection.
- 1(d) corresponds with the existence of an identity element, and is clearly true when the entries of $\mathrm{A}(2\pi)$ are calculated. Alternatively, since 2π is a full rotation, it makes sense that $\mathrm{A}(2\pi)$ would not change the position of any vector on which it acts.
- 1(e) corresponds with commutativity. If one accepts that Σ is closed, we can apply $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_1 + \theta_2)$ and the commutativity of real numbers to demonstrate the result.

$\Rightarrow$ Σ appears to be a commutative group under matrix multiplication.

Convene a classroom discussion to establish that Σ is in fact a group. Discuss the following connection to teaching:

Discuss This Connection to Teaching

The ability to "undo" something is key to many high school mathematics topics (e.g., inverse functions, solving equations by using additive/multiplicative inverses, etc.) By emphasizing this aspect of algebraic structure, prospective teachers can appreciate how working within a group (eventually, a field) guarantees that certain familiar operations are appropriate and valid.

We would like undergraduates to think about these conditions geometrically, but there are algebraic alternatives for advanced classes (or if time permits):

- *For closure, undergraduates will need to use the angle sum formulas for sine and cosine.*
- *For commutativity, undergraduates should refer to the fact that they have proved* $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_1 + \theta_2)$ *and appeal to the commutativity of* θ_1 *and* θ_2 *as real numbers.*
- *For inverses, undergraduates can use the fact that they have now shown both* $\mathrm{A}(\theta_2)\mathrm{A}(\theta_1) = \mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_1 + \theta_2)$ *and* $\mathrm{A}(2k\pi) = \mathrm{I}$ *to construct a reasonable algebraic argument.*

Class Activity: Problems 3 & 4 (20 minutes)

Allow the class to work on Problem 3 in small groups. Before they begin, you may wish to remind undergraduates that matrices in this set represent reflections about a line through the origin (and that line is determined by the angle θ).

Class Activity Problem 3 : Parts a, b, & c

3. Consider the set Φ given below.

$$\Phi = \left\{ \mathrm{B}(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

Answer the following questions by considering the geometric effect matrices in Φ have on vectors in $\mathbb{R}^2$. Do not make any calculations—your answers should be written sentences or labeled sketches.

(a) Let $\mathrm{B}(\theta_1)$ be a particular matrix in Φ. Do you think $\mathrm{B}(\theta_1)$ has a multiplicative inverse in Φ? That is, is there a value θ_2 for which $\mathrm{B}(\theta_2) = \mathrm{B}(\theta_1)^{-1}$? Explain why or why not.

Solution:
Yes, and $\theta_2 = \theta_1$. If you have reflected a vector over a line, to return it to its original position you can always reflect it over the same line; that is, each matrix in Φ is its own inverse.

(b) Does Φ contain the identity matrix? That is, is there a value θ for which $\mathrm{B}(\theta) = I$? Explain why or why not.

Solution:
Reflecting a vector about a line through the origin will almost always produce a different vector (the exception being when the vector coincides with the line of reflection). The identity matrix cannot be in Φ.

(c) Explain how your answers to Problems 3(a) and 3(b) can be used to determine that Φ is NOT closed under matrix multiplication.

Solution:
We see that $\mathrm{B}(\theta_1)\mathrm{B}(\theta_1) = \mathrm{I}$ by Problem 3(a), but I is not in Φ by Problem 3(b). Thus, there exists a pair of matrices in Φ whose product is not again in Φ.

Commentary:
As you circulate the classroom:

- For Problem 3(a), remind undergraduates to think geometrically: the inverse matrix of $\mathrm{B}(\theta_1)$ is the matrix in Φ which "undoes" the reflection given by $\mathrm{B}(\theta_1)$.
- For Problem 3(b), undergraduates may recognize that I is not in Φ by reasoning about the signs of the entries along the diagonal of $\mathrm{B}(\theta)$. If so, encourage them to also provide a geometric explanation of why no matrix in Φ could represent the identity transformation.
- For Problem 3(b), you might also prompt undergraduates to explain why it is not contradictory for it to be true that $\mathrm{B}(2\pi)(1,0) = (1,0)$ even though we know that $\mathrm{B}(2\pi) \neq \mathrm{I}$. This type of reasoning is important for helping undergraduates make sense of quantifiers in the group axioms.

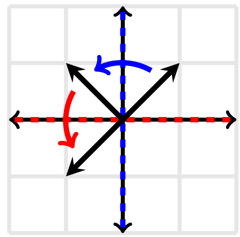
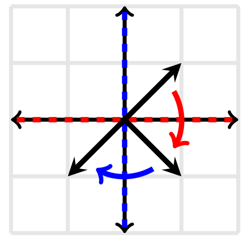
After most groups have completed Problems 3(a)–3(c), check in as a whole class to make sure everyone is on the same page. Discuss the following connection to teaching to set the stage before continuing on to Problem 3(d).

Discuss This Connection to Teaching

Problems such as Problem 3(d) are valuable because it's useful for everyone to think about how others use and reason with mathematics. It also gives prospective teachers or undergraduates who are considering work as a tutor or teaching assistant an opportunity to think about how they would respond to student work in ways that help students develop an understanding of the concept. Generating multiple questions for Luisa models that teachers often need to have several different questions prepared to facilitate different ways students might be thinking about the problem.

Class Activity Problem 3 : Part d

(d) Luisa sketches two diagrams, pictured below, to illustrate two different orders in which an arbitrary vector might be reflected over both the x and y axes.

Luisa then verifies arithmetically that $\mathrm{B}(\frac{\pi}{2})\mathrm{B}(\pi) = \mathrm{B}(\pi)\mathrm{B}(\frac{\pi}{2})$. Based on the diagrams and her supporting calculations, she concludes that elements of Φ commute under matrix multiplication.

i. Luisa's conclusion is incorrect. What understanding has Luisa demonstrated in her work? What has she proven?

Sample Responses:

- Luisa has demonstrated that she knows the matrices of the set Φ represent reflections over a line through the origin.
- Luisa understands matrix multiplication and how to graph vectors as they are transformed by a matrix.
- Luisa knows the definition of commutativity.
- Luisa has only proven only that the two matrices she has chosen, representing two specific reflections from the set Φ, commute with each other.

ii. Explain the error in Luisa's reasoning.

Sample Response:
Luisa has shown that a single pair of matrices in Φ commute, but because she did not choose two arbitrary matrices, she has not shown that ALL possible pairs of matrices in Φ commute.

iii. Give two questions you could ask Luisa to help her understand her error. Explain why your questions are helpful.

Sample Undergraduate Response(s):

- What does it mean for an entire set to be commutative under an operation? This question will help Luisa think about the quantifiers in the commutativity requirement for abelian groups.
- I want to prove that all horses are brown and bring you one brown horse. Have I convinced you? This question will help Luisa think about the quantifiers in the commutativity requirement for abelian groups.
- If we calculate the entries of $\mathrm{B}(\frac{\pi}{2})$ and $\mathrm{B}(\pi)$, what special property do we see these matrices have? Do all matrices in this set have this property? These questions will help Luisa think about the commutativity of diagonal and non-diagonal matrices, in particular whether reflections are always diagonal.

Give the class only a few minutes to consult their neighbors about Problem 4—their work on Problem 3 should have provided much of the insight for addressing this problem. As you monitor their work, select groups that will report out their responses. After an appropriate sequence of presentations from the groups, move on to the next part of the Class Activity.

Class Activity Problem 4

4. Explain whether Φ is a group under matrix multiplication by analyzing your responses to Problem 3. Does "order matter" when multiplying together a string of matrices from Φ?

   Solution:

   - Φ is not a group; by Problem 3(c) it is not closed, and by Problem 3(b) it does not contain an identity element.
   - Order does matter when applying reflections; by Problem 3(d), we see that reflections do not always commute.

Class Activity: Problems 5 & 6 (20 minutes)

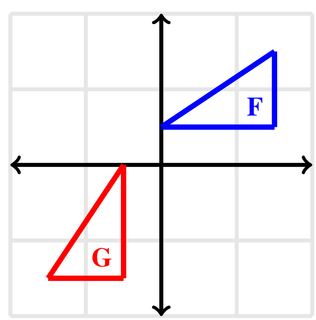
Allow undergraduates to work in small groups on Problem 5. As you monitor and facilitate their group work, consider the following prompts to encourage discussion. These prompts may also be posed for class discussion.

- How can we interpret this activity in terms of matrices acting on vectors in $\mathbb{R}^2$? Can you write Todd's steps using matrices from Σ and Φ?
- Can you find a single transformation, rather than a sequence, that will move triangle F back to its original position? What does this imply about the composition of a rotation and a reflection?

Class Activity Problem 5

5. Todd, a high school geometry student, is attempting to show that the two triangles pictured to the right are congruent. To do so, he must use some combination of reflections and rotations to move triangle F on top of triangle G. Todd concludes that he should:

   - Reflect F over the y-axis.
   - Rotate F counterclockwise $90°$ about the origin.



   To move F back to its original position, Todd says he can make these same two transformations in reverse order. That is, once F has been moved to the same position as G, he would:

   - Rotate F counterclockwise $90°$ about the origin.
   - Reflect F over the y-axis.

   (a) Why might Todd expect this procedure to work?

   Sample Responses:

   - Todd might not expect order to matter when applying transformations, especially if he has not yet learned about commutativity.
   - Todd might be thinking that if you reverse the order that you transform something, it will have the opposite effect.

   (b) Explain the error in Todd's reasoning.

   Sample Responses:

   - Todd's first step is okay, but next he needs to reflect over the x-axis instead of the y-axis.
   - Todd isn't thinking about the inverse of each individual transformation. He's just doing the same thing in the opposite order.

   (c) Find a sequence of transformations that will move F back to its original position. Explain, using vocabulary or notation from this course, how you know your steps are correct.

Solution(s):

- First, rotate F clockwise 90° about the origin. Then, reflect F over the y-axis. This will return F back to its original position because $[\mathrm{A}(\frac{\pi}{2})\mathrm{B}(\frac{\pi}{2})]^{-1} = \mathrm{B}(\frac{\pi}{2})^{-1}\mathrm{A}(\frac{\pi}{2})^{-1} = \mathrm{B}(\frac{\pi}{2})\mathrm{A}(\frac{-\pi}{2})$.
- You can also repeat the same sequence of transformations in the same order.

Problem 6 can be posed and answered as a class without any preceding group work.

Class Activity Problem 6

6. It turns out that the set of all rotations and reflections, $\Sigma \cup \Phi$, is itself a group under multiplication (you do not need to prove this). Does "order matter" when multiplying together a string of matrices from $\Sigma \cup \Phi$?

   Solution:
   Because $\Sigma \cup \Phi$ contains the reflections as a subset and we know that order matters for reflections, order also matters, in general, for $\Sigma \cup \Phi$.

Tie Problems 5 and 6 to the work of teaching by discussing the following connection:

Discuss This Connection to Teaching

Rigid motions are used in high school geometry to justify the congruence of geometric figures. Problems 5 and 6 require undergraduates to translate between high school geometry concepts and advanced mathematical language. This process emphasizes to prospective teachers that knowledge of matrices and underlying group structure can inform solutions in their future classrooms; in this case, a sequence of rigid motions can always be "undone" because of the underlying group structure, but that order does in fact matter when considering a sequence of rigid motions.

Wrap-Up (5 minutes)

Recap the lesson briefly for the class:

- We can represent particular rigid motions as sets of matrices and consider whether they are groups under matrix multiplication:
  - The set of rotations about the origin are an abelian group.
  - The set of reflections about lines through the origin are not a group, nor do they always commute.
  - The union of the two sets above is a group, but not an abelian one.
- Group structure can help indicate some situations in which "order doesn't matter": because rotations are an abelian group, we can compose those transformations however we like. We must be more careful when dealing with reflections.

We conclude the lesson by using an exit ticket. See Chapter 1 for guidance on how to conclude mathematics lessons using exit tickets.

Homework Problems

At the end of the lesson, assign the following homework problems.

Problem 1 prompts undergraduates to interpret another student's thinking in a way that allows them to generate a pair of guiding questions. In doing so, they practice interpreting the order of quantifiers in the identity group axiom; multiply quantified statements are a difficult topic for some undergraduates.

Homework Problem 1

1. Recall that Σ represents the set of rotations about the origin and that Φ represents the set of reflections across lines through the origin. These sets are given below:

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\} \quad \Phi = \left\{ \mathrm{B}(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

Reasoning geometrically, Jordan finds that $\mathrm{B}(\frac{\pi}{4})v = v$ for any vector v whose vertical and horizontal components are equal.

(a) What geometric understanding might Jordan have of the set Φ that enabled them to draw this conclusion without making any computations?

Sample Responses:

- Jordan recognized that Φ was the set of reflections about lines through the origin.
- Jordan knows that the line of reflection used by $\mathrm{B}(\frac{\pi}{4})$ to transform a vector lies $\frac{\pi}{4}$ radians above the positive x-axis.
- Jordan knows that the vector v must form an angle of either $\frac{\pi}{4}$ or $\frac{5\pi}{4}$ radians with the positive x-axis. In either case, v will coincide with the line over which $\mathrm{B}(\frac{\pi}{4})$ reflects and v will not be transformed.

(b) Jordan claims that the work above shows that $\mathrm{B}(\frac{\pi}{4})$ is the identity matrix. Explain the error in Jordan's reasoning.

Solution:
Jordan does not seem to understand that the identity matrix I must satisfy $\mathrm{I}v = v$ for ALL vectors $v \in \mathbb{R}^2$. This matrix only functions as an identity for vectors of the form (a, a).

(c) Give two questions you could ask Jordan to help them understand their error. Why would your questions be helpful?

Sample Responses:

- Does $\mathrm{B}(\frac{\pi}{4})$ act as the identity on $w = (1, 0)$? This question would be helpful because Jordan might not realize $\mathrm{B}(\frac{\pi}{4})$ is only the identity matrix if $\mathrm{B}(\frac{\pi}{4})v = v$ for ALL $v \in \mathbb{R}^2$.
- What are the entries of $\mathrm{B}(\frac{\pi}{4})$? This question would be helpful because it will show Jordan that the diagonal entries of any matrix $\mathrm{B}(\theta)$ cannot both be 1.

Undergraduates should understand that binary operations need not be "nice," i.e., only consisting of some combination of the usual, familiar operations. Problem 2 introduces undergraduates to an unconventional binary operation that is in fact well-behaved: order doesn't matter when applying $\diamond$.

Homework Problem 2

2. Consider the operation $\diamond$ given by $a \diamond b = a^{\log(b)}$ on the set of positive real numbers, $\mathbb{R}^+$.

(a) Is $\diamond$ closed on this set? If so, justify your conclusion. If not, provide a specific example of $a, b \in \mathbb{R}^+$ for which $a \diamond b \notin \mathbb{R}^+$.

Solution:
The operation $\diamond$ is closed since, $\forall\ a, b \in \mathbb{R}^+$, we have $\log(b) \in \mathbb{R}$, from which it follows that $a^{\log(b)} \in \mathbb{R}^+$.

