MSRI Mathematical Sciences Research Institute

Final Report
on the
Mathematical Sciences Research Institute
2015 Undergraduate Program (MSRI-UP)
supported by
NSA Grant
H98230-15-1-0039
May 2016

2015 Mathematical Sciences Research Institute – Undergraduate Program (MSRI-UP)

Final Report

Table of Contents

2015 Mathematical Sciences Research Institute – Undergraduate Program (MSRI-UP)
Final Report

1. Introduction

MSRI-UP continued in 2015, June 11 through July 24, with 17 students studying and researching Geometric Combinatorics Motivated by the Social Sciences. The summer program was staffed by lead director Duane Cooper, research leader Francis Su, postdoctoral fellow Mutiara Sondjaja, and graduate students Daniel Eckhardt and Lyda Urresta.

The MSRI-UP is a comprehensive program for undergraduates that aims to increase the number of students from underrepresented groups in mathematics graduate programs. MSRI-UP includes summer research opportunities, mentoring, workshops on the graduate school application process, and follow-up support.

The primary objective of the MSRI-UP is to identify talented students, especially those from underrepresented groups, who are interested in mathematics and make available to them meaningful research opportunities, the necessary skills and knowledge to participate in successful collaborations, and a community of academic peers and mentors who can advise, encourage, and support them through a successful graduate program. We achieve this through an intensive six-week summer program of mathematics research and other activities, along with maintenance of relationships with participating students for years beyond the summer program.

The MSRI-UP is coordinated by an experienced team of five directors, consisting in 2015 of Professors Federico Ardila of San Francisco State University, Duane Cooper of Morehouse College, Herbert Medina of Loyola Marymount University, Ivelisse Rubio of the University of Puerto Rico, Río Piedras, and Suzanne Weekes of Worcester Polytechnic Institute. The directors collaborate ongoingly and annually rotate direct leadership of the program. The program is supported by the leadership and staff of the Mathematical Sciences Research Institute in Berkeley, site of each summer's six-week program.

During the 2007-2014 summers, 134 students conducted 45 small group research projects in Computational Mathematics, Experimental Mathematics, Coding Theory, Elliptic Curves and Applications, Mathematical Finance, Enumerative Combinatorics, Algebraic Combinatorics, and Arithmetic Aspects of Elementary Functions. Most MSRI-UP participants who have graduated college proceeded to enter graduate programs in the mathematical sciences. In 2015, 17 undergraduates completed MSRI-UP, having learned about and engaged in research on Geometric Combinatorics motivated by the Social Sciences, led by Professor Francis Edward Su of Harvey Mudd College, who guided the students in conjunction with a postdoctoral fellow and two graduate students—carefully chosen role models—who contributed to the undergraduates' academic, personal, and professional development. In 2016, the MSRI-UP will continue with 18 undergraduates conducting research projects in Sandpile Groups, led by Professor Luis Garcia-Puente of Sam Houston State University.

2. Funding Information

The funding available to administer the 2015 MSRI-UP is summarized as follows:

1. National Security Agency	\$124,993[1]
2. The National Science Foundation	\$120,721[2]

In addition, the MSRI provided much additional support by allowing MSRI-UP to use classrooms, offices, and computers; by facilitating transportation; and providing administrative assistance.

3. Recruitment, Application and Admissions Procedures

The directors began recruiting for the 2015 MSRI-UP at the annual conference of the Society of Chicanos and Native Americans in Science (SACNAS) in Los Angeles in fall of 2014. Directors present distributed fliers and talked to dozens of students and faculty about the program. The MSRI-UP home page also provided information about and applications for the program. Recruitment of students also occurred that fall at the Math Alliance's Field of Dreams conference and then in January at the Joint Mathematics Meetings, specifically at the MAA's minority chairs breakfast, sessions of the National Association of Mathematicians (NAM), and the MAA Student Poster Session.

---

[1] Grant number H98230-15-1-0039.

The on-line application, which resided on the mathprograms.org site, consisted of four items: a completed student application form, transcripts, a statement of interest, and a letter of recommendation. The 2015 MSRI-UP received more than 400 applications.

Directors Ardila, Cooper, Medina, Rubio, and Weekes, reviewed each application and evaluated it using four criteria: 1.) the student's grades in mathematics courses; 2.) the student's mathematical background; 3.) the statement of interest; and 4.) the letter of recommendation. Based on these criteria, directors gave each applicant a score between 0 and 10. The scores were summed and averaged, and this score served as the initial measure for evaluating each applicant. The directors then proceeded to discuss individual applications and eventually reached a consensus on the eighteen[3] admittees for the program. Two students declined and one more was determined to be ineligible; the directors replaced them with alternate candidates.

4. Summary of Participant Demographics

Table 1 details some demographic information of the eighteen MSRI-UP students who began the program. The student participants were diverse by race and ethnicity, as well as by the types and geographic regions of their undergraduate institutions. The co-directors paid special attention during the selection process to attain racial and ethnic diversity and gender balance. Achieving this type of diversity and gender balance is important for creating the academic and research environment explained below and for achieving one of the MSRI-UP objectives.

5. Housing and Lodging for the Students

The students were housed in Stern Hall dormitory at the University of California, Berkeley. On weekdays, lunch was served at MSRI. The lunches at MSRI were shared with graduate students, faculty, and teachers participating in other MSRI summer programs. This allowed students to meet mathematicians at different stages of professional development. The students and the program's graduate students had breakfast and dinner at the dining facilities in the dormitories. On occasion, meals were shared by MSRI students and senior staff. Sharing meals with their

---

[2] Research Experience for Undergraduates (REU) program grant number DMS-1156499.

[3] The program is designed for eighteen students and indeed eighteen students began the program. One student decided to leave the program during Week 1.

MSRI-UP peers promoted mathematical discussions and enhanced the collaborative and intellectual environment of MSRI-UP.[4]

Table 1

2015 Mathematical Sciences Research Institute (MSRI-UP)

Student Data

Undergraduate Institution and State/Country		**Gender**	
University of Arizona	1	Male	11
Loyola Marymount University, California	1	Female	7
Pomona College, California	1		
San Francisco State University, California	1		
University of California, Los Angeles	1		
Clark Atlanta University, Georgia	1	**Major**	
Morehouse College, Georgia	2	Mathematics	18
Purdue University, Indiana	1	Economics (double major w/math)	1
University of Massachusetts	1	Computer Sci. (double w/math)	1
University of Michigan, Flint	1		
University of North Carolina	1		
Swarthmore College, Pennsylvania	1		
University of Tennessee	1	**Ethnicity**	
James Madison University, Virginia	1	Black/African American	7
University of Puerto Rico, Mayagüez	1	Hispanic/Latino	9
University of Puerto Rico, Río Piedras	1	Asian	1
Universidad de los Andes, Colombia	1	Pacific Islander	1

6. Pre-Research Seminar

During the first full week and part of the second week of MSRI-UP, students participated in a pre-research seminar consisting of lectures, tutorials, and problem-solving sessions. Professor Su planned the seminar so that he could familiarize students with motivation and fundamental concepts of the field of geometric combinatorics and also the main techniques that they would need to work on their research topics.

The pre-research phase was conducted in the Baker Boardroom, an excellent classroom-type facility at MSRI. Attached, as Appendix B to this report, is the program calendar, describing the structure of these first weeks and all weeks of the program.

Students engaged homework assignments a bit on site at the MSRI but mostly in evenings in the residence hall with the programs graduate student assistants present as an

---

available resource. Frequently, undergraduate groups were assigned and shuffled so that the students and the program staff could identify groups that collaborated well or did not.

7. Research Projects, Technical Reports, Posters, and MSRI Student Presentations

The focus of MSRI-UP is undergraduate research. From the second week till the end of the program, each student worked exclusively on an undergraduate research project in the field of geometric combinatorics that was carefully designed by the research leader, Prof. Su. There were five groups of three students plus one pair. Students wrote technical reports, prepared posters summarizing their work for use during the academic year ahead, and presented the results of their research in the MSRI-UP Student Colloquium the final Friday of the program.

At the end of Week 1 of the program, students received descriptions of their possible research projects. The students did preliminary reading and literature searches on the project topics, and they were requested to rank their top project choices. However, program staff composed the research teams, satisfying student preferences as much as possible while paying attention to interpersonal dynamics that had been observed during the pre-research seminar.

During the research phase of MSRI-UP, students worked in the offices assigned to them at MSRI. Each research team was assigned a support person from the academic staff of the program. Professor Su oversaw all the work of all six groups. Postdoctoral fellow Sondjaja supervised two research teams, as did each of the two graduate assistants. The undergraduates met with their support person for several hours each day, and each team met periodically with Professor Su to update him and receive guidance.

During the program, MSRI-UP participants were introduced to some of the techniques that are used while conducting successful research in the mathematical sciences. Indeed, students learned to work as part of a research team, develop an effective faculty advisor-student relationship, use computer software as tools, use the Internet as a resource, prepare and deliver an oral presentation, write a mathematics paper (technical report), and use LaTeX, including the Beamer package for presentations.

The outcome of all the students' hard work and dedication (and of course staff support) resulted in six technical reports and an equal number of oral presentations in the Student Colloquium session. The program of final research presentations, including abstracts of all six presentataions, and one of the six technical reports are included as Appendices C and D to this report. The tech. report describes a result that the team and staff have submitted for publication in a mathematics research journal.

8. Evaluation of Student Work

Close interaction with students allowed the academic staff to give individuals feedback on their work throughout the program. During the pre-research seminar, homework assignments were reviewed by the academic staff and critiqued by peers as students presented solutions to problems daily. During the research phase, each of the six research teams held daily meetings, for which students often prepared oral or written progress reports. Dr. Sondjaja, Mr. Eckhardt, or Ms. Urresta, who were serving as support for the research groups, were present at the progress meetings, as was Dr. Su frequently, for guidance.

Indeed, the program's academic staff gave students written feedback on drafts of their technical reports so that the finished product would be formatted as a professional publication. The academic staff and Dr. Cooper met with the teams for interim oral presentations during the weeks prior to the final presentations, helping the research teams with elements of substance and style in their deliveries.

Program staff met weekly to discuss the progress of and any concerns about individual and teams of students. The staff also at the program's end to assess the undergraduate performance at the program.

9. Colloquium Series

The 2015 MSRI-UP hosted five mathematicians for a colloquium series: Taleo Mayo, Princeton University; Gina-Maria Pomann, Duke University; Federico Ardila, San Francisco State University; Bobby Wilson, University of Chicago; and Khalilah Beal, University of California, Berkeley. The colloquium series stimulated the mathematical interests of the students and gave them a glimpse of current mathematical research. In addition to this, the speakers provided the students with additional role models and expanded their network of mentors. In particular, each speaker was asked to share personal stories, "offering reflections on your own journey to mathematics and advice for them to consider on their journeys." The speakers' schedules were arranged to maximize opportunities for them to engage the undergraduates in informal conversation, and many students took advantage of the opportunity to listen, ask, and learn, including joining the undergraduates on an excursion at week's end.

A special feature of the 2015 colloquium series is that Drs. Mayo, Pomann, and Wilson are three of MSRI-UP's first four alumni to earn the Ph.D. Also, graduate assistants Eckhardt and Urresta are MSRI-UP alumni. We delighted in having this significant presence of former program participants serving in advanced roles during the summer. The intent was to foster a feeling in the students that they are part of something special, this MSRI-UP family, and that much that the directors hope for them is, indeed, achievable.

10. Graduate School Workshops

Dr. Colette Patt, Director of Diversity Programs in the Physical Sciences at the University of California, Berkeley visited MSRI-UP and gave a workshop on applying for graduate school and attaining fellowship funding for graduate school. The workshop addressed questions/issues such as the significant differences between masters and doctoral programs, the funding opportunities available for most graduate programs, and the benefits of obtaining a graduate degree. In addition to this basic information, Dr. Patt also presented successful techniques for applying to graduate school. She discussed the elements that constitute a good statement of purpose, the types of professors from whom one should seek letters of recommendation, and successful techniques for addressing not-so-stellar semesters. Dr. Patt also discussed successful strategies for compiling a winning national fellowship application. She also provided the students with related written material. Her presentation and the information she provided were well received by the program students.

11. Additional Workshops & Panels

The program held workshops that were devoted to the development of skills that are important to every mathematician. Two were devoted to learning LaTeX, the typesetting program most widely used by mathematicians. The first workshop was an introduction arranged by MSRI-UP Dr. Cooper during the program's first week, and the second workshop, introducing some advanced tools and online resources for team editing, was arranged by Dr. Sondjaja. These skills were needed as MSRI-UP students prepared their technical report and transparencies using

LaTeX, gave an end-of-program oral presentation using Beamer presentation slides, and used LaTeX to prepare their research posters.

The program presented two panel discussions by current graduate students for the undergraduates. The panels were moderated by the program's two graduate student assistants. The first graduate student panel featured graduate students who were on site for one of the MSRI's summer graduate student workshops; the second featured mathematics and statistics doctoral students at various stages from the University of California, Berkeley. In both cases panelists spoke about their graduate-school experiences with the aim of "demystifying graduate school in mathematics." They provided insights on selecting a graduate department and succeeding in it.

A career panel featured four scholars with mathematical sciences Ph.D.s engaged in careers other than mathematical research. Panelists were a Dr. Jamylle Carter, Professor of Mathematics at Diablo Valley College, a community college, Eric Hsu, Professor of Mathematics at San Francisco State University, where he has extensive involvement in grade school math education, Dr. Tanya Moore, Education and Youth Services Specialist for the City of Berkeley, and Dr. Luis Serrrano, Software Engineer at Google. The guests drew many questions both during the panel and afterwards, over tea, where small group discussions with the students continued.

12. Recreational/Cultural Activities

In addition to all the academic activities described above, MSRI-UP students were treated to several recreational activities. These included visits to nearby San Francisco and Santa Cruz, an Oakland Athletics baseball game, a visit to the Exploratorium science center, and a walking tour of Berkeley. These carefully-planned recreational and cultural activities were essential to MSRI-UP's success, as they gave students the opportunity to put mathematics aside for a few hours so that they could come back later to their work with renewed vigor: They also helped to build the MSRI-UP mentored community, as all staff participated in the activities with the students.

13. Program Evaluation During MSRI-UP

Informal formative evaluation in the program started the first day of the program through conversations with students and staff. Frequently during the program, Professor Cooper met individually with each one of the students and staff of the program, conducting extensive discussions with Professor Su to learn about and share opinions regarding the research component. During the meetings with staff and students, the lead director had the opportunity to have more close contact with the students and staff, to listen to specific concerns, and to provide individual mentoring to the students. The staff's close interaction—especially the graduate assistants—with the students enabled them to gather informal feedback that also led to adjustments to improve the program.

The program staff had regular weekly meetings to discuss individual and group progress, and they held several impromptu lunch or other daytime meetings as issues arose that would benefit from immediate discussion and resolution. At the final staff meeting, individual student performances were discussed at length.

14. End-of-Program Evaluation

Each MSRI-UP student was required to complete a comprehensive, end-of-program, online evaluation. Indeed, the evaluation was an online instrument shared by numerous summer REU programs; a couple of additional evaluation prompts were added specifically for the MSRI-UP students. The evaluation form had both year-to-year formative evaluation questions designed for soliciting feedback in order to improve future institutes and summative-evaluation questions to measure the effectiveness of MSRI-UP in accomplishing the program objectives. The quantitative results of the end-of-program evaluation are provided in Appendix A.

Post-program conversations between the MSRI-UP staff and the Directors indicated that the staff felt that the institute was successful in accomplishing its objectives.

15. Post-Summer Conferences

MSRI-UP has a substantial post-summer component. Students are provided funding to attend academic conferences to present their research. In addition, each year the onsite director keeps students informed of conference opportunities and funding sources for attending such conferences.

Indeed, 15 of the 17 students who completed the program's, representing all 6 research project teams, presented their research at the 2015 SACNAS National Conference in Maryland in October, 2015. A majority of the students, representing all 6 research project teams, will present their work during the Student Poster Session at the Joint Mathematics Meetings in Seattle, Washington, in January, 2016. Students have also presented work at their home institutions and regional mathematics conferences.

16. Evidence Suggesting Long-Term Success of Program

MSRI-UP boasts 134 program alumni from its first eight summers, most of whom have graduated from college. Many students have continued for advanced study, as our program hopes many will choose to do. Eight of these studens have earned Ph.D.s, most in the mathematical sciences, all in STEM disciplines. They are Drs. Natalie Durgin, Talea Mayo, and Gina-Maria Pomann from the 2007 MSRI-UP and Drs. Natasha Cayco, Gerard Koffi, Nathan Kallus, Marcos Ortiz, and Bobby Wilson from the 2008 MSRI-UP. Three of them are tenure-track faculty members, 3 are postdoctoral researchers, 1 is a senior researcher in an academic lab, and 1 is a data scientist. Of the 121 other MSRI-UP students who have earned their bachelor's degrees, 62 are currently in a Ph.D. program, and 32 have received or are currently pursuing a master's degree.

17. Conclusion

Like the eight summers that preceded it, reviews of the MSRI-UP from its students, staff, and guests have been overwhelming positive. The program is certainly perceived as an overall success, though the real fruit—that of achieving the program's primary goal *to increase the number of graduate degrees in the mathematical sciences, especially doctorates, earned by U.S. citizens and permanent residents by cultivating heretofore untapped mathematical talent*—is starting to be realized.

The long-term data that will confirm that the MSRI-UP objectives contribute towards the goal of increasing the number of Latinos/Chicanos, African-American and Native Americans earning graduate degrees in the mathematical sciences will unfold over several years ahead. The directors are committed to maintain the relationships developed with each cohort of students in the program in order to monitor and collect data on the MSRI-UP students' academic progress and, whenever possible, to provide them with additional academic opportunities.

INSTRUMENT ANALYSIS

Sign out

My home page

Edit my profile

Privacy policy

Legal statement

User forum

Tips and tricks

Help

Instrument dashboard

On this page, you can view and download a summary of results from one or more SALG instruments. You can also view results across instruments.

Results displayed for the following instrument:

ID	Open Close	Instructor	Course	Semester	Description	N	
70761	Sat Jul 25, 2015 Tue Jul 28, 2015	Duane Cooper	MSRI-UP	2015	2015 Mathematical Sciences Research Institute Undergraduate Program	17	download

If you'd like to aggregate data across instruments, you can add another instrument to this analysis.

Cross-tabulate questions ?

Frequency distributions of scale results

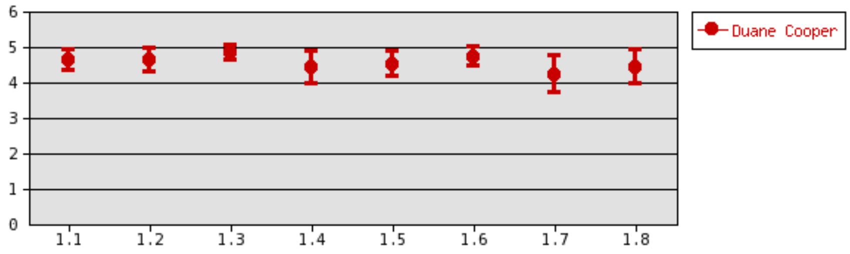
The table below lists the percentage of students responding in each category, along with the mean and number of responses for that item. If you'd like an more detailed analysis, click on the 'details' link to the right of that item.

Gains in THINKING AND WORKING LIKE A MATHEMATICIAN OR A STEM PROFESSIONAL: APPLICATION OF KNOWLEDGE TO RESEARCH WORK.

1. How much did you GAIN in the following areas as a result of your most recent research experience?	1:no gains	2:a little gain	3:moderate gain	4:good gain	5:great gain	9:not applicable	Mean	N	
1.1 Analyzing data for patterns.	0%	0%	6%	24%	65%	6%	4.6	16	details
1.2 Figuring out the next step in a research project.	0%	0%	12%	18%	71%	0%	4.6	17	details
1.3 Problem-solving in general.	0%	0%	0%	24%	76%	0%	4.8	17	details
1.4 Formulating a research question that could be answered with data.	0%	6%	6%	24%	53%	12%	4.4	15	details
1.5 Identifying limitations of research methods and designs.	0%	0%	12%	24%	65%	0%	4.5	17	details
1.6 Understanding the theory and concepts guiding my research project.	0%	0%	6%	18%	76%	0%	4.7	17	details
1.7 Understanding the connections among mathematical disciplines.	0%	12%	12%	18%	59%	0%	4.2	17	details
1.8 Understanding the relevance of research to my coursework.	0%	6%	18%	12%	65%	0%	4.4	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.



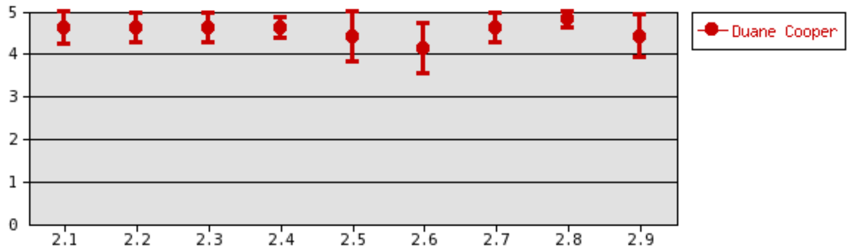
PERSONAL GAINS RELATED TO RESEARCH WORK

2. How much did you GAIN in the following areas as a result of your most recent research experience?	1:no gains	2:a little gain	3:moderate gain	4:good gain	5:great gain	9:not applicable	Mean	N	
2.1 Confidence in my ability to do research.	0%	6%	0%	24%	71%	0%	4.6	17	details
2.2 Confidence in my ability to contribute to mathematics.	0%	0%	12%	18%	71%	0%	4.6	17	details
2.3 Comfort in discussing mathematical concepts with	0%	0%	12%	18%	71%	0%	4.6	17	details

others.									
2.4 Comfort in working collaboratively with others.	0%	0%	0%	35%	65%	0%	4.6	17	details
2.5 Confidence in my ability to do well in future math courses.	6%	6%	6%	6%	76%	0%	4.4	17	details
2.6 Ability to work independently.	6%	0%	24%	12%	53%	6%	4.1	16	details
2.7 Developing patience with the slow pace of research.	0%	0%	12%	12%	76%	0%	4.6	17	details
2.8 Understanding what everyday research work is like.	0%	0%	0%	18%	82%	0%	4.8	17	details
2.9 Taking greater care in conducting procedures in the lab or field.	6%	0%	0%	29%	59%	6%	4.4	16	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