(b) Is $\diamond$ an associative operation on this set? If so, justify your conclusion. If not, provide a specific example of $a, b, c \in \mathbb{R}^+$ for which $a \diamond (b \diamond c) \neq (a \diamond b) \diamond c$.

Solution:
The operation $\diamond$ is associative. Below, we apply the fact that $a \diamond b = a^{\log(b)} = 10^{\log(a)\log(b)}$:
$a \diamond (b \diamond c) = a^{\log[10^{\log(b)\log(c)}]} = a^{\log(b)\log(c)} = 10^{\log(a)\log(b)\log(c)}$
$(a \diamond b) \diamond c = [10^{\log(a)\log(b)}]^{\log(c)} = 10^{\log(a)\log(b)\log(c)}$

(c) Is $\diamond$ a commutative operation on this set? If so, justify your conclusion. If not, provide a specific example of $a, b \in \mathbb{R}^+$ for which $a \diamond b \neq b \diamond a$.

Solution:
The operation is $\diamond$ commutative: $a \diamond b = a^{\log(b)} = [10^{\log(a)}]^{\log(b)} = [10^{\log(b)}]^{\log(a)} = b^{\log(a)} = b \diamond a$

(d) Does "order matter" under this operation? Explain why or why not.

Solution:
No. The operation $\diamond$ is clearly closed, and we have demonstrated that it is also both commutative and associative.

The impression that commutativity is "stronger" than associativity sometimes leads to the misconception that the former implies the latter. Averaging is a binary operation (familiar even to high school students such as Aisling) that is commutative but non-associative, providing an interesting example of when order does matter that is not likely already a part of undergraduates' concept image. Problem 3 gives undergraduates the opportunity to leverage their understanding of binary operations and their properties in a teaching application.

Homework Problem 3

3. Aisling, a high school student, has made an 84 and a 72 on her first two precalculus assignments. She calculates her average in the course to be a 78. The following day, she receives a 90 on her next assignment. She makes the following calculation to compute her new average:

$$\tfrac{1}{2}(78 + 90) = 84$$

(a) What error has Aisling made?

Solution:
Averaging is not an associative operation; instead of averaging her grades one at a time, she should take the average all at once by adding and dividing by 3.

(b) Show that the operation $*$, given by $a * b = \frac{1}{2}(a + b)$ where $a, b \in \mathbb{R}^+$, is commutative. Does "order matter" under this operation? Explain why or why not.

Solution:
The operation is $*$ commutative: $a * b = \frac{1}{2}(a + b) = \frac{1}{2}(b + a) = b * a$. However, "order matters" because $*$ is not associative.

(c) Consider the following questions that you might ask Aisling:

i. Explain why the question below might not help Aisling:

Should your average be lower than 84?

Sample Response:
Aisling doesn't know what her average should be; that is why she was calculating it. A teacher might be able to average three grades in their head pretty easily, but a student probably can't. This doesn't help Aisling see what she did wrong, it just makes her think her answer isn't right.

ii. Explain how the question below might help you advance Aisling's understanding:

What would your average be if you had made a 90, then a 72, then an 84?

Sample Response:

By asking Aisling to recalculate her average in a different order, she will probably get a different answer than her first calculation. Hopefully, she will see that "order matters" when computing an average—that is, you can't just average a bunch of numbers one after another and need to find a way to do it all at once.

Problem 4 formally establishes that we do not need to check that an element is both a left and a right inverse (as long as the operation in question is associative), and thus that a definition for a group that at first glance appears "weaker" is the same.

Homework Problem 4

4. Let G be a set with associative operation $*$ and with identity element e. Assume that every element of G has a left inverse: that is, $\forall\, a \in G, \exists\, b \in G$ such that $b * a = e$.

 (a) Show that b must also be a right inverse of a: that is, we also have $a * b = e$.

 Solution:

 Since $b \in G$, it must also be true that b has a left inverse c. That is, $\exists\, c \in G$ such that $c * b = e$. Then: $b * a = e \Rightarrow (b * a) * b = b \Rightarrow c * ((b * a) * b) = c * b \Rightarrow e * (a * b) = e \Rightarrow a * b = e$, as desired.

 (b) Explain how the associativity of $*$ plays a key role in your proof for 4(a).

 Solution:

 In the above proof, we assert that $(b * a) * b = b \Rightarrow c * ((b * a) * b) = c * b \Rightarrow e * (a * b) = e$. This is only true because associativity allows us to rewrite the expression $c * ((b * a) * b)$ as $(c * b) * (a * b)$, allowing us to apply the fact that c is the left inverse of b.

 (c) Examine a list of axioms that you've seen presented in the definition of a group. How does your work in this problem affect your understanding of these axioms?

 Sample Response:

 The group axiom which states that $\forall\, g \in G, \exists\, g^{-1} \in G$ such that $g * g^{-1} = g^{-1} * g = e$ is stronger than necessary. We need only verify that every element has a left inverse, that is, $\forall\, g \in G$, $\exists\, g^{-1} \in G$ such that $g^{-1} * g = e$. The fact that this left inverse is also the right inverse follows directly. Alternatively, the same proof as above slightly modified suffices to show that a right inverses are also left inverse. Essentially, group elements always commute with their inverse element. All of the above commentary requires that the set in question is already associative, however.

Translations are an important category of rigid motion that undergraduates may have used in high school geometry. Undergraduates might assume that translations can also be represented as matrices after seeing the such a treatment of the other rigid motions (reflections and rotations) in this lesson. In Problem 5 we justify that this is not true.

Homework Problem 5

5. We have encountered matrices which represent rotations and reflections of vectors in $\mathbb{R}^2$. Does there exist a 2×2 matrix which represents the translation of vectors? If so, write it down and justify how you know it represents translations. If not, explain.

 Solution:

 No. Matrices represent linear transformations, so if there existed a matrix A that was a translation, we would require $A\vec{0} = \vec{0}$. However, if A is any non-identity translation, this would not be the case.

Undergraduates are first given the opportunity to prove an important result algebraically that adds depth to results observed during the Class Activity. Problem 6 continues the trend of framing meaningful results about group structure using geometric language.

Homework Problem 6

6. Show that the product of any two reflection matrices is a rotation matrix. [Hint: You will need the angle subtraction formulas for sine and cosine]. Using this result, give a geometric description of when two reflection matrices will commute.

Solution:

$$\begin{aligned} B(\theta_1)B(\theta_2) &= \begin{bmatrix} \cos(2\theta_1) & \sin(2\theta_1) \\ \sin(2\theta_1) & -\cos(2\theta_1) \end{bmatrix} \begin{bmatrix} \cos(2\theta_2) & \sin(2\theta_2) \\ \sin(2\theta_2) & -\cos(2\theta_2) \end{bmatrix} \\ &= \begin{bmatrix} \cos(2\theta_1)\cos(2\theta_2) + \sin(2\theta_1)\sin(2\theta_2) & \sin(2\theta_2)\cos(2\theta_1) - \sin(2\theta_1)\cos(2\theta_2) \\ \sin(2\theta_1)\cos(2\theta_2) - \sin(2\theta_2)\cos(2\theta_1) & \cos(2\theta_1)\cos(2\theta_2) + \sin(2\theta_1)\sin(2\theta_2) \end{bmatrix} \\ &= \begin{bmatrix} \cos(2(\theta_1-\theta_2)) & \sin(2(\theta_2-\theta_1)) \\ \sin(2(\theta_1-\theta_2)) & \cos(2(\theta_1-\theta_2)) \end{bmatrix} \\ &= \begin{bmatrix} \cos(2(\theta_1-\theta_2)) & -\sin(2(\theta_1-\theta_2)) \\ \sin(2(\theta_1-\theta_2)) & \cos(2(\theta_1-\theta_2)) \end{bmatrix} \\ &= A(2(\theta_1-\theta_2)) \end{aligned}$$

Where we use the fact that sine is an odd function when moving from the third to the fourth line. By this result, $B(\theta_1)B(\theta_2) = B(\theta_2)B(\theta_1)$ when $A(2(\theta_1-\theta_2)) = A(2(\theta_2-\theta_1))$, or equivalently when $2(\theta_1-\theta_2) = 2(\theta_2-\theta_1) + 2k\pi$ for some $k \in \mathbb{Z}$. Simplifying, we see that two reflection matrices commute only when $\theta_1 = \theta_2 + k\frac{\pi}{2}$. That is, two reflection matrices will commute when the lines over which they reflect are either orthogonal or identical.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Assessment Problems 1 & 2

1. Consider the set Δ, given below.

$$\Delta = \left\{ \mathrm{C}(a) = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a \in \mathbb{R}^+ \right\}$$

(a) Describe the geometric effect that the matrix $\mathrm{C}(a)$ has on a vector $v \in \mathbb{R}^2$.

Solution:

The set Δ represents the dilations. That is, the matrix $\mathrm{C}(a)$ stretches or compresses a vector by a factor of a, depending on whether a is greater or less than 1, respectively.

(b) Is Δ an abelian group under matrix multiplication? Demonstrate why or why not.

Solution:

- The set is closed under multiplication: $\mathrm{C}(a)\mathrm{C}(b) = \mathrm{C}(ab)$, and $ab \in \mathbb{R}^+$ since both a and b are positive real numbers.
- The set of all 2×2 matrices is associative, so Δ, as a subset, must also be associative.
- It is clear that $\mathrm{C}(1) = \mathrm{I}$, so there is a group identity.
- If $a \in \mathbb{R}^+$, then $\frac{1}{a} \in \mathbb{R}^+$. Also: $\mathrm{C}(a)\mathrm{C}(\frac{1}{a}) = \mathrm{C}(\frac{a}{a}) = \mathrm{C}(1) = \mathrm{I}$, so each element has a group inverse.
- Finally, $\mathrm{C}(a)\mathrm{C}(b) = \mathrm{C}(ab) = \mathrm{C}(ba) = \mathrm{C}(b)\mathrm{C}(a)$, so Δ is abelian.

Thus, Δ is an abelian group, as desired.

2. Serena is working with the set Σ of rotations about the origin, given below.

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

She knows Σ is a group under multiplication. Reasoning geometrically, Serena argues that the matrices $A(-\frac{\pi}{4})$ and $A(\frac{7\pi}{4})$ both act as inverses of matrix $A(\frac{\pi}{4})$.

(a) What geometric understanding might Serena have of rotations about the origin that enabled her to draw this conclusion without making any computations?

Sample Responses:

- Serena knows that positive values of θ represent counterclockwise rotations and negative values of θ represent clockwise rotations.
- Serena recognizes that the inverse of $A(\frac{\pi}{4})$ could be conceptualized as either the matrix that rotates the opposite direction with equal magnitude, $A(-\frac{\pi}{4})$, or the matrix which rotates the rest of the way around the circle in the same direction, i.e. $A(2\pi - \frac{\pi}{4}) = A(\frac{7\pi}{4})$. Both of these matrices would return a vector to its original position after being transformed by $A(\frac{\pi}{4})$.

(b) Serena claims that her work above shows that the group of rotations is a counterexample to the claim that all group elements have a unique inverse. Explain the error in Serena's reasoning.

Sample Response:

She assumes that having differing θ values leads to unique matrices in A when really it's based on the evaluated $\sin(\theta)$ and $\cos(\theta)$ values.

(c) What question would you ask Serena to help her understand her error? Why is your question helpful?

Sample Responses:

- What are the matrix representations of $A(-\frac{\pi}{4})$ and $A(\frac{7\pi}{4})$? This question would be helpful to show that both angles would produce the same matrix.
- Where on the unit circle can you find $-\frac{\pi}{4}$ and $\frac{7\pi}{4}$? This question would be helpful to show that they are coterminal angles, so they have the same values for sine and cosine.
- Are the elements of the group real numbers or matrices? This question would help Serena realize that the parameters might not be the same, but that it would not matter if the resultant matrices are identical.

9.6 References

[1] Ada, T., & Kurtuluş, A. (2010). Students' misconceptions and errors in transformation geometry. *International Journal of Mathematical Education in Science and Technology*, *41*(7), 901-909.

[2] Lai, Y., & Donsig, A. (2018). Using geometric habits of mind to connect geometry from a transformation perspective to graph transformations and abstract algebra. *Connecting abstract algebra to secondary mathematics, for secondary mathematics teachers*, 263–289.

[3] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from `http://www.corestandards.org/`

[4] Seago, N., Jacobs, J., Driscoll, M., Nikula, J., Matassa, M., & Callahan, P. (2013). Developing teachers' knowledge of a transformations-based approach to geometric similarity. *Mathematics Teacher Educator*, *2*(1), 74-85.

9.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** GROUPS OF TRANSFORMATIONS (page 1 of 4)

Consider the sum $2 + 6 + 4 + 7$. When people say that "order does not matter" when computing such a sum they actually mean two things: the order of the individual terms of the sum can be rearranged without affecting the final result (for instance, $7 + 4 + 6 + 2$ and the original sum are sure to each give the same answer, 19) and, moreover, the order in which one chooses to compute the individual addition operations is unimportant (for instance, $((2+6)+4)+7$ and $2 + ((6 + 4) + 7)$ both yield the same final result of 19). This conclusion relies on the three fundamental beliefs of integer arithmetic:

- Integer addition is **closed**; that is, $a + b$ is itself an integer for all integers a and b.
- Integer addition is **commutative**; that is, $a + b = b + a$ for all integers a and b.
- Integer addition is **associative**; that is, $(a + b) + c = a + (b + c)$ for all integers a, b, and c.

1. Consider the set of all rotations about the origin of the plane.
   [Recall that transformations (e.g., rotations) are functions. As such, for rotations r_α and r_β on $\mathbb{R}^2$, the *composition* of r_α followed by r_β, $r_\beta \circ r_\alpha$ is defined by $r_\beta \circ r_\alpha(P) = r_\beta(r_\alpha(P))$ where $P \in \mathbb{R}^2$.]

   (a) Is this set closed under composition? Explain.

   (b) Do rotations commute with each other under composition? Explain.

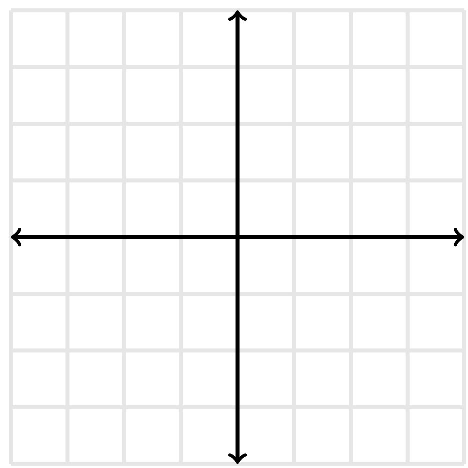
   (c) Do rotations about the origin satisfy the associative law under composition? Explain.

   (d) Does "order matter" when performing a series of rotations about the origin in the plane? Explain.