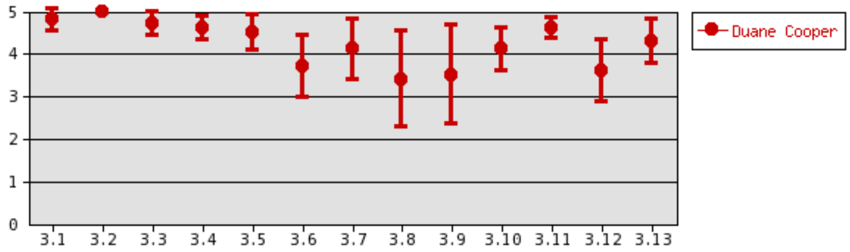


Gains in SKILLS

3. How much did you GAIN in the following areas as a result of your most recent research experience?	1:no gains	2:a little gain	3:moderate gain	4:good gain	5:great gain	9:not applicable	Mean	N	
3.1 Writing mathematical reports or papers.	0%	0%	6%	6%	88%	0%	4.8	17	details
3.2 Making oral presentations.	0%	0%	0%	0%	100%	0%	5.0	17	details
3.3 Defending an argument when asked questions.	0%	0%	6%	18%	76%	0%	4.7	17	details
3.4 Explaining my project to people outside my field.	0%	0%	6%	24%	71%	0%	4.6	17	details
3.5 Preparing a mathematics poster.	0%	0%	24%	0%	76%	0%	4.5	17	details
3.6 Keeping a detailed lab notebook.	0%	18%	18%	6%	29%	29%	3.7	12	details
3.7 Conducting observations in the lab or field.	6%	6%	12%	6%	53%	18%	4.1	14	details
3.8 Using statistics to analyze data.	12%	6%	6%	6%	24%	47%	3.4	9	details
3.9 Calibrating instruments needed for measurement.	12%	0%	6%	12%	18%	53%	3.5	8	details
3.10 Working with computers.	0%	6%	24%	24%	41%	6%	4.1	16	details
3.11 Understanding journal articles.	0%	0%	0%	35%	65%	0%	4.6	17	details
3.12 Conducting database or internet searches.	12%	12%	18%	12%	41%	6%	3.6	16	details
3.13 Managing my time.	6%	0%	12%	24%	59%	0%	4.3	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.



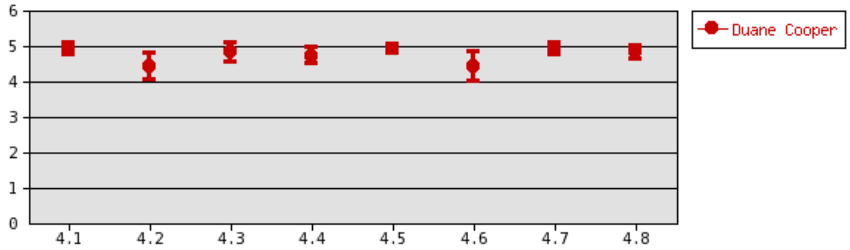
The following questions ask about your overall research experience and about any changes in your attitudes or behaviors as a researcher.

4. During your research experience HOW MUCH did you:	1:none	2:a little	3:some	4:a fair amount	5:a great deal	9:not applicable	Mean	N	
4.1 Engage in real-world mathematics research	0%	0%	0%	12%	88%	0%	4.9	17	details

4.2 Feel like a mathematician.	0%	0%	18%	29%	53%	0%	4.4	17	details
4.3 Think creatively about the project.	0%	0%	6%	12%	82%	0%	4.8	17	details
4.4 Try out new ideas or procedures on your own.	0%	0%	0%	29%	71%	0%	4.7	17	details
4.5 Feel responsible for the project.	0%	0%	0%	6%	94%	0%	4.9	17	details
4.6 Work extra hours because you were excited about the research.	0%	6%	6%	29%	59%	0%	4.4	17	details
4.7 Interact with mathematicians from outside your school.	0%	0%	0%	12%	88%	0%	4.9	17	details
4.8 Feel a part of a mathematics community.	0%	0%	0%	18%	82%	0%	4.8	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.



.

5. What year are you in college?	1:Freshman/rising sophomore	2:Sophomore/rising junior	3:Junior/rising senior.	4:Senior	5:Other :	Mean	N	
5.1 I am a:	0%	18%	65%	18%	0%	--	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

.

6. GPA	1:3.5 - 4.0	2:3.0 - 3.49	3:2.5 - 2.99	4:2.0 - 2.49	5:Below 2.0	6:Don't know	Mean	N	
6.1 What is your GPA?	65%	29%	0%	6%	0%	0%	--	17	details
6.2 What is your GPA in your math courses?	59%	35%	0%	6%	0%	0%	--	17	details

Summary of scale results

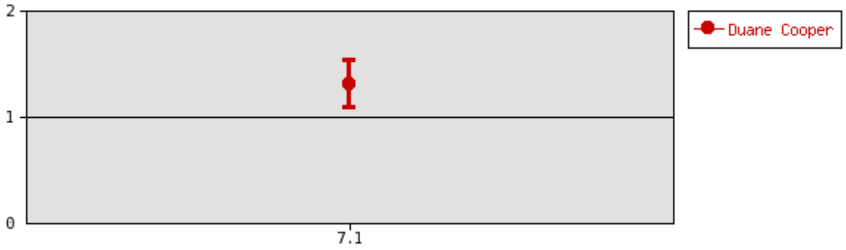
The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

.

7. Summer research	1:Never participated	2:1 summer	3:2 summers	4:3 summers	:	:	Mean	N	
7.1 How many times have you participated in SUMMER research, excluding this one?	71%	29%	0%	0%			1.3	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

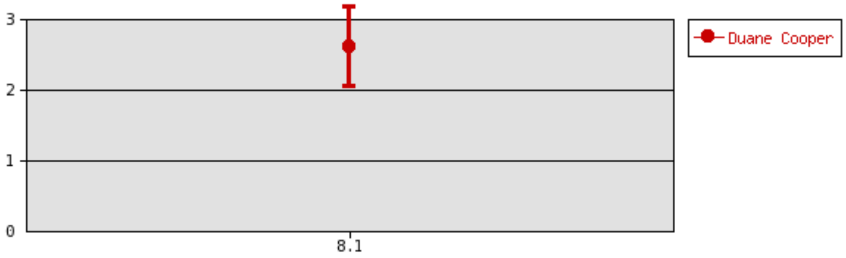


.

8. Stipend	1:Not at all important	2:Slightly important	3:Important	4:Very important	:	:	Mean	N	
8.1 How important was the stipend or money you were paid in allowing you to do research?	24%	24%	24%	29%			2.6	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.



.

9. Gender	1:Male	2:Female	3:Decline to answer	:	:	:	Mean	N	
9.1 What is your gender?	65%	35%	0%				--	17	details
9.2 Other	Enter codes for text answers						--	--	

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

.

10. Race	1:Native American	2:Asian American	3:African American	4:White	5:Other	6:Decline to answer	Mean	N	
10.1 What is your race?	0%	12%	41%	24%	18%	6%	--	17	details
10.2 Other	Enter codes for text answers						--	2	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

.

11. Ethnicity	1:Hispanic	2:Non-Hispanic	3:Decline to answer	: : :	Mean	N	
11.1 What is your ethnicity?	47%	53%	0%		--	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

Citizenship

12. What is your citizenship status?	1:US Citizen	2:US Permanent Resident	3:Decline to Answer	4:Other	: :	Mean	N	
12.1 Citizenship	94%	6%	0%	0%		--	17	details
12.2 Other	Enter codes for text answers					--	--	

Summary of scale results

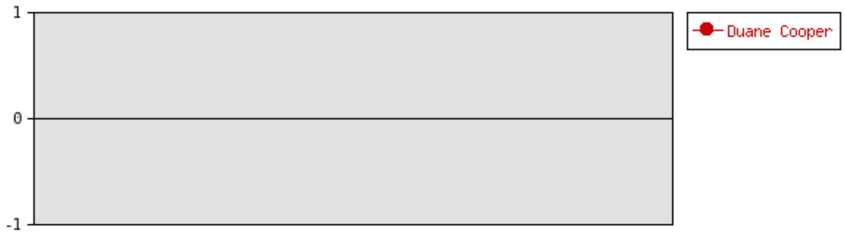
The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

.

13. Other demographics	1:Yes	2:No	3:Decline to answer	: : :	Mean	N	
13.1 Are you a first generation college student?	35%	65%	0%		--	17	details
13.2 Are you a student at a 2-year community college?	0%	100%	0%		--	17	details
13.3 Do you have a disability?	0%	100%	0%		--	17	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.

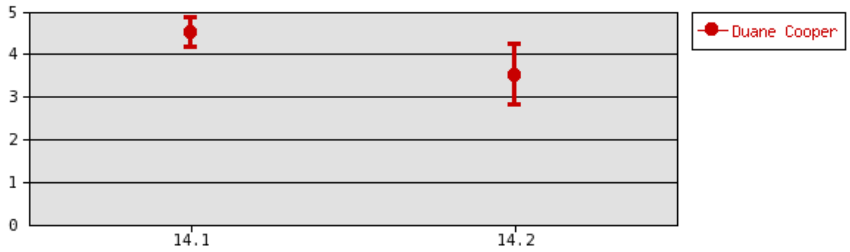


Research experience

14. Compared to your intentions BEFORE doing research, HOW LIKELY ARE YOU NOW to:	1:not more likely	2:a little more likely	3:somewhat more likely	4:much more likely	5:extremely more likely	9:not applicable	Mean	N	
14.1 enroll in a Ph.D. program in science, mathematics or engineering?	0%	0%	12%	29%	59%	0%	4.5	17	details
14.2 enroll in a masters program in science, mathematics or engineering?	18%	0%	12%	41%	18%	12%	3.5	15	details
14.3 Other. Please state your intended degree and, compared to your intentions BEFORE doing research, HOW LIKELY YOU ARE NOW to enroll in a graduate program leading to an advanced degree.	Enter codes for text answers						--	10	details

Summary of scale results

The graphic below lists the mean and confidence interval (±3 times the standard error) for each item.



MSRI-UP: exposure, development, community

15. (FYI, each reply is capped at 2000 characters.) Please take time to write a bit about these aspects of the 2015 MSRI-UP and what they meant to you--what you found informative, inspiring, valuable, worthwhile, useful, useless, fun, special, etc.--mathematically or otherwise and individually or collectively (e.g., you can address the impact of a specific visitor or the impact of the set of visitors).		Mean	N	
15.1 Colloquia (Drs. Talea Mayo, Gina-Maria Pomann, Federico Ardila, Bobby Wilson, Khalilah Beal)	Enter codes for text answers	--	17	details
15.2 Graduate School and Fellowship Workshop (Dr. Colette Patt), July 10	Enter codes for text answers	--	17	details
15.3 Graduate Student Panels, Parts 1 and 2, June 26 and July 2	Enter codes for text answers	--	17	details
15.4 Career Panel, July 15	Enter codes for text answers	--	17	details
15.5 Saturday Excursions (and one Friday)	Enter codes for text answers	--	17	details

MSRI-UP: research project, background

16. Please reflect on these aspects of your MSRI-UP research and pre-research experiences.		Mean	N	
16.1 Are you satisfied with your project topic, or do you wish you had requested or been assigned to a different problem? Please explain.	Enter codes for text answers	--	17	details
16.2 Are you satisfied with your project teammates, or do you wish you had requested or been assigned different people with whom to work? Please explain.	Enter codes for text answers	--	17	details

16.3 Please *compare*, positively or negatively, your *team* research to the work you may have produced and the experience you think you would have had working *by yourself* on the project.	Enter codes for text answers	--	17	details
16.4 In what ways was the advisement and guidance you received most effective or least effective?	Enter codes for text answers	--	17	details
16.5 Was the pre-research background phase (Week 1 plus) of the program "just right", or do you have suggestions that would have made it more effective to you? (e.g., move faster or slower, assign more homework or less homework, have more lectures or fewer lectures, etc.)?	Enter codes for text answers	--	17	details