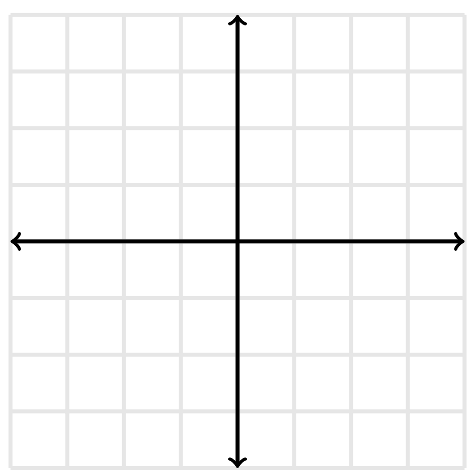
2. Consider the set Σ, given below.

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

(a) Calculate $A(\frac{\pi}{2})$. Choose three nonzero vectors v_1, v_2, and v_3 in $\mathbb{R}^2$ that are not all scalar multiples of one another. Compute $A(\frac{\pi}{2})v_1$, $A(\frac{\pi}{2})v_2$, and $A(\frac{\pi}{2})v_3$. Sketch all six vectors on the same coordinate plane.



(b) Repeat the process of 2(a) with $A(\theta)$ for a different nonzero value of θ and the same vectors.



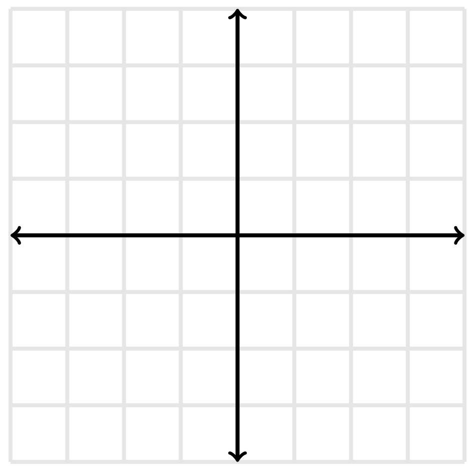
(c) Write a geometric description of how an arbitrary matrix from Σ acts on vectors in $\mathbb{R}^2$ based on your sketches in Problems 2(a) and 2(b).

PRE-ACTIVITY: GROUPS OF TRANSFORMATIONS (page 3 of 4)

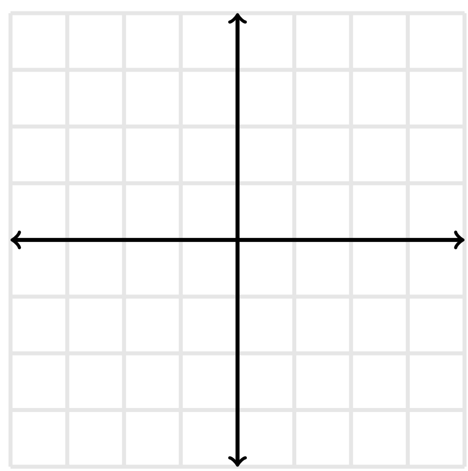
3. Consider the set Φ, given below.

$$\Phi = \left\{ \mathrm{B}(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

(a) Repeat the process of Problem 2(a) with matrix $B(\frac{\pi}{2})$.



(b) Repeat the process of Problem 3(a) with $B(\theta)$ for a different value of θ and the same vectors.



(c) Write a geometric description of how an arbitrary matrix from Φ acts on vectors in $\mathbb{R}^2$ based on your sketches in Problems 3(a) and 3(b).

4. Recall that 2×2 matrix multiplication is associative. On the other hand, 2×2 matrices do not always commute under matrix multiplication. Give an example of a pair of 2×2 matrices that do commute and a pair of 2×2 matrices that do NOT commute under matrix multiplication.

NAME: **CLASS ACTIVITY:** GROUPS OF TRANSFORMATIONS (page 1 of 4)

1. Consider the set Σ given below.

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

Restate the following equations in terms of the geometric effect matrices in Σ have on vectors in $\mathbb{R}^2$. Do not make any calculations—your answers should be written sentences or labeled sketches. A sample solution to part (a) and a partial solution to part (b) are provided.

(a) $\mathrm{A}(\theta_1)[\mathrm{A}(\theta_2)\mathrm{A}(\theta_3)] = [\mathrm{A}(\theta_1)\mathrm{A}(\theta_2)]\mathrm{A}(\theta_3)$

When rotating a vector by three different angles, the way that I group the rotations before doing them does not affect the overall rotation.

(b) $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_1 + \theta_2)$

Rotating a vector by _____ radians and then by _____ radians is equivalent to rotating it by ____________ radians.

(c) $\mathrm{A}(\theta)^{-1} = \mathrm{A}(-\theta)$

(d) $\mathrm{A}(2\pi) = \mathrm{I}$

(e) $\mathrm{A}(\theta_1)\mathrm{A}(\theta_2) = \mathrm{A}(\theta_2)\mathrm{A}(\theta_1)$

2. Explain whether Σ is an abelian group under matrix multiplication by analyzing your responses to Problem 1. Does "order matter" when multiplying together a string of matrices from Σ?

3. Consider the set Φ given below.

$$\Phi = \left\{ \mathrm{B}(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

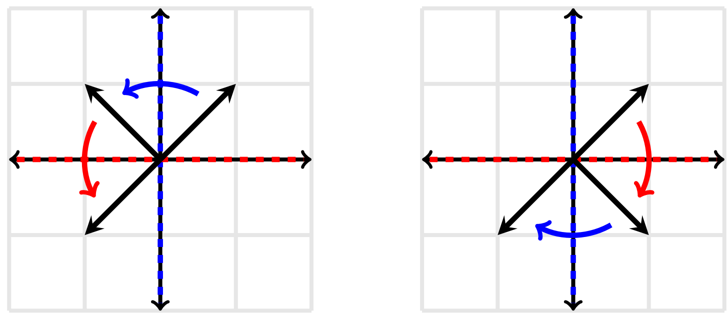
Answer the following questions by considering the geometric effect matrices in Φ have on vectors in $\mathbb{R}^2$. Do not make any calculations—your answers should be written sentences or labeled sketches.

(a) Let $\mathrm{B}(\theta_1)$ be a particular matrix in Φ. Do you think $\mathrm{B}(\theta_1)$ has a multiplicative inverse in Φ? That is, is there a value θ_2 for which $\mathrm{B}(\theta_2) = \mathrm{B}(\theta_1)^{-1}$? Explain why or why not.

(b) Does Φ contain the identity matrix? That is, is there a value θ for which $\mathrm{B}(\theta) = I$? Explain why or why not.

(c) Explain how your answers to Problems 3(a) and 3(b) can be used to determine that Φ is NOT closed under matrix multiplication.

(d) Luisa sketches two diagrams, pictured below, to illustrate two different orders in which an arbitrary vector might be reflected over both the x and y axes.



Luisa then verifies arithmetically that $\mathrm{B}(\frac{\pi}{2})\mathrm{B}(\pi) = \mathrm{B}(\pi)\mathrm{B}(\frac{\pi}{2})$. Based on the diagrams and her supporting calculations, she concludes that elements of Φ commute under matrix multiplication.

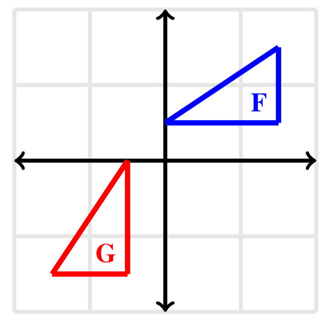
i. Luisa's conclusion is incorrect. What understanding has Luisa demonstrated in her work? What has she proven?

ii. Explain the error in Luisa's reasoning.

iii. Give two questions you could ask Luisa to help her understand her error. Explain why your questions are helpful.

4. Explain whether Φ is a group under matrix multiplication by analyzing your responses to Problem 3. Does "order matter" when multiplying together a string of matrices from Φ?

5. Todd, a high school geometry student, is attempting to show that the two triangles pictured to the right are congruent. To do so, he must use some combination of reflections and rotations to move triangle F on top of triangle G. Todd concludes that he should:



   - Reflect F over the y-axis.
   - Rotate F counterclockwise $90°$ about the origin.

   To move F back to its original position, Todd says he can make these same two transformations in reverse order. That is, once F has been moved to the same position as G, he would:

   - Rotate F counterclockwise $90°$ about the origin.
   - Reflect F over the y-axis.

   (a) Why might Todd expect this procedure to work?

   (b) Explain the error in Todd's reasoning.

   (c) Find a sequence of transformations that will move F back to its original position. Explain, using vocabulary or notation from this course, how you know your steps are correct.

6. It turns out that the set of all rotations and reflections, $\Sigma \cup \Phi$, is itself a group under multiplication (you do not need to prove this). Does "order matter" when multiplying together a string of matrices from $\Sigma \cup \Phi$?

NAME: **HOMEWORK PROBLEMS:** GROUPS OF TRANSFORMATIONS (page 1 of 2)

1. Recall that Σ represents the set of rotations about the origin and that Φ represents the set of reflections across lines through the origin. These sets are given below:

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\} \qquad \Phi = \left\{ \mathrm{B}(\theta) = \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

Reasoning geometrically, Jordan finds that $\mathrm{B}(\frac{\pi}{4})v = v$ for any vector v whose vertical and horizontal components are equal.

(a) What geometric understanding might Jordan have of the set Φ that enabled them to draw this conclusion without making any computations?

(b) Jordan claims that the work above shows that $\mathrm{B}(\frac{\pi}{4})$ is the identity matrix. Explain the error in Jordan's reasoning.

(c) Give two questions you could ask Jordan to help them understand their error. Why would your questions be helpful?

2. Consider the operation $\diamond$ given by $a \diamond b = a^{\log(b)}$ on the set of positive real numbers, $\mathbb{R}^+$.

(a) Is $\diamond$ closed on this set? If so, justify your conclusion. If not, provide a specific example of $a, b \in \mathbb{R}^+$ for which $a \diamond b \notin \mathbb{R}^+$.

(b) Is $\diamond$ an associative operation on this set? If so, justify your conclusion. If not, provide a specific example of $a, b, c \in \mathbb{R}^+$ for which $a \diamond (b \diamond c) \neq (a \diamond b) \diamond c$.

(c) Is $\diamond$ a commutative operation on this set? If so, justify your conclusion. If not, provide a specific example of $a, b \in \mathbb{R}^+$ for which $a \diamond b \neq b \diamond a$.

(d) Does "order matter" under this operation? Explain why or why not.

3. Aisling, a high school student, has made an 84 and a 72 on her first two precalculus assignments. She calculates her average in the course to be a 78. The following day, she receives a 90 on her next assignment. She makes the following calculation to compute her new average:

$$\tfrac{1}{2}(78 + 90) = 84$$

(a) What error has Aisling made?

(b) Show that the operation $*$, given by $a * b = \frac{1}{2}(a + b)$ where $a, b \in \mathbb{R}^+$, is commutative. Does "order matter" under this operation? Explain why or why not.

(c) Consider the following questions that you might ask Aisling:

i. Explain why the question below might not help Aisling:

Should your average be lower than 84?

ii. Explain how the question below might help you advance Aisling's understanding:

What would your average be if you had made a 90, then a 72, then an 84?

4. Let G be a set with associative operation $*$ and with identity element e. Assume that every element of G has a left inverse: that is, $\forall\, a \in G, \exists\, b \in G$ such that $b * a = e$.

(a) Show that b must also be a right inverse of a: that is, we also have $a * b = e$.

(b) Explain how the associativity of $*$ plays a key role in your proof for 4(a).

(c) Examine a list of axioms that you've seen presented in the definition of a group. How does your work in this problem affect your understanding of these axioms?

5. We have encountered matrices which represent rotations and reflections of vectors in $\mathbb{R}^2$. Does there exist a 2×2 matrix which represents the translation of vectors? If so, write it down and justify how you know it represents translations. If not, explain.

6. Show that the product of any two reflection matrices is a rotation matrix. [Hint: You will need the angle subtraction formulas for sine and cosine]. Using this result, give a geometric description of when two reflection matrices will commute.

NAME: **ASSESSMENT PROBLEMS:** GROUPS OF TRANSFORMATIONS (page 1 of 2)

1. Consider the set Δ, given below.

$$\Delta = \left\{ \mathrm{C}(a) = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} : a \in \mathbb{R}^+ \right\}$$

(a) Describe the geometric effect that the matrix $\mathrm{C}(a)$ has on a vector $v \in \mathbb{R}^2$.

(b) Is Δ an abelian group under matrix multiplication? Demonstrate why or why not.

ASSESSMENT PROBLEMS: GROUPS OF TRANSFORMATIONS (page 2 of 2)

2. Serena is working with the set Σ of rotations about the origin, given below.

$$\Sigma = \left\{ \mathrm{A}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} : \theta \in \mathbb{R} \right\}$$

She knows Σ is a group under multiplication. Reasoning geometrically, Serena argues that the matrices $A(-\frac{\pi}{4})$ and $A(\frac{7\pi}{4})$ both act as inverses of matrix $A(\frac{\pi}{4})$.

(a) What geometric understanding might Serena have of rotations about the origin that enabled her to draw this conclusion without making any computations?

(b) Serena claims that her work above shows that the group of rotations is a counterexample to the claim that all group elements have a unique inverse. Explain the error in Serena's reasoning.

(c) What question would you ask Serena to help her understand her error? Why is your question helpful?

10

Logarithms and Isomorphisms

Abstract (Modern) Algebra I

Andrew Kercher, *Simon Fraser University*
James A. M. Álvarez, *The University of Texas at Arlington*

10.1 Overview and Outline of Lesson

The defining characteristic of the logarithm, as taught in high school, is that it is the inverse of the exponential function and it is used as a tool for solving exponential equations. Often overlooked is that the logarithm was created as an isomorphism between the set of positive real numbers under multiplication and the set of all real numbers under addition. In this lesson, undergraduates learn that the logarithm was originally devised as a computational aid and that its use in this manner is possible because it is both a group homomorphism and a group isomorphism. This lesson highlights this distinction by contrasting the logarithm with the determinant function on invertible matrices, which fails to act as a computational aid because it is a group homomorphism but not a group isomorphism. In this lesson, undergraduates also explore the structure-preserving nature of isomorphisms.

1. Launch—Pre-Activity:
   Undergraduates complete this assignment prior to the lesson. The Pre-Activity addresses the historical genesis of the logarithm and other familiar functions that are not homomorphisms. Undergraduates investigate how to use a logarithm table to approximate different values of the function $\log(x)$.

2. Explore—Class Activity:

   - *Problems 1–4*
     After using a larger logarithm table to calculate more values of $\log(x)$, undergraduates investigate the determinant function on 2×2 invertible matrices. They find that, despite this function being a homomorphism, it cannot be used in the same way as $\log(x)$ to facilitate computation. To come to this conclusion, undergraduates analyze the work of a hypothetical classmate who is attempting to use the determinant function to compute matrix products; they then consider how they might help this student understand why his process will not work reliably. Ultimately, undergraduates establish that the bijectivity of $\log(x)$ is a key aspect of its ability to function as a computational aid.
   - *Discussion: Homomorphisms and Isomorphisms*
     The instructor motivates the formal definitions of both homomorphisms and isomorphisms and facilitates discussion to establish that isomorphic groups have the same algebraic structure. The instructor provides examples, such as the groups must have the same cardinality; if one group is commutative (resp. cyclic) so is the other; if one group has an element of order a, so must the other; etc.