Other MSRI-UP-specific prompts

17. Lastly, please answer Dr. Cooper's pet questions.		Mean	N	
17.1 *Mathematically,* what was the most valuable or memorable experience or incident during MSRI-UP to you?	Enter codes for text answers	--	17	details
17.2 *Otherwise,* what was the most valuable or memorable experience or incident during MSRI-UP to you?	Enter codes for text answers	--	17	details
17.3 Is there anything regarding any aspect of the program that you wish you had known *before* June 13 (arrival date)?	Enter codes for text answers	--	17	details
17.4 What advice would you give to a friend *applying* to the 2016 MSRI-UP and to a friend *accepted* to our program?	Enter codes for text answers	--	17	details

2015 MSRI-UP Calendar, Week 1

	Sunday 14-June	Monday 15-June	Tuesday 16-June	Wednesday 17-June	Thursday 18-June	Friday 19-June	Saturda 20-Jun
AM		8:40, "Hill Line" Shuttle	8:10, 8:40, 9:10,	8:10, 8:40, Shuttle	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	
AM		9-10, Lecture	Shuttle to MSRI	9-10, Lecture	Shuttle to MSRI	Shuttle to MSRI	
AM		10, Welcome to MSRI;	9:30-10:30, Lecture		9:30-10:30, Lecture	9:30-12, LaTeX Workshop	
0 AM		ID badges, keys, etc.		10:15-11:15, Lecture			10:00 - ur
0 AM		11-12, Lecture	11-12, Lecture	11:30 Begin hike for	11-12, Lecture		
n		Lunch	Lunch	12: MSRI Barbecue	Lunch	Lunch	San
PM		12:45, 1:25, 2:05, Library					Francisc
PM		tours, Paperwork,	1:30-3, Prob. Session/		1:30-3, Prob. Session/		
PM		Small group meetings	Groups/Presentations		Groups/Presentations	2-3:30, Colloquium,	
PM		2:45-3:30, Prob's, Gr'ps	3, Tea	3-5, Problem Session/	3, Tea	Talea Mayo	
PM		3:30, Tea	3:30-5, Prob. Session/	Group Work/Student	3:30-5, Prob. Session/	3:30, Tea	
PM		4-5, Problems, Groups	Groups/Presentations	Presentations	Groups/Presentations	Shuttle to Mining Circle	
PM	MSRI-UP	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	3:55, 4:25, 4:55,	
PM	Orientation	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	

2015 MSRI-UP Calendar, Week 2

	Sunday 21-June	Monday 22-June	Tuesday 23-June	Wednesday 24-June	Thursday 25-June	Friday 26-June	Saturda 27-Jun
AM		8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	
AM		Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	
AM		9:30-10:30, Lecture	9:30-10:30, Lecture	9:30-10:30, Lecture	9:30-10:30, Lecture	9:30-12,	
0 AM						*to be*	
0 AM		11-12, Reading Asst.	11-12, Lecture	11-12, Lecture	11-12, Lecture	*announced*	Berkele
n		Lunch	Lunch	Lunch	Lunch	Lunch	Walking
PM							Tour
PM		1:15-3:15, Reading	1:15-3:30, Research	1:15-3:30, Research	1:15-3:30, Research		
PM		Assignment, cont.	Team Meetings	Team Meetings	Team Meetings	2-3:30, Colloquium,	
PM		3:15, Tea	3, Tea	3, Tea	3, Tea	Gina-Maria Pomann	
PM		3:45-5, Groups/	3:30-5:15, Research	3:30-5:15, Research	3:30-5:15, Research	3:30, Tea	
PM		Research Teams	Team Meetings	Team Meetings	Team Meetings	Shuttle to Mining Circle	
PM		Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	3:55, 4:25, 4:55,	
PM		5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	

2015 MSRI-UP Calendar, Weeks 3 and 4

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturda
	28-June	29-June	30-June	1-July	2-July	3-July	4-July
AM		8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,		
AM		Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI		
AM		9:30-11:45, Research	9:30-11:45, Research	9:30-11:45, Research	9:30-11:45, Research		
0 AM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	Federal, Campus,	
0 AM						and MSRI-UP Holiday	
n		Lunch	Lunch	Lunch	Lunch		
PM					1-2, Graduate Student		
PM		1:15-3:30, Research	1:15-3:30, Research	1:15-3:30, Research	Panel, Part B	approx. 4:00, travel to:	4:30 - 8:
PM		Team Meetings	Team Meetings	Team Meetings	2-3:30, Colloquium,		
PM		3, Tea	3, Tea	3, Tea	Federico Ardila	6:05: MLBaseball:	Picnic/
PM		3:30-5:15, Research	3:30-5:15, Research	3:30-5:15, Research	3:30, Tea	Oakland Athletics	Barbecu
PM		Team Meetings	Team Meetings	Team Meetings	Shuttle to Mining Circle	vs.	Codornic
PM		Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	3:55, 4:25, 4:55,	Seattle Mariners	Park,
PM		5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	& fireworks	Berkele

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturda
	5-July	6-July	7-July	8-July	9-July	10-July	11-Jul
AM		8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	8:10, 8:40, 9:10,	
AM		Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	
AM		9:30-11:45, Research	9:30-11:45, Research	9:30-11:45, Research	9:30-11:45, Research	9:30-12, Graduate School	
0 AM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	and Fellowship	
0 AM						Workshop	
n		Lunch	Lunch	Lunch	Lunch	Lunch	Santa
PM							Cruz
PM		1:15-3:30, Research	1:15-3:30, Research	1:15-3:30, Research	1:15-3:30, Research		Beach
PM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	2-3:30, Colloquium,	Boardwa
PM		3, Tea	3, Tea	3, Tea	3, Tea	Bobby Wilson	
PM		3:30-5:15, Research	3:30-5:15, Research	3:30-5:15, Research	3:30-5:15, Research	3:30, Tea	
PM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	Shuttle to Mining Circle	
PM		Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	3:55, 4:25, 4:55,	
PM		5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	

2015 MSRI-UP Calendar, Weeks 5 and 6

	Sunday 12-July	Monday 13-July	Tuesday 14-July	Wednesday 15-July	Thursday 16-July	Friday 17-July	Saturda 18-Jul
AM		7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40,	7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40, 9:10,	
AM		Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	
AM		9:30-12, Research	9:30-12, Research	9-10:15, Team Meetings	9:30-12, Research	9:30-12, Research	
0 AM		Team Meetings	Team Meetings	10:30-12, Collquium,	Team Meetings	Team Meetings,	10-3,
0 AM				Khalilah Beal		Progress Reports	
n		Lunch	Lunch	Lunch	Lunch	Lunch & virtual seminar,	The
0 PM						Jennifer Johnson-Leung	Explora
PM		1-3, Research	1-3, Research		1-3:30, Team Meetings,	1-3, Team Meetings,	torium
PM		Team Meetings	Team Meetings	2-3:30, Career Panel	Progress Reports	Progress Report	
PM		3, Tea	3, Tea		3, Tea	3, Tea	
PM		3:30-5:15, Research	3:30-5:15, Research	3:30, Tea	3:30-5:15, Team Mtgs.,	3:30-5:15, Research	
PM		Team Meetings	Team Meetings	4-5:15, Team Meetings	Progress Report	Team Meetings	
PM		Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	
PM		5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	

	Sunday 19-July	Monday 20-July	Tuesday 21-July	Wednesday 22-July	Thursday 23-July	Friday 24-July	Saturda 25-Jul
AM		7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40, 9:10,	7:40, 8:10, 8:40, 9:10,	
AM		Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	Shuttle to MSRI	
AM		9:30-12, Research	9:30-12, Research	9:30-12, Research	9:30-12, Research	9:30, 10:15, 11, Final	
0 AM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	Research Presentations	
0 AM							
n		Lunch	Lunch	Lunch	Lunch	Lunch	
0 PM							
PM		1-3, Research	1-3, Research	1-3, Research	1-3, Research	1:15, 2, 2:45, Final	
PM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	Research Presentations	
PM		3, Tea	3, Tea	3, Tea	3, Tea		
PM		3:30-5:15, Research	3:30-5:15, Research	3:30-5:15, Research	3:30-5:15, Research	3:30, Tea	
PM		Team Meetings	Team Meetings	Team Meetings	Team Meetings	Shuttle to Mining Circle	5-7:30,
PM		Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	Shuttle to Mining Circle	3:55, 4:25, 4:55,	Farewe
PM		5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	5:25, 5:55, 6:25	Dinner

Mathematical Sciences Research Institute Undergraduate Program
(MSRI-UP)

2015 MSRI-UP

Geometric Combinatorics Motivated by the Social Sciences

Final Research Presentations

Friday, July 24
Mathematical Sciences Research Institute
Edward D. Baker Boardroom

MSRI - UP

The Mathematical Sciences Research Institute Undergraduate Program (MSRI-UP)

The MSRI-UP is a comprehensive program for undergraduates that aims to increase the number of students from underrepresented groups in mathematics graduate programs. MSRI-UP includes summer research opportunities, mentoring, workshops on the graduate school application process, and follow-up support.

The primary objective of the MSRI-UP is to identify talented students, especially those from underrepresented groups, who are interested in mathematics and make available to them meaningful research opportunities, the necessary skills and knowledge to participate in successful collaborations, and a community of academic peers and mentors who can advise, encourage, and support them through a successful graduate program. We achieve this through an intensive six-week summer program of mathematics research and other activities, along with maintenance of relationships with participating students for years beyond the summer program.

The MSRI-UP is coordinated by an experienced team of five directors, Professors Federico Ardila of San Francisco State University, Duane Cooper of Morehouse College, Herbert Medina of Loyola Marymount University, Ivelisse Rubio of the University of Puerto Rico, Río Piedras, and Suzanne Weekes of Worcester Polytechnic Institute, who collaborate ongoingly and who annually rotate direct leadership of the program. The program is supported by the leadership and staff of the Mathematical Sciences Research Institute in Berkeley, site of each summer's six-week program.

During the 2007-2014 summers, 134 students have conducted 45 small group research projects in Computational Mathematics, Experimental Mathematics, Coding Theory, Elliptic Curves and Applications, Mathematical Finance, Enumerative Combinatorics, Algebraic Combinatorics, and Arithmetic Aspects of Elementary Functions. Most MSRI-UP participants who have graduated college proceeded to enter graduate programs in the mathematical sciences. This year, 17 undergraduates are completing MSRI-UP, having learned about and engaged in research on Geometric Combinatorics Motivated by the Social Sciences, led by Professor Francis Edward Su of Harvey Mudd College, who has guided the students in conjunction with a postdoctoral fellow and two graduate students—carefully chosen role models—who contribute to the undergraduates' academic, personal, and professional development. In 2016, the MSRI-UP expects to continue with 18 undergraduates conducting research projects in Applied Algebraic Geometry led by Professor Luis Garcia Puente of Sam Houston State University.