- *Problems 5–8*
  Undergraduates first identify the two groups between which the logarithm is an isomorphism. Then, they explain (with proof) why no isomorphism can exist between a series of groups by appealing to differences in algebraic structure.

3. Closure—Wrap-Up:
   The instructor concludes the lesson by reiterating that homomorphisms "preserve the operation" of a group and that two groups with an isomorphism between them have matching algebraic structures.

10.2 Alignment with College Curriculum

Isomorphisms and homomorphisms play a crucial role in the study of algebraic structures. Undergraduates explore these ideas by attempting to make computations by using the homomorphism property of the logarithm and determinant functions. The lesson introduces the formal mathematical definitions of isomorphism and homomorphism, also using the logarithm and determinant functions as preliminary examples. The lesson addresses that the existence of an isomorphism between two groups ensures that their algebraic structures will fully match (e.g., identical orders for corresponding elements, commutativity between corresponding elements is preserved, etc.).

10.3 Links to School Mathematics

The topic of logarithms is often presented in high school mathematics as a tool for solving exponential equations. While this is certainly a useful application of the function, a more complete treatment addresses that the logarithm predates formal exponential notation and was originally devised to help simplify tedious multiplication in the field of astronomy. The homomorphism property of logarithms is often treated as a fact to be memorized by secondary school students, but understanding the historical circumstances that gave rise to the logarithm helps ground this property in a meaningful context. The determinant, often computed in secondary school, is examined as a function on the set of invertible matrices and is discovered to be a group homomorphism.

This lesson highlights:

- The value of understanding and deriving results that may have been taken for granted, such as certain properties of logarithms.
- The role of knowing whether a function is injective, surjective, or bijective when comparing different functions.

This lesson addresses several mathematical knowledge and practice expectations included in high school standards documents, such as the Common Core State Standards for Mathematics (CCSSM, 2010). For example, high school students are expected to know how to graph and evaluate logarithmic functions using both analytical techniques and technology (c.f. CCSS.MATH.CONTENT.HSF.IF.C.7.e and CCSS.MATH.CONTENT.HSF.LE.A.4). To provide a point of comparison for the logarithm function this lesson also makes use of a determinant function on 2×2 matrices; high school students must also know of the matrix determinant (c.f. CCSS.MATH.CONTENT.HSM.VM.C.10) but also the basic operations of matrix arithmetic (c.f. CCSS.MATH.CONTENT.HSM.VM.C). Finally, this lesson emphasizes the need for viable mathematical arguments and encourages undergraduates to look for and make use of structural similarities.

10.4 Lesson Preparation

Prerequisite Knowledge

Undergraduates should know:

- The definition of a group. Some or all of the following related vocabulary is recommended:
  - Abelian group;
  - Cyclic group;
  - Order of a group (resp. an element);
- Basic familiarity with logarithms and determinants of 2×2 matrices;
- The definition of a bijection and how to tell if a function is bijective.

Learning Objectives

In this lesson, undergraduates will encounter ideas about teaching mathematics, as described in Chapter 1 (see the five types of connections to teaching listed in Table 1.2). In particular, by the end of the lesson undergraduates will be able to:

- Define group homomorphism and group isomorphism and identify key characteristics of each;
- Provide examples of group homomorphisms and group isomorphisms, as well as identify when a group homomorphism is or is not a group isomorphism;
- Use algebraic structure to explain when two groups cannot be isomorphic;
- Analyze hypothetical student work and assess understandings of group homomorphisms and group isomorphisms;
- Pose guiding questions to help a hypothetical student investigate function properties that preserve algebraic structure.

Anticipated Length

One 75-minute class session.

Materials

The following materials are required for this lesson:

- Pre-Activity (assign as homework prior to Class Activity)
- Class Activity
- Homework Problems (assign at the end of the lesson)
- Assessment Problems (include on quiz or exam after the lesson)

All handouts for this lesson appear at the end of this lesson, and LaTeX files can be downloaded from maa.org/meta-math.

10.5 Instructor Notes and Lesson Annotations

Before the Lesson

Assign the Pre-Activity as homework to be completed in preparation for this lesson.

> *We recommend that you collect this Pre-Activity the day before the lesson so that you can review undergraduates' responses before you begin the Class Activity. This will help you determine if you need to spend additional time reviewing the solutions to the Pre-Activity with your undergraduates.*

Pre-Activity Review (10 minutes)

Introduce the Pre-Activity by discussing the following connection to teaching:

Discuss This Connection to Teaching

High school students often only learn about the logarithm function as the inverse of the exponential function. In the course of solving these problems, they often memorize the property $\log(x \cdot y) = \log(x) + \log(y)$ without a meaningful understanding of the significance of this property. Understanding that this property is why logarithms were invented helps justify its existence and lessens the burden of memorization by giving meaningful context.

Next, discuss the solutions to the Pre-Activity as needed: if you saw that most undergraduates completed each problem correctly, you do not need to spend much time reviewing the solutions. However, if there are common areas that need further attention, you can give undergraduates time to deepen their understanding before you start the Class Activity. To accomplish this, we recommend first allowing your class to compare their answers in small groups. Then, call them together to share their findings in a whole-class discussion. See Chapter 1 for guidance on facilitating group work and selecting and sequencing student work for use in whole-class discussion.

In this lesson, we make no distinction between logarithms of different bases. You may ask your class to assume that we are using the common log (i.e., $\log_{10}(x)$) throughout the lesson; if so, you might prompt them to consider whether each of their conclusions remains valid in other bases.

Pre-Activity Problem 1

At the beginning of the 17^{th} century, two different mathematicians were working to simplify tedious calculations in the field of astronomy (where numbers are often very large) by creating a function that would "translate" multiplication into addition. After all, it is much easier to add numbers like 177,320,045 and 9,317,032,566 by hand than it is to multiply them. This work by John Napier from Scotland and Joost Bürgi from Switzerland led to the creation of the function that we now call the logarithm. The precise nature of the logarithm has been refined over the years by other mathematicians, but its original use as a computational aid is exemplified by a familiar identity: for all positive real numbers x and y,

$$\log(x \cdot y) = \log(x) + \log(y)$$

It is useful to think of this property as "preserving an operation"; that is, the logarithm function preserves multiplication by translating it into addition.

1. Of course, most functions on the real numbers do not translate a multiplicative structure into an additive one.

   (a) Find a pair of positive real numbers x and y that demonstrates that $\sin(x \cdot y)$ need not equal $\sin(x) + \sin(y)$. Find a pair of non-zero values for which, by coincidence, this relationship does hold.

   Sample Response:
   Most pairs of positive real numbers x and y will not have this property. For example, for $x = \pi$ and $y = 2$, $\sin(x \cdot y) = \sin(2\pi) = 0$. But $\sin(x) + \sin(y) = \sin(\pi) + \sin(2) = \sin(2)$, which is clearly not equal to 0.
   Some pairs (x, y) that do work:

   - For $(x, y) = (\sqrt{\pi}, -\sqrt{\pi})$, we first have $\sin(x \cdot y) = \sin(-\pi) = 0$. Then, we apply the fact that the sine function is odd: $\sin(x)+\sin(y) = \sin(\sqrt{\pi})+\sin(-\sqrt{\pi}) = \sin(\sqrt{\pi})-\sin(\sqrt{\pi}) = 0$.
   - If we let $x = \pi$, then $\sin(x \cdot y) = \sin(x) + \sin(y) \Rightarrow \sin(\pi \cdot y) = \sin(y) \Rightarrow \pi y = y + 2k\pi$, where $k \in \mathbb{Z}$. Solving for y yields a set of solutions: $\{(\pi, \frac{2k\pi}{\pi-1}) : k \in \mathbb{Z}\}$.

   (b) Find a pair of positive real numbers x and y that demonstrates that $\sqrt{x \cdot y}$ need not equal $\sqrt{x}+\sqrt{y}$. Find a pair of non-zero values for which, by coincidence, this relationship does hold.

   Sample Response:
   Most pairs of positive real numbers x and y will not have this property. For example, for $x = 1$ and $y = 4$, $\sqrt{x \cdot y} = \sqrt{4} = 2$. But $\sqrt{x} + \sqrt{y} = \sqrt{1} + \sqrt{4} = 1 + 2 = 3$, which is clearly not equal to 2. On the other hand,

   - The pair $(4, 4)$ does work: $\sqrt{x \cdot y} = \sqrt{16} = 4 = 2 + 2 = \sqrt{4} + \sqrt{4} = \sqrt{x} + \sqrt{y}$.
   - If we let $x = k^2$ for some $k \in \mathbb{Z}$, then $\sqrt{x \cdot y} = \sqrt{x} + \sqrt{y} \Rightarrow |k|\sqrt{y} = |k| + \sqrt{y} \Rightarrow y = (\frac{|k|}{|k|-1})^2$. A set of solutions is thus $\{(k^2, (\frac{|k|}{|k|-1})^2) : k \in \mathbb{Z}\}$

   (c) Is there a pair of non-zero real numbers x and y for which $2(x \cdot y) = 2x + 2y$? If so, describe how to find all such pairs.

   Sample Response:
   Yes. Note that $2xy = 2x + 2y \Rightarrow xy = x + y \Rightarrow y = \frac{x}{x-1}$. Any pair $(x, \frac{x}{x-1})$ is a solution; for example, $(2, 2)$.

Commentary:
If your class is already comfortable with their solutions to these problems and you would like to initiate a more formal, abstract discussion (rather than an empirical one), you could use the following questions:

- Suppose f is a function on the real numbers such that $f(x \cdot y) = f(x) + f(y)$ for all x and y. What can you say about the value of $f(1)$?
- How can we use this conclusion to argue that the functions in parts (a)–(c) cannot preserve multiplication in the same way as the logarithm function?

With one minor additional hypothesis, it turns out that the logarithm is in fact the unique group homomorphism between $(\mathbb{R}^+, \cdot)$ and $(\mathbb{R}, +)$. See Dieudonne (1960, p. 82). Lastly, for further details on the history of logarithms, see Clark and Montelle (2011).

Problem 1 serves two important purposes: first, it establishes that a homomorphism must preserve the operation between any two arbitrary elements of a set; that is, undergraduates see that even trigonometric functions can be homomorphic on a subset of their domain. Second, it shows that there really was a need to "invent" a new function to translate multiplication into addition, since no extant functions at the time achieved this goal.

Pre-Activity Problems 2 & 3

2. A classmate claims that if everyone collectively forgot how to multiply two numbers together, logarithms would be useful for overcoming the memory lapse. That is, logarithms would be useful for computing products. Describe how you might still be able to compute the product $2 \cdot 3$ using the table of values for the logarithm function, given below, and the identity $\log(x \cdot y) = \log(x) + \log(y)$.

x	1	2	3	4	5	6	7	8	9
$\log(x)$	0	.301	.477	.602	.699	.778	.845	.903	.954

Sample Response:
I can use the chart to see that $\log(2) = .301$ and $\log(3) = .477$. These values added together are .778, which we see is the value for $\log(6)$. Since $\log(2 \cdot 3) = \log(2) + \log(3) = \log(6)$, and because we know the logarithm function is bijective (and thus injective), we conclude $2 \cdot 3 = 6$.

3. Another classmate claims that since 72 is the product of 8 and 9, the table can be used to calculate $\log(72)$. They claim that this is possible even though 72 is not an entry in the first row of the table. Moreover, they claim that the table can be used to compute the logarithm of any number which is the product of numbers represented in the first row of the table. Use the table and the identity $\log(x \cdot y) = \log(x) + \log(y)$ to:

   (a) Calculate $\log(15)$.

   Solution:
   Since $15 = 5 \cdot 3$, we have $\log(15) = \log(5 \cdot 3) = \log(5) + \log(3) = .699 + .477 = 1.176$.

   (b) Calculate $\log(24)$ in two different ways.

   Solution(s):
   - Since $24 = 8 \cdot 3$, we have $\log(24) = \log(8 \cdot 3) = \log(8) + \log(3) = .903 + .477 = 1.380$.
   - Since $24 = 6 \cdot 4$, we have $\log(24) = \log(6 \cdot 4) = \log(6) + \log(4) = .778 + .602 = 1.380$.

(c) Estimate $\log(17)$. Why isn't it possible to calculate $\log(17)$ precisely?

Sample Response:
Since 17 is prime, we cannot factor it and use the chart. Instead, we might calculate $\log(16) = \log(4) + \log(4) = 1.204$ and $\log(18) = \log(9) + \log(2) = 1.255$ and then average them. By this method of estimation, $\log(17) \approx 1.230$.

Commentary:
Consider using some of the following questions to engage groups who finish quickly:

- Is it true that $\log(x \cdot y \cdot z) = \log(x) + \log(y) + \log(z)$? How do you know?
- What's the fewest number of entries on a logarithm chart such as this one that you would need to calculate the log values of all the integers up to 20? 100?
- How precise is your answer to Problem 3(c)? How could you make it more accurate?

To transition to the Class Activity, ask one or more groups to share their solution to Problem 2. Allow the other groups to comment on whether they agree or disagree with the procedure. Tell the class that in today's activity they will compare and contrast the logarithm with other functions that "preserve an operation" to better see what additional properties the logarithm has that make it especially appropriate for these kinds of calculations.

Class Activity: Problems 1–4 (30 minutes)

Prepare undergraduates for the Class Activity by asking them to brainstorm everything they know about logarithms and logarithmic functions. Write all undergraduate answers on the board—if necessary, discuss discrepancies and highlight correct and precise responses. Some possible undergraduate responses are listed below; if it is not supplied by your class, make sure that you lead them to generate the first item (possibly by using the second or third items).

- Bijective
- Injective/ Surjective
- Inverse of the exponential function
- Monotonically increasing
- Concave down
- The natural logarithm has base e and is written $\ln(x)$; the common logarithm has base 10 and is written $\log(x)$.
- The graph of the logarithm function is stretched or shrunk vertically by changing the base of the logarithm.
- $\log(1) = 0$; $\log(x) < 0$ for $x < 1$ and $\log(x) > 0$ for $x > 1$

Leave this list on the board. Undergraduates will need to refer to it in Problem 4.

Distribute the Class Activity. Allow your class to work in small groups on Problem 1. As you monitor their work in groups, you might use some of the following questions to prompt discussion:

- How might you use $\log(x \cdot y) = \log(x) + \log(y)$ to explain why it makes sense that the values of $\log(x)$ are negative when $x < 1$?
- What difficulties did you encounter when attempting to find the product of 1.2 and 1.2 using this chart? How did you overcome them?

For further details on the use of logarithms by NASA in the Apollo program, see Nadworny (2014).

Class Activity Problem 1

The steps used by astronomers and other scientists to calculate the product $x \cdot y$ using a logarithm function looked like the following:

- First, calculate $\log(x)$ and $\log(y)$.
- Next, add these values together.
- Finally, find a positive real number whose logarithm is the sum $\log(x) + \log(y)$. This number is $x \cdot y$.

For the first and last step, charts called log tables were created which recorded, as efficiently as possible, $\log(x)$ for all real numbers x. A (simplified) log table is included below.

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0	n/a	-1.000	-0.699	-0.523	-0.398	-0.301	-0.222	-0.155	-0.097	-0.046
1	0.000	0.041	0.079	0.114	0.146	0.176	0.204	0.230	0.255	0.279
2	0.301	0.322	0.342	0.362	0.380	0.398	0.415	0.431	0.447	0.462
3	0.477	0.491	0.505	0.519	0.531	0.544	0.556	0.568	0.580	0.591
4	0.602	0.613	0.623	0.633	0.643	0.653	0.663	0.672	0.681	0.690
5	0.699	0.708	0.716	0.724	0.732	0.740	0.748	0.756	0.763	0.771
6	0.778	0.785	0.792	0.799	0.806	0.813	0.820	0.826	0.833	0.839
7	0.845	0.851	0.857	0.863	0.869	0.875	0.881	0.886	0.892	0.898
8	0.903	0.908	0.914	0.919	0.924	0.929	0.934	0.940	0.944	0.949
9	0.954	0.959	0.964	0.968	0.973	0.978	0.982	0.987	0.991	0.996

For example, $\log(1.3)$ is found at the intersection of the row marked “1” and the column marked “.3”, so $\log(1.3) = 0.114$. Log tables (and a related tool, the slide rule) were in use well into the 1900s, and were even used by NASA to make calculations for the Apollo 11 moon landing.

1. Use the above process and the log table provided to find the product of the following numbers:

   (a) 1.5 and 2.0

   Solution:
   $\log(1.5) + \log(2.0) = 0.176 + 0.301 = 0.477 \Rightarrow 1.5 \cdot 2.0 = 3.0$

   (b) 0.5 and 8.6

   Solution:
   $\log(0.5) + \log(8.6) = -0.301 + 0.934 = 0.633 \Rightarrow 0.5 \cdot 8.6 = 4.3$

   (c) 1.2 and 1.2

   Sample Response:
   $\log(1.2) + \log(1.2) = 0.079 + 0.079 = 0.158 \Rightarrow 1.2 \cdot 1.2 \approx 1.4$

Once you have seen that most groups have finished Problem 1, initiate a classroom discussion in which you ask groups to describe their process for 1(c). Because 0.158 is not on the chart, undergraduates may have different strategies for finding the value of 1.2^2. For example, undergraduates may say that:

- 0.158 is closer to 0.146 than 0.176, so $1.2^2 \approx 1.4$.
- Because $0.176 - 0.146 = 0.030$ and $0.158 - 0.146 = 0.012$, we can linearly approximate the value of 1.2^2 as $1.4 + \frac{0.012}{0.030} \times 0.1 = 1.4 + 0.04 = 1.44$.

Afterward, allow your class to work in small groups on Problem 2. If your class is unfamiliar with linear algebra concepts, before giving them time to work on this task you might choose to first remind them how to find the determinant of a 2×2 matrix and/ or how to tell if a 2×2 matrix is invertible using the determinant.

Class Activity Problem 2

2. Another setting in advanced mathematics where a function “preserves” a difficult operation is found in linear algebra. Consider the set $GL_2(\mathbb{R}) = \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : a, b, c, d \in \mathbb{R}, ad - bc \neq 0\}$, which consists of all invertible 2×2 matrices with real number entries, and the determinant function $det : GL_2(\mathbb{R}) \to \mathbb{R}$,

which assigns to each matrix its determinant.

Recall from linear algebra that for any two matrices M and N in $GL_2(\mathbb{R})$, we have that $\det(\mathrm{MN}) = \det(\mathrm{M}) \cdot \det(\mathrm{N})$. Choose two 2×2 matrices that you know are invertible and verify this identity holds for your two matrices.

Solution:
Answers will vary. Undergraduate responses should largely mirror the following example:

$$\begin{aligned}\det(\mathrm{MN}) &= \det\left(\left[\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right]\left[\begin{smallmatrix}2&0\\0&1\end{smallmatrix}\right]\right)\\ &= \det\left(\left[\begin{smallmatrix}2&1\\0&1\end{smallmatrix}\right]\right)\\ &= 2\\ &= 1\cdot 2\\ &= \det(\mathrm{M})\cdot\det(\mathrm{N})\end{aligned}$$

When you call the class back together to compare answers, instead of asking each group to report out individually, ask whether any group found a pair of matrices for which this identity did NOT appear to hold. Address their proposed counterexample on the board. If instead everyone agrees that this identity should hold for any two matrices, you may choose not to do your own example on the board in the interest of time. Let undergraduates work in small groups on the next two problems simultaneously. Make sure to circulate the classroom to observe their progress.

Class Activity Problems 3 & 4

3. Finn hopes that the determinant function (see previous question) can help them to avoid multiplying matrices, much like the logarithm helped astronomers avoid multiplication of large, real numbers. Finn models the following strategy based on the process given before Problem 1:

   - First, calculate $\det(\mathrm{M})$ and $\det(\mathrm{N})$
   - Next, multiply these values together.
   - Finally, find a matrix whose determinant is the product $\det(\mathrm{M}) \cdot \det(\mathrm{N})$. This matrix is MN.

   (a) Why won't Finn's strategy work?

   Solution:
   At the last step, there are multiple matrices that have the same determinant. Simply finding ONE matrix whose determinant is the same as the product $\det(\mathrm{M}) \cdot \det(\mathrm{N})$ doesn't guarantee that this matrix is the product MN.

   (b) What question might you ask Finn to help them see the flaw in this plan? Why do you think this question would be helpful?

   Sample Responses:

   - Can two matrices have the same determinant? This question would help Finn see that a matrix is not uniquely determined by its determinant.
   - Is the determinant function injective? Why is this property important? This sequence of questions would help Finn consider whether det has an inverse function that would allow their procedure to work.
   - What is the determinant of $\left[\begin{smallmatrix}1&0\\0&4\end{smallmatrix}\right]$? What about $\left[\begin{smallmatrix}2&0\\0&2\end{smallmatrix}\right]$? By asking Finn to calculate the determinant of two different matrices, they will see that the determinant function cannot be injective and that the last step of the process isn't reasonable.

4. What is a key property that the logarithm function applied to positive real numbers possesses that the determinant function applied to invertible matrices does NOT possess?

   Solution:
   The key idea is that the logarithm function is bijective, but that the determinant function is not. Another sample response: "You can have two matrices with the same determinant so the determinant function isn't one-to-one."

Commentary:
As you monitor the undergraduates during their group work,

- If undergraduates are having trouble on Problem 3, encourage them to attempt to apply Finn's strategy to the matrices M and N which they chose in Problem 2.
- Ask undergraduates to refer to the list of properties of the logarithm function, which should still be on the board, when they reach Problem 4.

Questions you might use to prompt discussion:

- How can we use $\det(\mathrm{MN}) = \det(\mathrm{M}) \cdot \det(\mathrm{N})$ to describe the relationship between the determinant of a matrix and the determinant of its inverse?
- Do you think the homomorphism property of the determinant function still holds for non-invertible matrices? Why or why not?

Call the class back together and allow volunteers to report out on Problems 3(a) and 4 especially. Point out that, while the determinant map considered here still "preserves an operation" like the logarithm, because it is not also bijective it cannot be used to "avoid" the original operation: in this case, matrix multiplication.

Before transitioning to the next portion of the lesson, discuss the following connection to teaching with your class.

Discuss This Connection to Teaching

Problems such as Problem 3(b) are important because it is useful for all undergraduates to think about how others use and reason with mathematics. Posing questions gives undergraduates the opportunity to think about how they would respond to another person's work in ways that would help that person develop an understanding of the concept, which improves both their own understanding of the content and their ability to communicate technical mathematics.

Discussion: Homomorphisms and Isomorphisms (10 minutes)

To bridge to the next class activity, take time to have undergraduates understand that the logarithm map is a bijection, but the determinant map is not. This key difference can be one way to motivate the definitions of homomorphisms and isomorphisms of groups. This can be orchestrated by sequencing student work from the previous problems or other ways that the instructor feels are appropriate for their context. Use the notation from your own notes or your own textbook, but emphasize or highlight the following:

- Homomorphisms "preserve the operation"; that is, one can do the group operation either in the domain (before applying the homomorphism) or in the codomain (after applying the homomorphism) and expect the same result.
- Because they both "preserve the operation," the image of a group homomorphism (and hence, a group isomorphism) is itself a group. (A proof of this is included as a sequence of homework exercises.)
- Isomorphisms, since they are bijections, also preserve the cardinality of the sets.

At this point in the class discussion, discuss the following connection to teaching.

Discuss This Connection to Teaching

High school students often use the inverse relationship between exponents and logarithms to solve problems involving exponents and logarithms. Prospective teachers can leverage their understanding that the logarithm map is a bijection to discuss why it is appropriate to apply logarithms when solving exponential equations or use exponential functions when solving logarithmic equations.

After introducing definitions of homomorphisms and isomorphisms of groups, emphasize that one of the many uses of an isomorphism is that isomorphic groups share the same structural characteristics. Consider using simpler language first to provide some of the following illustrative examples of such characteristics: If G and H are isomorphic, then

- G and H have the same number of elements (i.e., the same cardinality).
- If the elements of G commute, then so do the elements of H.
- If G has an element g such that $g^n = e_G$, then H has an element h such that $h^n = e_H$.
- If G is generated by a single element, so is H.
- If every element of G is its own inverse, every element of H is its own inverse.

If undergraduates are already familiar with any of the key vocabulary necessary (order of an element, cyclic, abelian, etc.) to list these properties with mathematical precision, you may instead choose to write them formally. Proofs of these properties are included in the additional homework section of this document. Another useful analogy, if your undergraduates are comfortable working with Cayley tables, is that an isomorphism between two groups acts as a "key" that lets you easily fill in one table assuming you already know what the other looks like. The Klein 4-group and $\mathbb{Z}_2 \times \mathbb{Z}_2$ illustrate this concept well.

Class Activity: Problems 5–8 (20 minutes)

Problem 5 can be asked and answered at the front of the room without requiring the undergraduates to consult each other in groups. Make sure that undergraduates are specific about the group operations of the domain and codomain. If you feel it would benefit your class, you might also ask: "Between which two groups is the determinant function a homomorphism?"

Class Activity Problem 5

5. Between which two groups is the logarithm function an isomorphism?

   Solution:
   The logarithm is an isomorphism from $(\mathbb{R}^+, \cdot)$ to $(\mathbb{R}, +)$.

Use Problem 5 to discuss the following connection to teaching before moving on to the remainder of the Class Activity:

Discuss This Connection to Teaching

In mathematics, it is important to use precise language in order to avoid ambiguity. When we name a group, we are sure to include its operation; when we claim a function is an isomorphism, it must be an isomorphism between two spaces. Secondary school teachers must also take steps to ensure that both they and their students use mathematical language in an unambiguous way. For example, in geometry, students often say that two objects (such as triangles) are "the same." In this case, teachers who have developed sensitivity for precise language may prompt their students to be more specific. That is, are the triangles "the same" because they are congruent or because they are similar? Or are they the same type of triangle (e.g., obtuse, right) but otherwise of a different shape?

Allow undergraduates to work in small groups on Problems 6, 7, and 8 simultaneously. If appropriate, you may want to discuss Problem 6 as a class first. Consider using the following sequence of prompts to help motivate this discussion:

- Use our list to describe the group structure of $(GL_2(\mathbb{R}), \cdot)$ and $(\mathbb{R}^*, \cdot)$.
- What would we know about the structure of the two groups if there were a group isomorphism between them?

As undergraduates work, walk around the classroom and check for interesting approaches and help correct misunderstandings. Also, as you monitor your class,

- Encourage groups that finish quickly to look for more than one structural difference between the groups in each question.
- Where appropriate, ask groups to justify their claims about the structure of the groups in these problems, at least informally. For example:
  - If $(\mathbb{Q}^+, \cdot)$ were cyclic, what would its generator be? Why doesn't this make sense?
  - Is there a bijection between $\mathbb{R}$ and $\mathbb{Z}$ when they are viewed just as sets?

Class Activity Problems 6, 7 & 8

6. Now that Finn understands that the function det is not a group isomorphism and will not help avoid matrix multiplication, they want to try and find another function from $(GL_2(\mathbb{R}), \cdot) \to (\mathbb{R}^*, \cdot)$ which is a group isomorphism, where $\mathbb{R}^*$ denotes the set of nonzero real numbers. Does such a function exist? Why or why not?

   Solution:
   No. $(GL_2(\mathbb{R}), \cdot)$ is not an abelian group, but $(\mathbb{R}^*, \cdot)$ is. Since the structure of the two groups is different, there can be no isomorphism between them.

7. Finn shifts focus to finding a "better logarithm"; that is, a group isomorphism from $(\mathbb{R}^+, \cdot) \to (\mathbb{Z}, +)$ that allows them to multiply positive real numbers by adding integers instead of real numbers. Does such a function exist? Why or why not?

   Solutions:

   - No. $\mathbb{R}^+$ is an uncountable set, but $\mathbb{Z}$ is countable. Since the two groups have different orders, there can be no isomorphism between them.
   - No. $(\mathbb{Z}, +) = \langle 1 \rangle$, while $(\mathbb{R}^+, \cdot)$ is not cyclic (it is uncountable and all cyclic groups are at most countable). Since the two groups have different algebraic structures, there can be no isomorphism between them.

8. Maybe, Finn says, we could at least find a group isomorphism from $(\mathbb{Q}^+, \cdot) \to (\mathbb{Z}, +)$ that lets us multiply positive rational numbers by adding integers. Does such a function exist? Why or why not?

   Solution:
   No. $(\mathbb{Z}, +) = \langle 1 \rangle$, but $(\mathbb{Q}^+, \cdot)$ is not cyclic. Since the structure of the two groups is different, there can be no isomorphism between them.

Wrap-Up (5 minutes)

Recap the lesson briefly for the class:

- Homomorphisms "preserve the operation" of a group. Consequently, the image of a group homomorphism (and hence, a group isomorphism) is itself a group.
- Two groups with an isomorphism between them have matching algebraic structures (e.g., identical orders for corresponding elements, commutativity between corresponding elements is preserved, etc.)

Emphasize that, as we saw with the logarithm function, isomorphisms also allow one to avoid working in a particular group with a difficult operation by mapping elements into the codomain group and working with its operation before mapping back. In this way, we can understand everything essential about a group by studying a different group to which it is isomorphic.

Discuss This Connection to Teaching

Students are familiar with the fact that linear functions of the form $h(x) = cx$, where $c \in \mathbb{R}$, preserve structure. For example, if $f(x) = 2x$, we have $f(a+b) = 2(a+b) = 2a+2b = f(a)+f(b)$. However, they often overgeneralize this understanding when working with $g(x) = x^2$ even though $g(a+b) = (a+b)^2 \neq a^2 + b^2 = g(a) + g(b)$. Prospective teachers' advanced understanding of group isomorphisms supports their capacity to discuss how the logarithm function, say, "converts" multiplication problems into addition problems which can also be a common way to motivate the technique of logarithmic differentiation in calculus.

We end the lesson using an exit ticket. See Chapter 1 for guidance about how to conclude mathematics lessons with exit tickets.