MSRI-UP Final Presentations
Geometric Combinatorics Motivated by the Social Sciences

Schedule of Presentations

9:30 am
Double-*n* Circular Societies
Edwin Baeza, Nikaya Smith, Sarah Yoseph

10:15 am
An Analogue of the Median Voter Theorem for Approval Voting
Ethan Bush, Kyle Duke, Miles Stevens

11 am
A Matroid Generalization of Sperner's Lemma
Gabriel Andrade, Andrés Rodriguez Rey, Alberto Ruiz

11:45-1:15
Lunch

1:15 pm
Committee Selection with Approval Voting and Hypercubes
Caleb Bugg, Gabriel Elvin

2 pm
The Banquet Seating Problem
Michelle Rosado Pérez, Ashley Scruse, A.J. Torre

2:45 pm
A Volume Argument for Tucker's Lemma
Beauttie Kuture, Oscar Leong, Christopher Loa

3:30
Afternoon Tea

Geometric Combinatorics Motivated by the Social Sciences

Abstracts

Double-n Circular Societies

Edwin Baeza, Nikaya Smith, Sarah Yoseph

A *society* is a geometric space with a collection of subsets that represent voter preferences. We call this space the *spectrum* and these preference sets *approval sets.* The agreement proportion is the largest fraction of approval sets that intersect in a common point. Klawe et al. considered linear societies where approval sets are the disjoint union of two intervals, or double intervals. We examine arc-shaped double intervals on circular societies. We consider the case of pairwise-intersecting intervals of equal length and call these double-n circular societies. What is the minimal agreement proportion for double-n societies? We show that the asymptotic agreement proportion is bounded between 0.3333 and 0.3529and conjecture that the proportion approaches 1/3.

An Analogue of the Median Voter Theorem for Approval Voting

Ethan Bush, Kyle Duke, Miles Stevens

The Median Voter Theorem is a well-known result in social choice theory for majority-rule elections. We develop an analogue in the context of approval voting. We consider voters to have preference sets that are intervals on a line, called *approval sets*, and the *approval winner* is a point on the line that is contained in the most approval sets. We define *median voter* by considering the left and right end points of each voters approval sets. We consider the case where approval sets are equal length. We show that if the pairwise agreement proportion is at least 3/4, then the median voter interval will contain the approval winner. We also prove that under an alternate geometric condition, the median voter interval will contain the approval winner, and we investigatevariants of this result.

A Matroid Generalization of Sperner's Lemma

Gabriel Andrade, Andrés Rodriguez Rey, Alberto Ruiz

In a 1980 paper, Lovász generalized Sperner's lemma for matroids. He claimed that a triangulation of a d-simplex labeled with elements of a matroid M must contain at least one "basis simplex". We present a counterexample to Lovász's claim when the matroid contains singleton dependent sets and provide an additional sufficient condition that corrects Lovász's result. Furthermore, we show that under some conditions on the matroids, there is an improved lower bound on the number of basis simplices. We present further work to sharpen this lower bound by looking at M's lattice of flats and by proving that there exists a group action on the simplex labeled by M with $\mathcal{S}_n$.

Committee Selection with Approval Voting and Hypercubes

Caleb Bugg, Gabriel Elvin

In this paper we will examine elections of the following form: a committee of size k is to be electedwith two candidates running for each position. Each voter submits a ballot with his or her ideal committee, which generates their *approval set.* The approval sets of voters consist of committees that are "close" to their ideal preference. We define this notion of closeness with Hamming distance in a hypercube: the number of candidates by which a particular committee differs from a voter's ideal preference. We establish a tight lower bound for the popularity of the most approved committeeand consider restrictions on voter preferences that may increase that popularity. Our approach considers both the combinatorial and geometric aspects of these elections.

The Banquet Seating Problem

Michelle Rosado Pérez, Ashley Scruse, A.J. Torre

Suppose you want to seat $n = mk$ people around k tables with m people at each table. Each person gives you a list of j people next to whom they would enjoy sitting. What is the smallest j for which you can always make a seating arrangement that would seat each person next to one of the people on their list? In this paper we show that j must be strictly more than half of n, the total number of people. Our key tool is a particular 'blue-green-red' lemma that helps us construct 'worst-case scenario' seating arrangements. We consider cases with two tables and more than two tablesand explore seating arrangements with particular kinds of preferences.

A Volume Argument for Tucker's Lemma

Beauttie Kuture, Oscar Leong, Christopher Loa

Sperner's lemma is a statement about labeled triangulations of a simplex. McLennan and Tourky (2007) provided a novel proof of Sperner's Lemma using a volume argument and a piecewise linear deformation of a triangulation. We adapt a similar argument to prove Tucker's Lemma on a triangulated cross-polytope P in the 2-dimensional case where vertices of P have different labels. The McLennan-Tourky technique would not directly apply because the natural deformation distorts the volume of P; we remedy this by inscribing P in its dual polytope, triangulating it, and considering how the volumes of deformed simplices behave.

2015 MSRI-UP Staff

Duane Cooper, Morehouse College
Lead Director

Francis Edward Su, Harvey Mudd College
Faculty Research Leader

Mutiara Sondjaja, New York University
Postdoctoral Researcher

Daniel Eckhardt, Rensselaer Polytechnic Institute
Graduate Student Assistant

Lyda Pam Urresta, University of Notre Dame
Graduate Student Assistant

MSRI-UP Directors

Federico Ardila, San Francisco State University
Duane Cooper, Morehouse College
Herbert Medina, Loyola Marymount University
Ivelisse Rubio, University of Puerto Rico, Río Piedras
Suzanne Weekes, Worcester Polytechnic Institute

MSRI-UP Thanks:

The MSRI—the Institute and its staff
The National Science Foundation
The National Security Agency

Double-n Circular Societies

Edwin Baeza
Purdue University

Nikaya Smith
UNC-Chapel Hill

Sarah Yoseph
Loyola Marymount University

24 July 2015

[1]

Abstract

A *society* is a geometric space with a collection of subsets that represent voter preferences. We call this space the *spectrum* and these preference sets *approval sets*. The agreement proportion is the largest fraction of approval sets that intersect in a common point. Klawe et al. considered linear societies where approval sets are the disjoint union of two intervals, or double intervals. We examine arc-shaped double intervals on circular societies. We consider the case of pairwise-intersecting intervals of equal length and call these double-n circular societies. What is the minimal agreement proportion for double-n societies? We will show that the asymptotic agreement proportion is bounded between 0.2500 and 0.3529, and conjecture that the proportion approaches 1/3.

1 Introduction

Consider a voting system where participants indicate their preferences on a political spectrum that is some geometric space (often, a line in $\mathbb{R}^1$, but could be other spaces as well). Points on the spectrum are called *platforms*. Each voter has an approval set, which consists of all platforms on the spectrum which they approve. Following [1], we define a *society* to be a tuple containing the voters, political spectrum, and approval sets. The *agreement proportion* of a society is the fraction of voters whose approval sets intersect at some point. We wish to draw conclusions about the guaranteed agreement proportion, for any society whose approval sets satisfy some condition.

A first example is to consider a line as a political spectrum and model each voter's approval set as an interval. If we assume that these intervals are pairwise intersecting, then Helly's theorem [4] would imply that there is a platform that all voters approve. We can think of this as saying that all voters would "agree" on this platform.

---

[1] This work was conducted during the 2015 Mathematical Sciences Research Institute Undergraduate Program (MSRI-UP), supported by grants from the National Science Foundation (DMS 1156499) and the National Security Agency (H98230-15-1-0039). We also thank supervisors Duane Cooper, Francis Edward Su, and Daniel Eckhardt.

Many researchers have considered various modifications on the kinds of approval sets allowed and conditions that they satisfy. Rather than looking at pairwise intersecting societies, Berg et al. [1] studied (k,m)-agreeable societies, where for every m voters, there is a set where k of them agree. They discovered a guaranteed agreement proportion of $\frac{k-1}{m-1}$. Hardin [2] expanded this approach by considering arc-shaped approval sets on a circle, and found an agreement proportion of $\frac{k-1}{m}$. Klawe et al. [3] considered preferences on a line where the approval sets were not convex, but rather two disjoint intervals. They found that the agreement proportion of these societies is at least 0.268. For certain kinds of double interval societies, those arising from *double-n strings* (made precise in Section 3), they show that the asymptotic agreement proportion is between 8/23 and 5/13.

In our work, we consider societies where the spectrum is a circle and the preference sets are unions of two disjoint intervals. What is the guaranteed agreement proportion for such societies?

We consider the case where the intervals are all of equal length, and our main result (Theorem 4.2) shows that the asymptotic agreement proportion lies between 1/4 and 6/17.

This paper is organized as follows. We first give motivation for studying such societies. Then, we discuss the connection between these societies and certain strings we call double-n strings. Then we prove our results and make some conjectures.

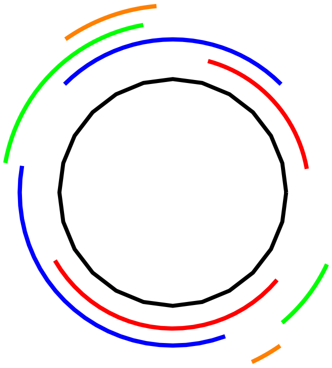


Figure 1: A Double Interval Circular Society.

2 Motivation

In linear societies, each end of the spectrum consists of opposing platforms, for example, liberal and conservative views. Circular societies allow for these opposing platforms to “blend” (think of a color wheel), which is useful for modeling political ideologies. Circular societies are also optimal for visualizing recurring phenomena, such as time. For example, scheduling work shifts for a 24-hour restaurant, or finding a meeting time on a twelve-month calendar.

Why double intervals? Perhaps a work team needs to find two work shifts per-person for a 24-hour schedule. Or, in the case of a political spectrum, a voter may have a range of both moderate and extreme views. In either case, voter approval sets consist of disjoint intervals on a circle. In our research project, we examine the double-n society on a circular model. Rather than working with disjoint intervals specifically, we constructed double-n strings which simplified the process of finding results about our circular societies.

3 Double-n Circular Societies and Double-n Strings

We limit our investigation in this paper to circular societies where approval sets are two disjoint intervals that are the same length. Even in this case, the combinatorics is rich, as evidenced by Klawe et al.'s similar analysis [3] for double-interval societies on lines. We call such societies *double-n circular societies* because of their connection to *double-n strings*, which we explain below.

Definition 3.1. A *double-n circular society* is a society where all approval sets are two disjoint intervals of equal length.

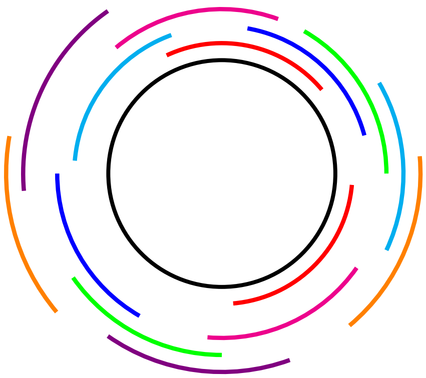
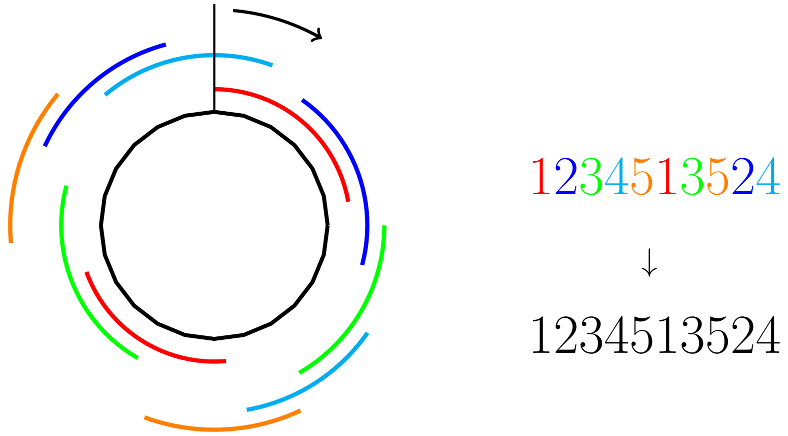


Figure 2: A Double-n Circular Society

Definition 3.2. A *double-n string* is a string of length $2n$ such that each of n symbols appears twice.

Example 1. Consider the strings 12344123 and 3141534252. They are respectively double-4 and double-5 strings.

In Figure 3 below, we see that these double-n strings can be constructed from double-n circular societies.



A Double-n
Circular Society

Corresponding
Double-n String

Figure 3: Constructing a Double-n String

Note that in a circular society, double-n strings are "loops", that is, the last symbol is considered adjacent to the first symbol.

Definition 3.3. The *distance* between two symbols i and j is the smallest number of symbols separating any instance of them, where the distance between two adjacent symbols for instance is 1. The *diameter* is the maximum of all such distances between any two distinct symbols.

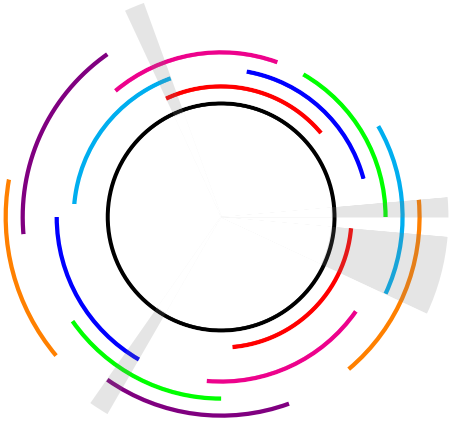


Figure 4: The double-7 string $\cdots 12345167325746 \cdots$ with diameter $d = 2$. The agreement proportion is given by $(d + 1)/n = 3/7$.

The minimal diameter $\delta(n)$ gives the number of intervals that intersect with some element i (exclusive), so to find the *agreement number*, or maximum number of intervals that intersect, we must add one. Just as in the linear case, we can see that the smallest agreement number possible is $\delta(n)+1$. Note the fraction $(\delta(n)+1)/n$ is the smallest possible agreement proportion for double-n circular societies.

$$\cdots\ \underbrace{1\,2\,3\,4}_{\text{Occurrence 1}}\ \_\_\ \cdots\_\_\ \underbrace{4\,1\,3\,2}_{\text{Occurrence 2}}\ \_\_\cdots$$

Figure 5: The *distance* between the first occurrences of 1 and 4 is 3. The *diameter* among all instances of {1,2,3,4} is 1.

3.1 Patterns in Double-n Strings

In this section, we explore some examples of double n-strings to give motivation for our results later. First, we establish an important lemma that shows that these double-n strings have some structure.

Lemma 3.4. *Let S be a double-n string with diameter d. Then there exists a substring in S of at least size $(n-d)$ with no repeated symbols.*

For instance, 12341572653764 is a double-7 string with $d = 2$. Note that 72653 is a non-repeating substring of length 5.

Proof of Lemma 3.4. Suppose that there is no such substring. Then every substring of S of size $n-d$ has at least one symbol that repeats twice. So, pick a substring L_1 of size $n-d$. The supposition above shows that there is a symbol, say x, that repeats twice in L_1. We can form a new substring L_2 by adjoining the adjacent d elements to either end of L_1. Note that L_2 is of size $n+d$. Now consider the substring $L_3 = S \setminus L_2$. Since the size of L_3 is $2n-(n+d) = n-d$, we have by our supposition that there is an element, say y, that occurs twice in L_3. Now consider the distance between x and y. By construction, this distance is at least $d+1$, a contradiction. □

We are interested in constructing double-n strings that minimize the diameter d. Let's take a closer look at $n = 5$, with the set $\{1, 2, 3, 4, 5\}$. Can we construct one with diameter 1?

We know by Lemma 3.4 there is a non-repeating substring of length 4, so we can assume without loss of generality that this is the substring 1234. There are now constraints on the placement of other symbols imposed by adjacencies already achieved in that substring. A particular symbol has exactly 4 adjacent spots (2 for each instance of that symbol), and there are exactly 4 other symbols that need to fill those 4 spots. So it is at the outset unclear whether we will be able to do this. However here is one that works:

$$\cdots\mathbf{1\,2\,3\,4\,1\,5\,2\,4\,5\,3}\cdots$$

Note that if we wanted a double-6 string, we know that it cannot have diameter 1, because any symbol would need to be adjacent to 5 other symbols but there are only 4 adjacent slots for each symbol. Here is a double 6-string with diameter 2.

$$\cdots \mathbf{1\,2\,3\,4\,1\,6\,5\,2\,6\,5\,3\,4} \cdots$$

A similar problem arises with $d = 2$ and $n = 9$, so we must increase the diameter to $d = 3$.

$$\cdots \mathbf{1\,2\,3\,4\,5\,6\,7\,1\,8\,5\,9\,2\,6\,4\,8\,3\,9\,7} \cdots$$

One would think that $n = 9$, $d = 2$ is a possible construction, because any symbol can have at most $4d$, or 8 other symbols adjacent to it. However in this case, any substring of size $2d+1$ must be distinct. It is not possible to satisfy Lemma 3.4 with this construction. Therefore, we conjecture that the number of $n-d$ distinct elements cannot exceed $2d + 2$, which is discussed in Section 4.8.

4 The Asymptotic Agreement Proportion of Double-n Strings.

Our focus in this section is to find asymptotic results about the agreement proportion. We begin with the following definition.

Definition 4.1. The *asymptotic agreement proportion* Δ_c for a double-n circular society is

$$\Delta_c = \lim_{n\to\infty} \frac{\delta(n)+1}{n} = \lim_{n\to\infty} \frac{\delta(n)}{n}\,.$$

Theorem 4.2. *For a double-n circular society*

$$\frac{1}{4} \le \Delta_c \le \frac{6}{17}\,.$$

Lemma 4.3. *For any positive integer n, we have $\delta(n) \ge (n-1)/4$.*

Proof. Consider any double-n string with diameter d. Now note that any symbol i is adjacent to at most $4d$ symbols (see Figure 6 below). Now as i must be adjacent to all $n-1$ symbols, we must have that $4d \ge n-1$.

$\square$

$$\_\ \cdots \_\, i \underbrace{\_\ \cdots \_}_{d \text{ symbols}} \quad \cdots \quad \_\ \cdots \_\, i \_\ \cdots \_$$

Figure 6: The adjacencies of i.

We also have the trivial upper bound $\Delta_c \le n/2$. This can be seen by considering the double-n string $1\,2\,\ldots\,n\,1\,2\,\ldots n$, which shows that $\delta(n) \le n/2$ for $n \ge 2$. It can be shown that this equality is strict for $n \ge 3$.

Lemma 4.4. *Let $n \geq 3$. Then $\delta(n) < n/2$.*

Proof. We first note that the strings 123123, 12341324, and 1234152453 show the lemma is true for $n = 3, 4$, and 5. Now for $n > 5$, we partition $1, 2, \ldots n$ into 5 strings S_1, S_2, S_3, S_4, S_5, explained below, and examine the diameter of the double-n string $T(n) = S_1, S_2, S_3, S_4, S_1, S_5, S_2, S_4, S_5, S_3$.

For $n > 5$, let $q = \lfloor \frac{n}{5} \rfloor$. First suppose that $n = 5q$ and let S_i be consecutive strings, each of length q. It is easily verified that the diameter of $T(n)$ is $2q - 1$ and that $2q-1 < n/2$. Next, if $n = 5q+1$, let S_1, S_2, S_3, S_4 be strings of length q as before and S_5 be the string $4q + 1, \ldots, 5q + 1$ of length $q + 1$. The diameter of $T(n)$ is now $2q$ and, similarly, $2q < n/2$. For $n = 5q + 2$, let S_1, S_2, S_3 be strings of length q and S_4, S_5 have length $q + 1$. The diameter of $T(n)$ is now $2q + 1 < n/2$. If $n = 5q + 3$, let S_1, S_2 be strings of length q and S_3, S_4, S_5 have length $q + 1$. It follows that the diameter of $T(n)$ is $2q + 1$ and $2q + 1 < n/2$. Finally, if $n = 5q + 4$, let S_1 be a string of length q and S_2, S_3, S_4, S_5 have length $q + 1$. The diameter $T(n)$ is still $2q + 1$ and $2q + 1 < n/2$.

□

Theorem 4.5. *For $n > 0$, consider a double-n string with minimal diameter δ. Then $n \leq 3\delta + 2$.*

Proof. The result is trivial for $n \leq 2$. The strings 123123, 12314324, and 1234152453 of minimal diameter 1 show that this is true for $n \leq 5$.

For $n > 5$, let $q = \lfloor \frac{n}{5} \rfloor$ and let $T(n)$ be the double-n string constructed in the proof of Lemma 4.4. Since δ is by minimality less than or equal to the diameter of $T(n)$, we can obtain a relationship between δ and n. We hope to show $\delta \geq \frac{n-2}{3}$.

Suppose that $n = 5q$. It was shown above that the diameter of $T(n)$ is $2q - 1$. Since $q > 1$, it follows that $2q - 1 \geq \frac{n}{3} - \frac{2}{3}$. Next, if $n = 5q + 1$, the diameter of $T(n)$ is $2q$ and $2q \geq \frac{n}{3} - \frac{2}{3}$ follows from $q > -1$. Now if $n = 5q + 2$, $T(n)$ has diameter $2q + 1$ and $2q + 1 \geq \frac{n}{3} - \frac{2}{3}$. It was shown above that if $n = 5q + 3$ or $n = 5q + 4$, then the diameter of $T(n)$ is again $2q + 1$. It follows that $2q + 1 \geq \frac{n}{3} - \frac{2}{3}$ for both $n = 5q + 3$ and $n = 5q + 4$.

□

Lemma 4.6. *For any symbol i in a double-n string, let r_i be the number of symbols which occur twice within diameter d of each occurrence of i. Then $r_i \leq 4d - n + 1$.*

Proof. For each occurrence of i, there are $2d$ adjacent spaces in which r_i of the elements can repeat. Since the element i must be adjacent to the other elements, we must have that $n - 1 \leq 4d - r_i$. So, $r_i \leq 4d - n + 1$.

□

Theorem 4.7. *For any $n > 0$ there exists a double-n string with diameter $d \leq 6\lceil \frac{n}{17} \rceil - 1$.*

Proof. Consider the double-17 string

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 2, 13, 14, 5, 16, 15, 7, 8, 17, 1, 9, 3, 13, 14, 4, 16, 15, 6, 17, 10, 11, 12.$$

Note that this string has diameter 5. We will use this string to construct a general double-n string. Let $k = \lceil \frac{n}{17} \rceil$. Replace each symbol i in the double-17 string with the following substring:

$$k(i-1)+1, k(i-1)+2, \ldots, k(i-1)+k\,.$$

We omit any symbols greater than n. Since the double-17 string had diameter 5, any two substrings are separated by at most 4 substrings of size k. Now consider any two symbols $1 \leq i, j \leq n$. Suppose that they are as far apart as possible. Then they are at the opposite of their respective substrings, so they are separated by at most $2k-1$ symbols. Thus we have that

$$d \leq 4k + (2k-1) = 6k - 1 = 6\left\lceil \frac{n}{17} \right\rceil - 1$$

$\square$

Proof of Theorem 4.2. The first inequality follows from Lemma 4.3. The second follows from Theorem 4.7. $\square$

4.1 Finding Restrictions for Double-n Strings

Considering Lemma 3.4, we know that there exists a unique sequence of size $n-d$ for some double-n string.