Homework Problems

At the end of the lesson, assign the following homework problems.

Undergraduates need to know that the fact that homomorphisms preserve the operation means that they also preserve some aspects of the structure of a group. Problem 1 illustrates that the image of homomorphism acting on a group will be a subgroup of the codomain.

Homework Problem 1

1. Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups and $f : G \to H$ be a homomorphism.
   (a) Show that the image of the group identity of G under f must be the group identity of H (that is, $f(e_G) = e_H$).
   Solution:
   First, note that $f(e_G) = f(e_G \cdot_G e_G)$. Because f is a homomorphism, we may rewrite this equation as $f(e_G) = f(e_G) \cdot_H f(e_G)$. By definition, then, $f(e_G) = e_H$.
   (b) Use this result to show that, given an arbitrary element a in G, the image of the inverse of a under f must be the inverse of the image of a under f (that is, $\forall a \in G$, $f(a^{-1}) = f(a)^{-1}$).
   Solution:
   Let $a \in G$.Using the result from 1(a) and the fact that f is a homomorphism, $e_H = f(e_G) = f(a \cdot_G a^{-1}) = f(a) \cdot_H f(a^{-1})$. Similar work shows that $e_H = f(a^{-1}) \cdot_H f(a)$. Then, by the uniqueness of a group inverse, we have that $f(a^{-1}) = f(a)^{-1}$, as desired.
   (c) Use the previous two results to show that the image of G under f is itself a group with the operation of H.
   Solution:
   - Closure follows directly from the fact that f is a homomorphism: Let $f(a), f(b) \in f(G)$. Then, $f(a) \cdot_H f(b) = f(a \cdot_G b) \in f(G)$.
   - Associativity of $f(G)$ is inherited from associativity of H.
   - By 1(a), $e_H = f(e_G) \in f(G)$.
   - By 1(b), an element $f(a) \in f(G)$ has inverse $f(a^{-1})$

   Thus, $f(G)$ is a group under the same operation as H.

In Problem 2, undergraduates prove one of the claims made during the discussion of the Class Activity. Before doing so, they analyze a hypothetical student's attempt at a proof and identify an error; importantly, they also evaluate how the given proof might not be incorrect in a different context.

Homework Problem 2

2. Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups and $f : G \to H$ be a homomorphism. Furthermore, let G be an abelian group. Anjali writes the following proof, which she claims shows that H is also an abelian group:

Let g_1 and g_2 be two distinct elements of G. Then:

$g_1 \cdot_G g_2 = g_2 \cdot_G g_1$ ← G is abelian

$f(g_1 \cdot_G g_2) = f(g_2 \cdot_G g_1)$

f is a homomorphism

$f(g_1) \cdot_H f(g_2) = f(g_2) \cdot_H f(g_1)$

$f(g_1)$ and $f(g_2)$ are in H

$h_1 \cdot_H h_2 = h_2 \cdot_H h_1$

So, H is also an abelian group.

(a) Explain why Anjali's proof does not show that H is an abelian group. What has she proven instead?

Solution:
At the last line, h_1 and h_2 are not arbitrary elements of H; in fact, they may not even be distinct elements. Anjali has instead proven that elements in the image of G under f commute with each other.

(b) Given the same groups $(G, \cdot_G)$ and $(H, \cdot_H)$ let $f : G \to H$ now be an **isomorphism**. Add to Anjali's proof to show that, under these conditions, H is now an abelian group.

Solution:
Before claiming that H is abelian, Anjali should add: "Because f is injective, we know that h_1 and h_2 are two distinct elements of H. Because f is also surjective, we know that every element of H is in the image of f; thus, h_1 and h_2 also represent arbitrary elements of H."

Problem 3 helps undergraduates appreciate that saying "the logarithm is a homomorphism" is not sufficiently specific; one needs to specify the groups and operations involved. Furthermore, it gives undergraduates the opportunity to explore how they might translate their advanced mathematical understanding of logarithms into language appropriate for a high school student.

Homework Problem 3

3. Imagine that you are a high school mathematics teacher who has noticed that one of your students, Hai, has assumed that $\log(x + y) = \log(x) + \log(y)$ when simplifying logarithmic expressions on a homework assignment.

(a) Find a pair of real numbers x and y demonstrating that Hai's assumption is not always true. For which pairs of positive real numbers (x, y) is Hai's assumption true?

Sample Response:
For $x = 1$ and $y = 2$, $\log(x + y) = \log(3)$ but $\log(x) + \log(y) = \log(1) + \log(2) = \log(2)$. Clearly, $\log(3) \neq \log(2)$. More generally, choosing $x = 1$ and $y = a$ will yield $\log(1 + a) =$

$\log(a)$, which can never be true given that the logarithm function is a bijection. Then:

$$\log(x+y) = \log(x) + \log(y)$$
$$10^{\log(x+y)} = 10^{\log(x)+\log(y)} = 10^{\log(x)} \cdot 10^{\log(y)}$$
$$x + y = x \cdot y$$
$$y = \frac{x}{x-1}$$

So any pair of real numbers (x, y) with $x, y > 0$ satisfying $y = \frac{x}{x-1}$ will also satisfy $\log(x+y) = \log(x) + \log(y)$.

(b) Using what you have learned about logarithms in this lesson, how might you help Hai understand that $\log(x \cdot y) = \log(x) + \log(y)$ is the correct identity? Why is your explanation helpful? Make sure your explanation is appropriate for a high school student.

Sample Responses:

- I could explain that the logarithm was created in order to help astronomers turn multiplication problems into addition problems. That means that one side of the identity needs to feature multiplication as an operation somewhere.
- I know that the logarithm function is a group isomorphism, so it should be injective. I could show Hai my work for part (a) to demonstrate that his identity leads to a logarithm function that is NOT injective; then, I might show him that choosing $x = 1$ in $\log(x \cdot y) = \log(x) + \log(y)$ does not create such a problem.

The familiar additive group of integers modulo n is shown to be related to the additive group of integers by a natural mapping. In Problem 4(b), undergraduates need to both argue from the definition of an isomorphism and from the concept of an isomorphism as structure-preserving to make their point. This demonstrates a thorough and versatile understanding of the material.

Homework Problem 4

4. Recall that $\mathbb{Z}_n$ is the set of *equivalence classes* on the integers, where two integers are in the same equivalence class if and only if they both have the same (smallest, non-negative) remainder when divided by n. The set $\mathbb{Z}_n$ contains n such equivalence classes which, canonically, are represented by the possible remainders when an integer is divided by n: $\{0, 1, \ldots, n-2, n-1\}$.

(a) Let θ be the map from $(\mathbb{Z}, +)$ to $(\mathbb{Z}_n, +)$ given by $\theta(z) = r$, where z is an integer and r is its remainder when divided by n. Show that θ is a group homomorphism. You may assume without proof that $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +)$ are in fact groups.

Solution:

Let $a, b \in \mathbb{Z}$. Then by the division algorithm, $\exists\, q_1, q_2, r_1, r_2 \in \mathbb{Z}$ such that $a = q_1 \cdot n + r_1$ and $b = q_2 \cdot n + r_2$, where $0 < r_1, r_2 < n$. Note that $\theta(a+b) = \theta((q_1+q_2) \cdot n + r_1 + r_2)$. If $r_1 + r_2 < n$, then $\theta((q_1+q_2) \cdot n + r_1 + r_2) = r_1 + r_2 = \theta(a) + \theta(b)$, as desired.

If $r_1 + r_2 \geq n$, then $\exists\, q_3, r_3 \in \mathbb{Z}$ such that $r_1 + r_2 = q_3 \cdot n + r_3$, where $0 < r_3 < n$. Then, $\theta((q_1+q_2) \cdot n + r_1 + r_2) = \theta((q_1+q_2+q_3) \cdot n + r_3) = r_3$. Note that here, r_3 represents the equivalence class containing those integers that differ from r_3 by a multiple of n. In particular, this equivalence class includes the element $r_3 + q_3 \cdot n = r_1 + r_2 = \theta(a) + \theta(b)$. So, $\theta(a+b) = \theta(a) + \theta(b)$, as desired.

(b) Explain why θ cannot be an isomorphism in two ways: by showing that θ is not a bijection and by finding a difference in the algebraic structures of $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +)$ (other than that their orders differ).

Solution:

- θ is not injective: $\theta(n) = \theta(2n) = 0$, but $n \neq 2n$ given that $n > 1$.
- The order of the elements are also different: for any nonzero $z \in (\mathbb{Z}, +)$, $|z| = \infty$; however, every nonzero element of $(\mathbb{Z}_n, +)$ has finite order.

Problem 5 accomplishes two goals: it ties important calculus knowledge to the abstract algebra curriculum and, in part (b), it helps undergraduates informally verify that the kernel of a homomorphism is always a subgroup of the domain.

Homework Problem 5

5. Let F be the set of continuous, real-valued functions on the interval $[0, 1]$. Let σ be the map from $(F, +)$ to $(\mathbb{R}, +)$ given by $\sigma(f) = \int_0^1 f(x)dx$ for all f in F. The operation "+" on F is defined by $(f + g)(x) = f(x) + g(x)$ for all f, g in F and $x \in [0, 1]$. The operation "+" on $\mathbb{R}$ is the usual real number addition.

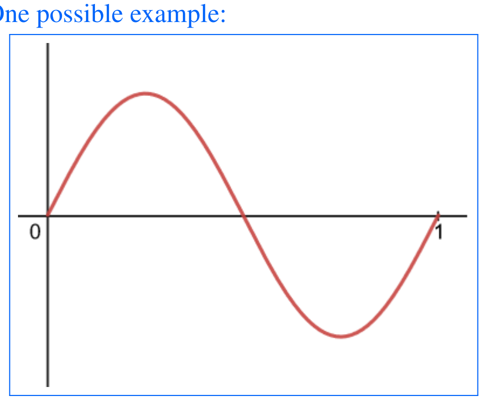
   (a) Show that σ is a group homomorphism. You may assume without proof that $(F, +)$ and $(\mathbb{R}, +)$ are in fact groups.

   Solution:
   Let $f, g \in F$. Then: $\sigma(f + g) = \int_0^1 (f + g)(x)dx = \int_0^1 [f(x) + g(x)]dx = \int_0^1 f(x)dx + \int_0^1 g(x)dx = \sigma(f) + \sigma(g)$, as desired.

   (b) Describe the function that represents the identity element in $(F, +)$. Now, choose a different element of $(F, +)$ which maps to the identity element of $(\mathbb{R}, +)$ under the map σ. Draw a graph of your chosen function and explain how you know it meets this criteria.

   Solution:

   - The function that acts as the identity element in $(F, +)$ is the function whose values are equal to 0 for every $x \in [0, 1]$.
   - Undergraduate graphs will vary, but the integral of each graphed function over the interval $[0, 1]$ should be zero. One possible example:



   - Note that the identity of $(\mathbb{R}, +)$ is 0. If for $f \in F$ we have that $\sigma(f) = 0$, then the area bounded between f and the x-axis is the same both above and below the x-axis.

   (c) Let $ker(\sigma)$ be the set of **all** elements of $(F, +)$ that map to the identity element of $(\mathbb{R}, +)$ under the map σ. Show that $ker(\sigma)$ is a subgroup of $(F, +)$.

Solution:

- Let $f, g \in ker(\sigma)$. Then $\sigma(f+g) = \sigma(f) + \sigma(g) = 0 + 0 = 0$, so $ker(\sigma)$ is closed.
- $ker(\sigma)$ inherits associativity from F.
- Let e be the zero function given by $e(x) = 0$ for all $x \in [0,1]$. This is the identity element of $(F, +)$ described in part (a). Then, $\sigma(e) = 0$, so $e \in ker(\sigma)$.
- Let $f \in ker(\sigma)$. Since $f(x) + (-f)(x) = (f - f)(x) = e(x)$ for all $x \in [0,1]$, the function $-f$ is the inverse of f. To see that $-f$ must also be in $ker(\sigma)$, note that $\sigma(-f) = \int_0^1 (-f)(x)dx = -\int_0^1 (f)(x)dx = 0$.

Thus, $ker(\sigma)$ is a group.

Assessment Problems

The following two problems address ideas explored in the lesson, with a focus on connections to teaching and mathematical content. You can include these problems as part of your usual course quizzes or exams.

Assessment Problems 1 & 2

1. Let $M_2(\mathbb{R}) = \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : a, b, c, d \in \mathbb{R}\}$ and define the function $tr : (M_2(\mathbb{R}), +) \to (\mathbb{R}, +)$ to be the map $tr(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]) = a + d$. That is, tr assigns to each matrix its trace.

   (a) Show that tr is a group homomorphism. You may assume without proof that $(M_2(\mathbb{R}), +)$ and $(\mathbb{R}, +)$ are in fact groups.

   Solution:

$$\begin{aligned} tr(\mathrm{M} + \mathrm{N}) &= tr(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] + \left[\begin{smallmatrix} e & f \\ g & h \end{smallmatrix}\right]) \\ &= tr(\left[\begin{smallmatrix} a+e & b+f \\ c+g & d+h \end{smallmatrix}\right]) \\ &= a + e + d + h \\ &= (a + d) + (e + h) \\ &= tr(\mathrm{M}) + tr(\mathrm{N}) \end{aligned}$$

   (b) Is tr an isomorphism? Why or why not?

   Solution:

   The function tr is not injective, and thus not a bijection. Two matrices can certainly have the same trace but represent fundamentally different linear operators:

   $tr([\begin{smallmatrix} 1 & 7 \\ 0 & 1 \end{smallmatrix}]) = 2$

   $tr([\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}]) = 2$

2. Abina claims that there must be an isomorphism between $(\mathbb{Z}_4, +)$ and $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ since the orders of the groups are the same.

   (a) Based on her claim, what do you think Abina understands about isomorphisms?

   Sample Responses:

   - Abina knows that isomorphisms are bijections, and so both the domain and the codomain must have the same order.
   - Abina knows that isomorphisms preserves structure, e.g., order.
   - I think she understands that if two groups are isomorphic that they have the same order, but is thinking she can flip the "if, then" statement and say "If two groups have the same order, then they are isomorphic."

(b) Explain why the two groups in question cannot be isomorphic.

Solution(s):

- Every element of $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ is its own inverse. However, the inverse of 1 in $(\mathbb{Z}_4, +)$ is 3 since $1 + 3 = 4 = 0$ (mod 4).
- While $(\mathbb{Z}_4, +)$ is cyclic, $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ is not. The groups generated by the elements of $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ are:

  $\langle(0,0)\rangle = \{(0,0)\}$
  $\langle(1,0)\rangle = \{(0,0),(1,0)\}$
  $\langle(0,1)\rangle = \{(0,0),(0,1)\}$
  $\langle(1,1)\rangle = \{(0,0),(1,1)\}$

  Since none of these is the whole group, $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ is not cyclic.

(c) What question would you ask Abina to help her understand her mistake? Why do you think your question would be helpful?

Sample Responses:

- What is the difference between an isomorphism and a bijection? This question will help Abina realize that she is not attending to the entire definition of isomorphism.
- What is a homomorphism, and how is it different from an isomorphism? This question will help Abina realize that she is not attending to the entire definition of isomorphism.
- Is $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ cyclic? This question will help Abina identify that the algebraic structure of the groups differ in a key way.
- What is the order of each element of $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$? Of $(\mathbb{Z}_4, +)$? This question will help Abina identify that the algebraic structure of the groups differ in a key way.
- We know that $(\mathbb{Z}, +)$ and $(\mathbb{Q}^+, \cdot)$ have the same order. What is different about the structure of these two groups? This question will prompt Abina to compare the current situation to a similar problem from the class activity.

10.6 References

[1] Clark, K. M. & Montelle, C. (2011) *Logarithms: The early history of a familiar function*. Retrieved from https://www.maa.org/press/periodicals/convergence/logarithms-the-early-history-of-a-familiar-function-introduction.

[2] National Governors Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards for Mathematics.* Authors. Retrieved from http://www.corestandards.org/.

10.7 Lesson Handouts

Handouts for use during instruction are included on the pages that follow. LaTeX files for these handouts can be downloaded from maa.org/meta-math.

NAME: **PRE-ACTIVITY:** LOGARITHMS AND ISOMORPHISMS (page 1 of 2)

At the beginning of the 17^{th} century, two different mathematicians were working to simplify tedious calculations in the field of astronomy (where numbers are often very large) by creating a function that would "translate" multiplication into addition. After all, it is much easier to add numbers like 177,320,045 and 9,317,032,566 by hand than it is to multiply them. This work by John Napier from Scotland and Joost Bürgi from Switzerland led to the creation of the function that we now call the logarithm. The precise nature of the logarithm has been refined over the years by other mathematicians, but its original use as a computational aid is exemplified by a familiar identity: for all positive real numbers x and y,

$$\log(x \cdot y) = \log(x) + \log(y)$$

It is useful to think of this property as "preserving an operation"; that is, the logarithm function preserves multiplication by translating it into addition.

1. Of course, most functions on the real numbers do not translate a multiplicative structure into an additive one.

   (a) Find a pair of positive real numbers x and y that demonstrates that $\sin(x \cdot y)$ need not equal $\sin(x) + \sin(y)$. Find a pair of non-zero values for which, by coincidence, this relationship does hold.

   (b) Find a pair of positive real numbers x and y that demonstrates that $\sqrt{x \cdot y}$ need not equal $\sqrt{x} + \sqrt{y}$. Find a pair of non-zero values for which, by coincidence, this relationship does hold.

   (c) Is there a pair of non-zero real numbers x and y for which $2(x \cdot y) = 2x + 2y$? If so, describe how to find all such pairs.

PRE-ACTIVITY: LOGARITHMS AND ISOMORPHISMS (page 2 of 2)

2. A classmate claims that if everyone collectively forgot how to multiply two numbers together, logarithms would be useful for overcoming the memory lapse. That is, logarithms would be useful for computing products. Describe how you might still be able to compute the product $2 \cdot 3$ using the table of values for the logarithm function, given below, and the identity $\log(x \cdot y) = \log(x) + \log(y)$.

x	1	2	3	4	5	6	7	8	9
$\log(x)$	0	.301	.477	.602	.699	.778	.845	.903	.954

3. Another classmate claims that since 72 is the product of 8 and 9, the table can be used to calculate $\log(72)$. They claim that this is possible even though 72 is not an entry in the first row of the table. Moreover, they claim that the table can be used to compute the logarithm of any number which is the product of numbers represented in the first row of the table. Use the table and the identity $\log(x \cdot y) = \log(x) + \log(y)$ to:

   (a) Calculate $\log(15)$.

   (b) Calculate $\log(24)$ in two different ways.

   (c) Estimate $\log(17)$. Why isn't it possible to calculate $\log(17)$ precisely?

NAME: **CLASS ACTIVITY:** LOGARITHMS AND ISOMORPHISMS (page 1 of 4)

The steps used by astronomers and other scientists to calculate the product $x \cdot y$ using a logarithm function looked like the following:

- First, calculate $\log(x)$ and $\log(y)$.
- Next, add these values together.
- Finally, find a positive real number whose logarithm is the sum $\log(x) + \log(y)$. This number is $x \cdot y$.

For the first and last step, charts called log tables were created which recorded, as efficiently as possible, $\log(x)$ for all real numbers x. A (simplified) log table is included below.

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0	n/a	-1.000	-0.699	-0.523	-0.398	-0.301	-0.222	-0.155	-0.097	-0.046
1	0.000	0.041	0.079	0.114	0.146	0.176	0.204	0.230	0.255	0.279
2	0.301	0.322	0.342	0.362	0.380	0.398	0.415	0.431	0.447	0.462
3	0.477	0.491	0.505	0.519	0.531	0.544	0.556	0.568	0.580	0.591
4	0.602	0.613	0.623	0.633	0.643	0.653	0.663	0.672	0.681	0.690
5	0.699	0.708	0.716	0.724	0.732	0.740	0.748	0.756	0.763	0.771
6	0.778	0.785	0.792	0.799	0.806	0.813	0.820	0.826	0.833	0.839
7	0.845	0.851	0.857	0.863	0.869	0.875	0.881	0.886	0.892	0.898
8	0.903	0.908	0.914	0.919	0.924	0.929	0.934	0.940	0.944	0.949
9	0.954	0.959	0.964	0.968	0.973	0.978	0.982	0.987	0.991	0.996

For example, $\log(1.3)$ is found at the intersection of the row marked "1" and the column marked ".3", so $\log(1.3) = 0.114$. Log tables (and a related tool, the slide rule) were in use well into the 1900s, and were even used by NASA to make calculations for the Apollo 11 moon landing.

1. Use the above process and the log table provided to find the product of the following numbers:

   (a) 1.5 and 2.0

   (b) 0.5 and 8.6

   (c) 1.2 and 1.2

2. Another setting in advanced mathematics where a function "preserves" a difficult operation is found in linear algebra. Consider the set $GL_2(\mathbb{R}) = \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : a, b, c, d \in \mathbb{R}, ad - bc \neq 0\}$, which consists of all invertible 2×2 matrices with real number entries, and the determinant function $det : GL_2(\mathbb{R}) \to \mathbb{R}$, which assigns to each matrix its determinant.

   Recall from linear algebra that for any two matrices M and N in $GL_2(\mathbb{R})$, we have that $\det(\mathrm{MN}) = \det(\mathrm{M}) \cdot \det(\mathrm{N})$. Choose two 2×2 matrices that you know are invertible and verify this identity holds for your two matrices.

3. Finn hopes that the determinant function (see previous question) can help them to avoid multiplying matrices, much like the logarithm helped astronomers avoid multiplication of large, real numbers. Finn models the following strategy based on the process given before Problem 1:

   - First, calculate $\det(\mathrm{M})$ and $\det(\mathrm{N})$
   - Next, multiply these values together.
   - Finally, find a matrix whose determinant is the product $\det(\mathrm{M}) \cdot \det(\mathrm{N})$. This matrix is MN.

   (a) Why won't Finn's strategy work?

   (b) What question might you ask Finn to help them see the flaw in this plan? Why do you think this question would be helpful?

4. What is a key property that the logarithm function applied to positive real numbers possesses that the determinant function applied to invertible matrices does NOT possess?

5. Between which two groups is the logarithm function an isomorphism?

6. Now that Finn understands that the function det is not a group isomorphism and will not help avoid matrix multiplication, they want to try and find another function from $(GL_2(\mathbb{R}), \cdot) \to (\mathbb{R}^*, \cdot)$ which is a group isomorphism, where $\mathbb{R}^*$ denotes the set of nonzero real numbers. Does such a function exist? Why or why not?

7. Finn shifts focus to finding a "better logarithm"; that is, a group isomorphism from $(\mathbb{R}^+, \cdot) \to (\mathbb{Z}, +)$ that allows them to multiply positive real numbers by adding integers instead of real numbers. Does such a function exist? Why or why not?

8. Maybe, Finn says, we could at least find a group isomorphism from $(\mathbb{Q}^+, \cdot) \to (\mathbb{Z}, +)$ that lets us multiply positive rational numbers by adding integers. Does such a function exist? Why or why not?

NAME: **HOMEWORK PROBLEMS:** LOGARITHMS AND ISOMORPHISMS (page 1 of 2)

1. Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups and $f : G \to H$ be a homomorphism.
   (a) Show that the image of the group identity of G under f must be the group identity of H (that is, $f(e_G) = e_H$).
   (b) Use this result to show that, given an arbitrary element a in G, the image of the inverse of a under f must be the inverse of the image of a under f (that is, $\forall a \in G$, $f(a^{-1}) = f(a)^{-1}$).
   (c) Use the previous two results to show that the image of G under f is itself a group with the operation of H.

2. Let $(G, \cdot_G)$ and $(H, \cdot_H)$ be groups and $f : G \to H$ be a homomorphism. Furthermore, let G be an abelian group. Anjali writes the following proof, which she claims shows that H is also an abelian group:

   Let g_1 and g_2 be two distinct elements of G. Then:

   $g_1 \cdot_G g_2 = g_2 \cdot_G g_1 \leftarrow$ G is abelian

   $f(g_1 \cdot_G g_2) = f(g_2 \cdot_G g_1)$ f is a homomorphism

   $f(g_1) \cdot_H f(g_2) = f(g_2) \cdot_H f(g_1)$ $f(g_1)$ and $f(g_2)$ are in H

   $h_1 \cdot_H h_2 = h_2 \cdot_H h_1$

   So, H is also an abelian group.

   (a) Explain why Anjali's proof does not show that H is an abelian group. What has she proven instead?
   (b) Given the same groups $(G, \cdot_G)$ and $(H, \cdot_H)$ let $f : G \to H$ now be an **isomorphism**. Add to Anjali's proof to show that, under these conditions, H is now an abelian group.

3. Imagine that you are a high school mathematics teacher who has noticed that one of your students, Hai, has assumed that $\log(x + y) = \log(x) + \log(y)$ when simplifying logarithmic expressions on a homework assignment.
   (a) Find a pair of real numbers x and y demonstrating that Hai's assumption is not always true. For which pairs of positive real numbers (x, y) is Hai's assumption true?
   (b) Using what you have learned about logarithms in this lesson, how might you help Hai understand that $\log(x \cdot y) = \log(x) + \log(y)$ is the correct identity? Why is your explanation helpful? Make sure your explanation is appropriate for a high school student.

4. Recall that $\mathbb{Z}_n$ is the set of *equivalence classes* on the integers, where two integers are in the same equivalence class if and only if they both have the same (smallest, non-negative) remainder when divided by n. The set $\mathbb{Z}_n$ contains n such equivalence classes which, canonically, are represented by the possible remainders when an integer is divided by n: $\{0, 1, \ldots, n-2, n-1\}$.
   (a) Let θ be the map from $(\mathbb{Z}, +)$ to $(\mathbb{Z}_n, +)$ given by $\theta(z) = r$, where z is an integer and r is its remainder when divided by n. Show that θ is a group homomorphism. You may assume without proof that $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +)$ are in fact groups.
   (b) Explain why θ cannot be an isomorphism in two ways: by showing that θ is not a bijection and by finding a difference in the algebraic structures of $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +)$ (other than that their orders differ).

5. Let F be the set of continuous, real-valued functions on the interval $[0, 1]$. Let σ be the map from $(F, +)$ to $(\mathbb{R}, +)$ given by $\sigma(f) = \int_0^1 f(x)dx$ for all f in F. The operation "+" on F is defined by $(f + g)(x) = f(x) + g(x)$ for all f, g in F and $x \in [0, 1]$. The operation "+" on $\mathbb{R}$ is the usual real number addition.

(a) Show that σ is a group homomorphism. You may assume without proof that $(F, +)$ and $(\mathbb{R}, +)$ are in fact groups.

(b) Describe the function that represents the identity element in $(F, +)$. Now, choose a different element of $(F, +)$ which maps to the identity element of $(\mathbb{R}, +)$ under the map σ. Draw a graph of your chosen function and explain how you know it meets this criteria.

(c) Let $ker(\sigma)$ be the set of **all** elements of $(F, +)$ that map to the identity element of $(\mathbb{R}, +)$ under the map σ. Show that $ker(\sigma)$ is a subgroup of $(F, +)$.

NAME: **ASSESSMENT PROBLEMS:** LOGARITHMS AND ISOMORPHISMS (page 1 of 2)

1. Let $M_2(\mathbb{R}) = \{\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] : a, b, c, d \in \mathbb{R}\}$ and define the function $tr : (M_2(\mathbb{R}), +) \to (\mathbb{R}, +)$ to be the map $tr(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]) = a + d$. That is, tr assigns to each matrix its trace.

   (a) Show that tr is a group homomorphism. You may assume without proof that $(M_2(\mathbb{R}), +)$ and $(\mathbb{R}, +)$ are in fact groups.

   (b) Is tr an isomorphism? Why or why not?

ASSESSMENT PROBLEMS: LOGARITHMS AND ISOMORPHISMS (page 2 of 2)

2. Abina claims that there must be an isomorphism between $(\mathbb{Z}_4, +)$ and $(\mathbb{Z}_2 \times \mathbb{Z}_2, +)$ since the orders of the groups are the same.

   (a) Based on her claim, what do you think Abina understands about isomorphisms?

   (b) Explain why the two groups in question cannot be isomorphic.

   (c) What question would you ask Abina to help her understand her mistake? Why do you think your question would be helpful?

11

Summary of Research Findings

Elizabeth G. Arnold, *Colorado State University*
Elizabeth A. Burroughs, *Montana State University*
James A. M. Álvarez, *The University of Texas at Arlington*

11.1 Introduction

The META Math project has focused on adding *secondary mathematics teaching* explicitly to the list of application areas addressed in undergraduate mathematics major courses by creating teaching materials that help instructors integrate national recommendations for teacher preparation. The nine lessons at the center of this book include applications to teaching and have been implemented in a variety of settings nationwide. We have researched their use in four mathematics major courses—single variable calculus, introduction to statistics, discrete mathematics or introduction to proof, and abstract algebra—whose population of undergraduates consists of prospective secondary mathematics teachers, mathematics majors, and non-mathematics majors. Throughout the project, we employed a qualitative, multiple-case study methodology to investigate the nature of instructors' and undergraduates' experiences with connections between undergraduate mathematics and school mathematics, during and after they encountered explicit attention to particular examples of connections to teaching in the META Math lessons. We collected undergraduates' written work from the lessons and interviewed both instructors and a subset of their undergraduate students to understand their experiences teaching or learning from these lessons. These data informed multiple revisions of the lessons, which are what appear in this volume. In this chapter, we focus on highlighting some of our research findings (additional details about the research methodology can be found in Burroughs et al., 2023).