Conjecture 4.8. For any double-n string with minimal diameter δ we have that

$$2\delta \leq n - \delta \leq 2\delta + 2.$$

If true, this conjecture would imply that $\Delta_c = 1/3$.

The following table gives our examples of double-n strings with minimal diameter δ and an agreement proportion (A. P.) of $\frac{\delta+1}{n}$ in the last column. Note that the values in the table agree with Theorem 4.5.

n	δ	**Double-n String**	**A. P.**
3	1	1 2 3 1 2 3	.6667
4	1	1 2 3 4 2 1 3 4	.5000
5	1	1 2 3 4 1 5 2 4 5 3	.4000
6	2	1 2 3 4 1 6 5 2 6 5 3 4	.5000
7	2	1 2 3 4 5 1 6 2 5 7 3 6 4 7	.4286
8	2	1 2 3 4 5 1 6 8 2 5 7 3 8 6 4 7	.3750
9	3	1 2 3 4 5 6 7 1 8 9 4 3 2 7 6 5 9 8	.4444
10	3	1 2 3 4 5 6 7 1 8 9 10 4 3 2 7 6 5 10 9 8	.4000
11	3	1 2 3 4 5 6 7 8 1 9 10 2 6 8 11 3 4 9 10 5 11 7	.3636
12	4	1 2 3 4 5 6 7 8 1 2 9 10 5 11 12 4 9 10 3 8 6 7 11 12	.3636
13	4	1 2 3 4 5 6 7 8 9 1 10 11 3 8 2 12 13 9 7 10 11 4 6 5 12 13	.3846
14	4	1 2 3 4 5 6 7 8 9 10 1 13 14 6 2 11 12 7 8 3 13 14 4 5 11 12 9 10	.3571
15	5	1 2 3 4 5 6 7 8 9 10 11 12 2 13 14 5 15 7 8 1 9 3 13 14 4 15 6 10 11 12	.4000
16	5	1 2 3 4 5 6 7 8 9 10 11 12 2 13 14 5 16 15 7 8 1 9 3 13 14 4 16 15 6 10 11 12	.3750
17	5	1 2 3 4 5 6 7 8 9 10 11 12 2 13 14 5 16 15 7 8 17 1 9 3 13 14 4 16 15 6 17 10 11 12	.3529

5 Conclusion and Open Questions

Our study of circular societies expanded on previous research in voting theory. Rather than studying linear societies, we considered societies on a circle. In particular, we studied societies where all approval sets consist of two disjoint sets of equal size. From these societies we construct double-n strings and use them to establish and conjecture that:

$$2\delta \leq n - \delta \leq 2\delta + 2.$$

While we were unable to prove the lower bound of our conjecture, we have produced a bound for the asymptotic agreement proportion for voter size n. We found that:

$$\frac{1}{4} \leq \Delta_c \leq \frac{6}{17}.$$

This means that the minimal agreement proportion falls within 0.3333 and 0.3529. If we were to prove our conjecture, however, we would find an exact value of 1/3.

There are still many questions to be explored. How would this bound change if the voter intervals were not the same length? What if some approval sets were disjoint and others were not? Can we establish bounds for approval sets in three dimensions? A significant inquiry for the construction of double-n strings is whether there is a programming method for finding any string of size n so that asymptotic agreement proportions can be more accurately found.

6 Acknowledgments

We would like to thank Dr. Francis Su, Mr. Daniel Eckhardt, Dr. Duane Cooper, and the rest of the MSRI-UP staff. This work was supported by the National Science Foundation DMS 0754872 and the National Security Agency H98230-11-1-0213

References

[1] Deborah E. Berg, Serguei Norine, Francis Edward Su, Robin Thomas, and Paul Wollan. Voting in agreeable societies. *Amer. Math. Monthly*, 117(1):27–39, 2010.

[2] Christopher S. Hardin. Agreement in circular societies. *Amer. Math. Monthly*, 117(1):40–49, 2010.

[3] Maria Margaret Klawe, Kathryn L. Nyman, Jacob N. Scott, and Francis Edward Su. Double-interval societies. In *The mathematics of decisions, elections, and games*, volume 624 of *Contemp. Math.*, pages 135–146. Amer. Math. Soc., Providence, RI, 2014.

[4] Csaba D Toth, Joseph O'Rourke, and Jacob E Goodman. *Handbook of discrete and computational geometry*. CRC press, 2004.

An Analogue of the Median Voter Theorem for Approval Voting

Ethan Bush
University of Michigan - Flint

Kyle Duke
James Madison University

Miles Stevens
Morehouse College

February 5, 2016

Abstract

The Median Voter Theorem is a well-known result in social choice theory for majority-rule elections. We develop an analogue in the context of approval voting. We consider voters to have preference sets that are intervals on a line, called *approval sets*, and the *approval winner* is a point on the line that is contained in the most approval sets. We define *median voter* by considering the left and right end points of each voter's approval sets. We consider the case where approval sets are equal length. We show that if the pairwise agreement proportion is at least $\frac{3}{4}$, then the median voter interval will contain the approval winner. We also prove that under an alternate geometric condition, the median voter interval will contain the approval winner, and investigate variants of this result.

[1]

---

[1] This work was conducted during the 2015 Mathematical Sciences Research Institute Undergraduate Program (MSRI-UP), supported by grants from the National Science Foundation (DMS 1156499) and the National Security Agency (H98230-15-1-0039). We also thank supervisors Duane Cooper, Francis Edward Su, and Dan Eckhardt.

1 Introduction

In a majority rule voting system, any candidate receiving more than 50% of the votes will win the election. Voters are said to have single-peaked preferences if there exists some ideal choice with all other choices being preferred less than that ideal choice. Under a two-candidate, majority-rule voting system with single peaked preferences, the candidate who places himself on the "philosophical median" of the political spectrum will win the election [3]. This is the Median Voter Theorem.

Approval voting is another single winner election method first suggested in 1978 by Steven Brams and Peter Fishburn [2]. Under approval voting, a voter can vote for any number of candidates. Each candidate is considered independently of his competitors. Approval voting is not used in larger political elections in the United States, but it has been adopted by the Mathematical Association of America, the Institute for Operations Research and the Management Sciences, the American Statistical Association and many other organizations to elect their leaders. It has some advantages.

Consider a situation where there exists three candidates A, B, and C. Let's say A is the first choice for 51% of the voters. B is the first choice for 39% of the voters and these voters dislike A. C is the first choice for 10% of the population, but C is the second choice for every voter who prefers A or B. Under a single vote election method, candidate A would win the election and many voters would be unsatisfied. In approval voting, if each voter votes for their top two choices, candidate C could win the election and leave more voters satisfied with the results. Because each candidate gets either a yes or no from each voter, a candidate cannot be considered a voter's ideal preference. Therefore, voter preferences are not single peaked. Additionally, voters can vote for multiple candidates, so the Median Voter Theorem does not apply.

Is there an analogue of the Median Voter Theorem for approval voting? In this paper, we develop some analogues. To do this, we must define what we mean by "median voter" and then we will suggest conditions where an approval winner is approved by the median voter.

Our model for approval voting follows that of Berg et al.[1]. We imagine the set of all possible preferences to be a geometric space, such as line, and we will call this space the *spectrum*. Each element of the spectrum is a *platform*. Assume that there is a finite set V of voters and each has an *approval set* I_v of $platforms$. A *society* is a spectrum, together with a set of voters and their approval sets. If the spectrum is a line, we call it a *linear society*.

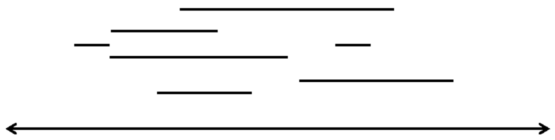


Figure 1: A linear society with interval approval sets (indicated as segments separated from the line for ease of viewing).

The *agreement number* of a platform is the number of voters who agree on that platform. The *agreement number* of a society is the maximum agreement number over all platforms. The platform that has the highest agreement number is called the *approval winner*. The *agreement proportion* $\rho(S)$ of society S is the agreement number of a society divided by the number of voters. The *pairwise agreement proportion* $\alpha(S)$ of a society S is the number of pairwise intersections divided by the total number of pairs of voters.

We show in Theorem 3.1 that if the pairwise agreement proportion is high enough, then the median voter interval will contain the approval winner. We also show in Theorem 4.1 that under certain geometric conditions involving looking at sets to the "right" and "left" of the approval winner, we can say that the median voter interval will contain the approval winner.

2 Developing a Notion of Median Voter

Our task is to define the notion of the median voter for approval voting. We developed three ways to define a median voter, and hence the median voter interval.

Right Median Method. Label the approval sets in numerical order from the left-most right endpoint to the right-most right end point, from $1, 2, ..., n$. See Figure 2.

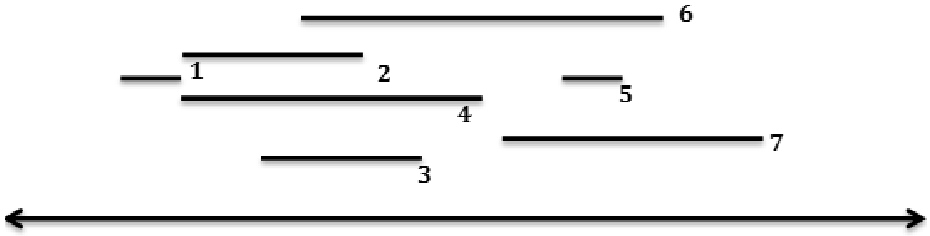


Figure 2: Labeling approval sets by their right endpoint.

When n is odd, the median voter is the interval labeled $I_{\frac{n+1}{2}}$. In Figure 3, it is set I_4. When n is even the median voter is the set containing the union of the approval sets labeled $I_{\frac{n}{2}}$ and $I_{\frac{n}{2}+1}$.

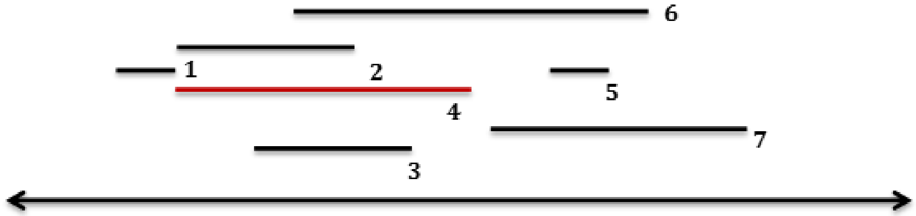


Figure 3: The right median voter interval.

Left Median Method. Label the approval sets in numerical order from the left-most left endpoint to the right-most left endpoint, from $1, 2, ..., n$. See Figure 3

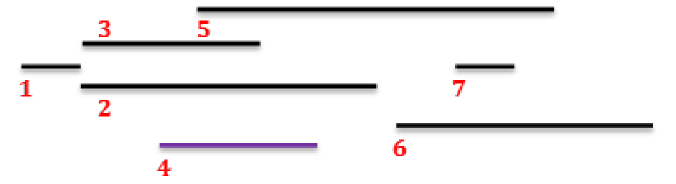


Figure 4: Labeling approval sets by their left endpoint. The left median approval interval is set number 4.

Notice in Figure 5 that both the left and right median voter intervals are different, but both contain the winner.