The nine META Math lessons were implemented at a diverse set of institutions and settings. The ten distinct institutions included in our case study fall roughly into four groups: two private institutions with enrollments of less than 2,000 students, three predominantly undergraduate public institutions with enrollments ranging from 4,000 to 18,000 students, four public high research activity doctoral institutions with enrollments ranging from 8,000 to 27,000 students, and one public very high research activity doctoral institution with enrollment greater than 43,000 students. For each course included in the case study, no class was larger than 50 undergraduates. For calculus and statistics, three of the classes had more than 30 undergraduates enrolled and one had fewer than 20 undergraduates enrolled. For abstract algebra and discrete mathematics or introduction to proof, two classes had more than 35 undergraduates enrolled, two had between 20 and 30 undergraduates enrolled, and two had fewer than 15 undergraduates enrolled. Undergraduates who participated in the study are from diverse backgrounds (e.g., rural, urban, historically disenfranchised in STEM), with at least three of the participating institutions carrying the designation of Hispanic Serving Institution. The variety of institutions and settings in which the materials were implemented gives us confidence that the materials have broad applicability.

11.2 Research Findings

The goal of the META Math project was to increase faculty capacity to guide prospective secondary mathematics teachers in experiencing and understanding specific examples of connections between undergraduate mathematics and the school mathematics they will teach. In our research, we defined the meaning of a *connection to teaching* and categorized such connections into five broad types (Arnold et al., 2020; Table 11.1). Details about how these connections appear in lesson materials are described in Chapter 1.

Connection to Teaching	**Description**
Content Knowledge	Undergraduates use course content in applied teaching contexts or to answer mathematical questions in the course.
Explaining Mathematical Content	Undergraduates justify mathematical procedures or theorems and use of related mathematical concepts.
Looking Back / Looking Forward	Undergraduates explain how mathematics topics are related over a span of K–12 curriculum through undergraduate mathematics.
School Student Thinking	Undergraduates evaluate the mathematics underlying a hypothetical student's work and explain what that student may understand.
Guiding School Students' Understanding	Undergraduates pose or evaluate guiding questions to help a hypothetical student understand a mathematical concept and explain how the questions may guide the student's learning.

Table 11.1. Five types of connections between undergraduate mathematics and teaching secondary mathematics

11.2.1 Instructors' Experiences with Connections to Teaching

The instructors who implemented these materials indicated that the lessons fit well in their existing undergraduate curriculum and were enriching for their undergraduates. Many of the instructors have not been formally trained to prepare prospective teachers, and in their interviews, they discussed how the materials helped them to prepare the prospective teachers in their courses and to address specific examples of connections to teaching in their instruction. From their perspectives, the lesson annotations marked as "Discuss This Connection to Teaching" bridged a gap between the undergraduate mathematics they were teaching and the school mathematics prospective teachers will teach. These annotations highlighted focal points of the lesson that were particularly important for prospective teachers, made instructors aware of how specific topics in undergraduate mathematics relate to school mathematics, and gave instructors guidance on how to directly address specific examples of connections to teaching in their courses.

Collectively, most instructors identified instances of the following three types of connections as being most valuable: *Looking Back / Looking Forward*, *School Student Thinking*, and *Guiding School Students' Understanding.* Instructors viewed the *Looking Back / Looking Forward* connections as a way to understand the "big picture" of how concepts taught in undergraduate mathematics relate to concepts taught in school mathematics. *School Student Thinking* connections were a highlight of the lessons for multiple instructors; they recognized this type of connection in teaching applications that included human beings as characters (e.g., the "Henry" problem from *The Binomial Theorem* lesson or the "Alex, Jordan, and Kelly" problem from the *Inverse Functions and Their Derivatives* lesson). Instructors described the novelty of this type of connection in mathematics content courses and a desire to integrate similar teaching applications in their future courses. Instructors also described a high level of class engagement and appreciated the kinds of discussion that emerged when their undergraduates worked on teaching applications that incorporated *School Student Thinking* connections.

In the early versions of the lessons, *Guiding School Students' Understanding* connections were framed as prompts for undergraduates to pose a set of questions to ask a hypothetical student with the intention of guiding the student's mathematical understanding. Based on input from the instructors who used these early versions of the materials, we have revised these prompts to include some sample questions to a hypothetical student and ask undergraduates to evaluate the quality of the questions. We have also included more lesson annotations about how an instructor can guide undergraduates in learning how to pose questions that would guide a hypothetical student's understanding. The versions of the lessons included in this volume make use of this new format for these kinds of teaching applications.

11.2.2 Undergraduates' Experiences with Connections to Teaching

During interviews, we asked undergraduates to describe how they thought the lessons fit into the course they were enrolled in. While undergraduates generally described that the structure of these lessons was different than other lessons their instructors taught, they indicated that the content of the lessons was suited to their course. They recognized that these lessons often encouraged a different way of learning mathematics, but, overall, from their perspectives, they learned the mathematical concepts within these lessons as well as they learned other concepts from non-META Math lessons in their course. The most noticeable differences, according to undergraduates, centered on the *Active Engagement* and the *Recognition of Mathematics as a Human Activity* principles (see Chapter 1) that we incorporated into the design of our materials. Undergraduates often highlighted the active learning components within the class activities, the opportunities to examine hypothetical student work, the opportunities to think about the *why* behind mathematical concepts, or the hands-on experiences in the statistics lessons.

Across the four content areas, undergraduates identified instances of all five types of connections to teaching that were included in the lessons. But, when we interviewed undergraduates and asked them to describe their experiences learning from the lessons, many primarily described instances related to the *Explaining Mathematical Content*, *School Student Thinking*, and *Guiding School Students' Understanding* connections. Undergraduates viewed the *Explaining Mathematical Content* connections as opportunities to explain their thinking and understand why various mathematical truths hold. From their perspectives, the mathematical tasks that embedded *Explaining Mathematical Content* connections prompted them to explain why concepts they learned in high school "are the way they are." As an example, many undergraduates recalled finding the inverse of a function using the common "switch x and y and solve for y" method that was discussed in the *Inverse Functions and Their Derivatives* lesson. Undergraduates were intrigued by the other methods presented in the lesson and wanted to understand more about these multiple mathematical perspectives.

As their instructors did, many undergraduates described the novelty of teaching applications that addressed *School Student Thinking* connections. They indicated that these lessons were the first time that they encountered a mathematical task where they were prompted to analyze and make sense of hypothetical student work. A majority of the undergraduates we interviewed found these kinds of teaching applications useful and engaging, although some undergraduates reported that instances where the content of the hypothetical student work was something they already had a deep understanding of as being less useful to their learning. Those who found these teaching applications to be useful typically described the value in seeing how others approach mathematics and common mistakes that might be made when applying certain concepts.

Undergraduates described similar experiences as their instructors with *Guiding School Students' Understanding* connections: they found teaching applications that integrated this type of connection to be the most challenging. None of the undergraduates had previous experience posing or evaluating a set of questions to guide the understanding of a hypothetical student, and many found it difficult to pose, in their own words, "good" questions. Part of the challenge was that undergraduates recognized that a good question must do more than tell the hypothetical student what they did wrong and what exactly they needed to do to fix their work. The undergraduates wanted to respect the student's work and ask a question that built from the student's understanding. The revised teaching applications addressing *Guiding School Students' Understanding* connections included in this volume include opportunities for undergraduates to examine examples of teachers' effective questions. (More resources for learning about questioning can be found in Álvarez et al., 2020, and National Council of Teachers of Mathematics, 2014, p. 41.)

The undergraduates' experiences with the five types of connections provide evidence that incorporating connections to teaching in mathematics lesson materials is a way to engage undergraduates in learning mathematics. Overall, a majority of the undergraduates we interviewed, across all four content areas, described their learning experiences with connections between undergraduate mathematics and school mathematics as beneficial. The benefits they described ranged from deepening their own mathematical or statistical understanding; making the lesson more engaging and interactive; requiring them to think more critically about a problem; giving them opportunities to see multiple perspectives; helping them understand common errors people make and serving as a warning to not make similar mistakes themselves; and helping them learn to communicate and explain mathematics. For prospective teachers, in particular, this experience provided them with opportunities to work on mathematical tasks from the stance of a secondary teacher, and they were given opportunities to see how the mathematics they are learning as undergraduates will help them when they teach school mathematics. Further, we learned that although many of the undergraduates were not interested in teaching secondary mathematics, they did envision themselves teaching in some other capacity, such as tutoring or becoming a college professor. Thus, our materials and the connections to teaching resonated in ways beyond secondary teaching. As stated by one undergraduate, these connections to teaching are applicable to everyone because "teaching is a universal skill."

11.2.3 The Human Context of Mathematics

Robust preparation of secondary teachers requires attention to how students interact with mathematics content. Teachers' interactions with students interweave their mathematical content knowledge with their capacity to respond to and guide student thinking. Prospective teachers need opportunities to engage in these practices simultaneously, to recognize their students' strengths over deficits, and to reflect on their perceptions of students as learners. Many of the teaching applications in our materials embed hypothetical learners as characters so that the mathematical content is presented in the context of guiding learners. Our analysis of undergraduate interviews highlighted a theme that emerged from their experiences learning from our lessons; namely, seeing the human context of mathematics.

In Álvarez et al. (2020), we identified two primary ways that the role of human characters in teaching applications contributed to the undergraduates' learning of mathematics: (1) scaffolding undergraduates' mathematical knowledge and (2) engaging undergraduates in practices central to teaching, such as analyzing student work, valuing student work, posing questions to help guide a student's understanding, and acknowledging that when students make errors, they are often basing their reasoning on justifications that make sense to them. We incorporated teaching applications that included human beings as characters in each content area, and we observed generally positive reactions to these teaching applications from undergraduates across all four content areas. For example, after engaging in the discrete mathematics or introduction to proof lessons, many undergraduates demonstrated that they held respect for the characters' mathematical work and what the characters understood. After engaging in the calculus lessons, undergraduates identified with the named characters and talked about them as if they were peers. After engaging in the statistics lessons, undergraduates stated how seeing different ways of thinking and doing statistics, as presented through the hypothetical students' work, helped improve their own learning, and that they learned to focus on students' assets and recognize not only the mistakes that were made but also what mathematical understanding the character demonstrated. After engaging in the abstract algebra lessons, undergraduates appreciated how these types of teaching applications encouraged communication skills among peers and helped them to identify common misconceptions related to the concepts the undergraduates were learning, such as misapplying the zero product property in non-integral domains.

We also found that undergraduates, including those who did not plan to teach secondary mathematics, were thinking about human beings while working on these applications; undergraduates' interviews and written responses suggested that they imagined the human characters to be real people with real mathematical thoughts. Some instructors, echoing the undergraduates' perspectives, stated that these types of teaching applications offered undergraduates a different and unique way to engage with mathematical content compared to mathematical tasks devoid of any human beings as characters.

We view these findings as confirmation that including teaching applications that involve human beings as characters in mathematics content courses is a useful approach to give undergraduates opportunities to engage in practices that require both mathematical expertise and skills for probing student thinking or finding meaning in learners' perspectives. With these applications, undergraduates are simultaneously learning about mathematics and learning about teaching mathematics in mathematics content courses. These lessons provide a way for instructors and all undergraduates, regardless of whether or not they are prospective secondary teachers, to grow in their understanding of mathematics as a human endeavor.

11.3 Summary

In all four content areas of single variable calculus, introduction to statistics, discrete mathematics or introduction to proof, and abstract algebra, we found that directly addressing connections to teaching through teaching applications in these courses did not detract, in the eyes of undergraduates or of instructors, from the mathematical content that was learned by all undergraduates in the course. Instructors found that the content of the lessons related to content from school mathematics while still attending to the rigor and goals of their undergraduate mathematics courses. In fact, the focus on connections to teaching led to some undergraduates self-reporting that they were encouraged to think more deeply about the mathematical content. According to many of the undergraduates we interviewed, teaching is a universal skill, and there's no better way to learn the material than when you are asked to analyze student work, explain mathematical concepts to others, and ask questions that build from someone's mathematical work to help guide and deepen their mathematical understanding.

Overall, we found that the META Math lessons can enrich undergraduate mathematics courses by providing powerful mathematical learning for all undergraduates while attending to the specific needs of prospective secondary mathematics teachers. We recommend the inclusion of these five types of connections within teaching applications across undergraduate courses as a means for mathematicians to make explicit connections between the content of undergraduate mathematics courses and the content of school mathematics.

11.4 References

[1] Álvarez, J. A., Arnold, E. G., Burroughs, E. A., Fulton, E. W., & Kercher, A. (2020). The design of tasks that address applications to teaching secondary mathematics for use in undergraduate mathematics courses. *The Journal of Mathematical Behavior*, *60*, 100814.

[2] Burroughs, E. A., Arnold, E. G., Álvarez, J. A. M., Kercher, A., Tremaine, R., Fulton, E., & Turner, K. (2023). Encountering ideas about teaching and learning mathematics in undergraduate mathematics courses. *ZDM–Mathematics Education*. `doi.org/10.1007/s11858-022-01454-3`.

[3] National Council of Teachers of Mathematics. (2014). Principles to actions: Ensuring mathematical success for all. NCTM.

About the Editors

James A. M. Álvarez earned his BS in mathematics and physics from Texas A & M University-Commerce and his PhD in mathematics at The University of Texas at Austin. As a professor in the Department of Mathematics at The University of Texas at Arlington, he spent more than two decades directing the graduate program for secondary mathematics teachers, designing courses and course materials at the graduate and undergraduate levels aimed at enhancing engagement in mathematical problem solving and development of mathematical knowledge for teaching (MKT). His research and professional interests have focused on the mathematical preparation of secondary mathematics teachers, mathematical problem solving, and efforts for increasing student success in gateway mathematics courses. He was a lead writer for the MAA's *Instructional Practices Guide* (2018) and has served as Chair of the Special Interest Group of the MAA on MKT. He is a past recipient of the E. Glenadine Gibb Achievement Award by the Texas Council of Teachers of Mathematics for his contributions to the improvement of mathematics education at the state and national levels.

Elizabeth G. Arnold earned her BA and MS in mathematics from Cal Poly Humboldt and her MS in statistics and PhD in mathematics education from the Department of Mathematical Sciences at Montana State University. Her research centers on the mathematical preparation and development of pre-service and in-service K–12 mathematics teachers, with a focus on mathematical knowledge for teaching secondary mathematics, mathematical modeling, and teaching and learning statistics. Throughout her career, she has taught undergraduate mathematics content courses whose population consists of general mathematics majors and pre-service teachers, content courses specifically designed for pre-service teachers, and methods of teaching courses. She is also a co-author on *Becoming a Teacher of Mathematical Modeling, Grades K–5*, and *Becoming a Teacher of Mathematical Modeling, Grades 6–12*, both published in 2021 by NCTM.

Elizabeth A. Burroughs earned her BA in mathematics and English from the University of North Carolina and her MA and PhD in mathematics from the University of New Mexico. As a professor at Montana State University in the Department of Mathematical Sciences, she has taught a variety of undergraduate mathematics courses and graduate mathematics education courses. A former high school mathematics teacher, Beth has focused her research on the preparation and professional development of mathematics teachers at all levels. She was a lead writer for the MAA's *Instructional Practices Guide* (2018) and a member of the writing team for the Association of Mathematics Teacher Educators' *Standards for Preparing Teachers of Mathematics* (2017). She is a co-author of *Becoming a Teacher of Mathematical Modeling, Grades K–5*, and *Becoming a Teacher of Mathematical Modeling, Grades 6–12*, both published in 2021 by NCTM.