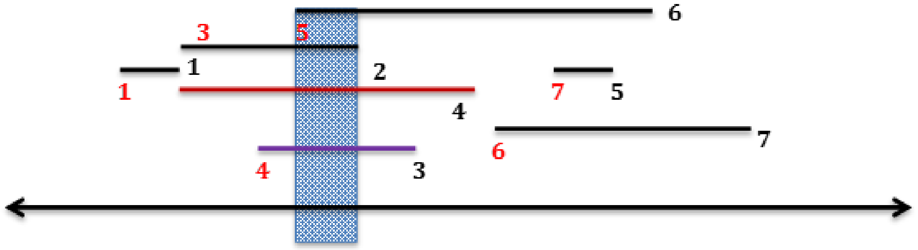


Figure 5: There is an interval of approval winners (shaded) and the left and right median approval intervals contain these winners.

Midpoint Median Method. Label the approval sets in numerical order from the left-most midpoint to the right-most midpoint, from $1, 2, ..., n$. See Figure 6

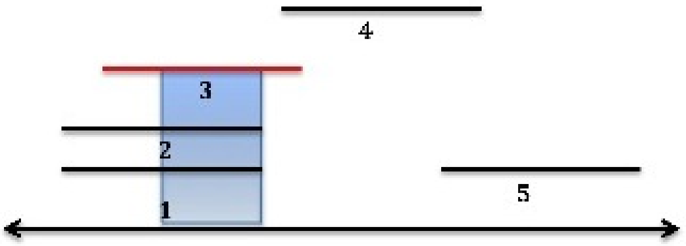


Figure 6: The midpoint median method labels, midpoint median voter interval (red), and approval winners (shaded).

All three methods produce the same results when the the spectrum contains approval sets of the same length.

3 A Pairwise Agreement Proportion Condition

The left, right, and midpoint median methods will lead to the same results, when approval sets are all the same length, so the left median method will be used unless specified otherwise.

By our definition of the median voter, consider the case where we have two evenly sized collections of approval sets, A and B, such that each approval set in A pairwise intersects and each approval set in B pairwise intersects. Also, consider that there exists one isolated approval set between A and B. Using a society like Figure 7, we can show that there exists some relationship between the pairwise agreement proportion and the location of the winning platform.

Figure 7: A society of approval sets where the median voter interval does not contain an approval winner.

We can find the pairwise agreement proportion, α, for all linear societies such that two equally sized collections of sets that agree on a platform exist on the extrema of the spectrum and there exists one isolated approval set between the outer sets.

Let m be the number of approval sets that share a common platform on the extrema.

$$\alpha = \frac{2\binom{m}{2}}{\binom{2m+1}{2}}.$$

$$= \frac{m-1}{2m+1}$$

$$\lim_{m\to\infty} \frac{m-1}{2m+1} = 0.5$$

We can see from the proceeding example where the median voter interval does not contain the approval winner, the pairwise agreement proportion is below a certain threshold. This suggests our next result: that if the pairwise agreement proportion were high enough, we can show that the median voter interval does capture the approval winner.

Theorem 3.1. *Let all approval sets be equal length. If the pairwise agreement proportion is at least $\frac{3}{4}$, then the median voter interval will contain the approval winner.*

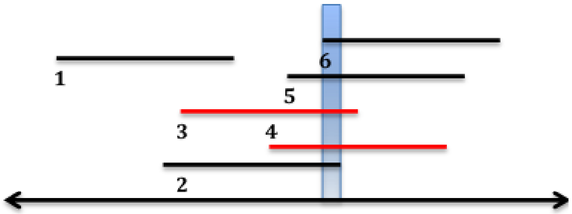


Figure 8: A society with a pairwise agreement proportion of $\frac{4}{5}$.

Proof. We use the Fractional Helly Theorem which states that if at least $\alpha\binom{n}{2}$ of the sets pairwise intersect, then a common point is shared by at least $(1-(1-\alpha)^{(1/2)})n$ of the sets, where α is the pairwise agreement proportion[1].

If α is greater than or equal to $\frac{3}{4}$, then it follows that at least $\frac{1}{2}$n sets share a common point.

The next lemma gives us the desired conclusion. We separate it as a lemma because it may have independent interest. □

Lemma 3.2. *Let all sets be equal length. If at least half of the the approval sets share a common point, then the median voter interval will contain the winner.*

For ease of description, we will call a collection of pairwise intersecting sets on a line a *clique*. (We borrow this terminology from graph theory, since the graph of intersections of these sets will have a clique of pairwise intersecting vertices.)

Proof. Let there be $2k+1$ approval sets, where k is a non-negative integer, and at least half of the sets share a common point. Then, there exists some clique, C, containing at least half of approval sets such that C is the largest collection of sets that all pairwise intersect one another.

Let us say that the median voter interval exists somewhere in the spectrum. Using our median methods, the median voter interval is the median number of the numbered sets when the number of approval sets is odd, so the median is the $(k+1)-th$ element. So, there exists k elements with left endpoints left of the median voter interval's left endpoint and k elements with left endpoints right of the median voter interval's left endpoint. $k < \frac{n}{2}$.

Assume the median voter interval is not contained in C. C is defined as containing at least half of the approval sets and the highest number of sets that share a common point. Then the median voter interval is contained in a clique of equal size as C or a clique smaller than C. If the median voter interval is contained in a clique of equal size as C, then the median voter interval contains the approval winner. By definition, there are an even amount of sets on each side of the median voter interval. Thus, all of the approval sets contained in C must be on one side of the median voter interval. There could not possibly be an even amount of sets on both sides of the median voter interval and therefore a contradiction must exist. □

This theorem does not apply to societies that have uneven length approval sets. See Figure 9.

4 A Clique Size Condition

Now we consider the size of the clique that contains the approval winner relative to the number of approval sets not that do not contain the approval winner.

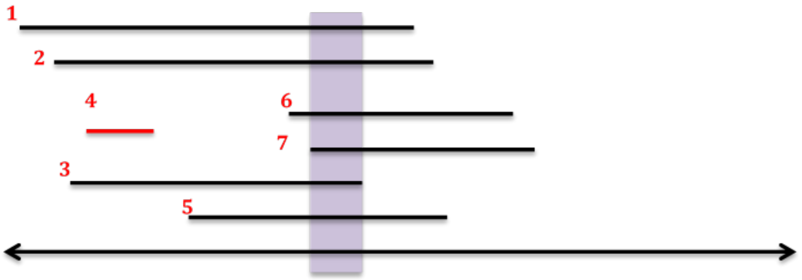


Figure 9: A linear society of varying length approval sets with a pairwise agreement proportion over $\frac{3}{4}$, but the median voter interval I_4 does not contain the approval winner

Theorem 4.1. *Let all approval sets be equal length. Let C be the collection of approval sets belonging to the largest clique. Let L be the collection of approval sets on the left side of C such that the right endpoint of every approval set contained in L is left of the approval winner contained in C. Let R be the collection of approval sets on the the right side of C such that every left endpoint contained in R is right of the approval winner contained in C. The median voter interval is contained in the largest clique C if and only if $\big||L| - |R|\big| \leq |C|$, so the median voter interval contains the winner if and only if $\big||L| - |R|\big| \leq |C|$.*

Proof. Let all approval sets be equal length. Given $\big||L| - |R|\big| \leq |C|$ and $|L| + |C| + |R| = 2k+1$ where k is a non-negative integer. Then $|L| \leq |C|+|R|$ and $|R| \leq |C|+|L|$. We will show the median voter interval is contained in C if $\big||L| - |R|\big| \leq |C|$.

Assume the median voter interval is contained in L. Then there exists at least $k+1$ sets in L because there must be k elements on each side of the median voter interval and L also contains the median voter interval. This statement implies that $|L| > |R| + |C|$ because $k + 1$ is greater than half of the approval sets, but $|L| \leq |R| + |C|$, so this cannot exist. Set R cannot contain the median voter interval by the same argument. Therefore the median voter interval must belong to C if $\big||L| - |R|\big| \leq |C|$. Considering $|R| \leq |C| + |L|$ will lead to the same result.

We will show that the median voter interval is not contained in C if $\big||L| - |R|\big| > |C|$.

Given that $\big||L| - |R|\big| > |C|$ and $|L| + |C| + |R| = 2k+1$, where k is a non-negative integer. Then $|L| > |C| + |R|$ and $|R| > |C| + |L|$. Assume the median voter interval is contained in C. Then, less than half of the elements are contained in L, but L contains greater than half of the elements because $|L| > |C| + |R|$. So, if the median is contained in C, then $|L| > |C| + |R|$ cannot be satisfied. Therefore, the median voter interval is not contained in C if $\big||L| - |R|\big| > |C|$. The result is the same when considering $|R| > |C| + |L|$. ☐

However, this theorem does not hold for unequal length approval sets.

Theorem 4.2. *If $\big||L| - |R|\big| \leq |C|$, the approval winner will be contained in the convex hull of the left and right median intervals.*

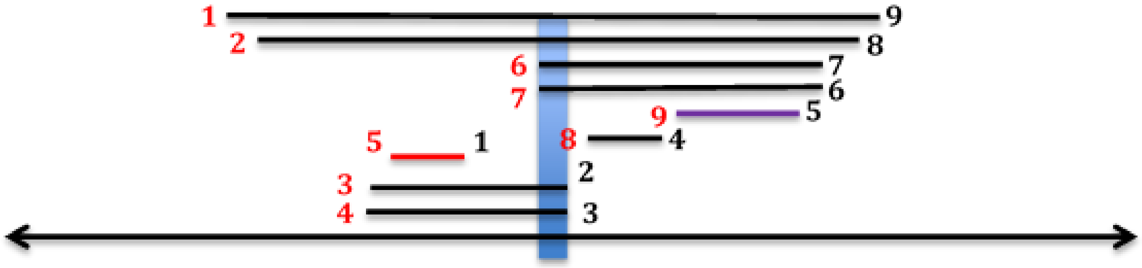


Figure 10: A society of varying length approval sets where neither the left or right median intervals contain the approval winner.

Proof. Given $\big||L| - |R|\big| \leq |C|$ and $|L| + |C| + |R| = 2k + 1$ where k is a non-negative integer. Then, $|L| \leq |C| + |R|$. So, $|L| \leq k$ and $|R| \leq k$. Because R is defined as all approval sets with a left endpoint to the right of the approval sets contained in C, then if the left endpoint method is used then the sets contained in R will be labeled sequentially and since $|R| \leq k$, the left median interval cannot be contained in R. So, the left median must be contained in either L or C. We can use a similar argument to show that the right median must be contained in either R or C. Because the left median interval is either contained in L or C and the right median interval is contained in R or C, then the convex hull of the left and right median intervals will always contain a set in C, and therefore the approval winner.

□

5 A Clique Size Condition for Agreement Graphs

In Theorem 4.2, we developed a condition that guarantees the approval winner exists in the convex hull of the left endpoint and right endpoint median voter intervals. The convex hull of these sets can be large, compared to the average length of approval sets in the data. After attempts at finding a smaller interval that would always contain the winner, we have concluded that while using interval plots, the convex hull of the left endpoint median and right endpoint median is the smallest interval that is guaranteed to contain the winner with sets of any length. Is there a different way to plot the approval sets to guarantee the winner in a smaller interval? Also, does there exist a single median method where the median will capture the approval winner? This motivates our transition from considering the location of endpoints to considering the location of pairwise intersections.

Agreement graphs allow for a labeling system that considers the location of the pairwise intersections while ignoring the length of the approval intervals. We form agreement graphs by plotting each voter's approval interval as a vertex and the intersection of voter approval intervals as edges connecting the vertices. Each voter will be plotted from the left-most left end point to the right-most left endpoint, but not necessarily labeled in this manner. A clique is a set of vertices such that every pair of vertices is connected by an edge. The element of each clique with the left-most intersection will be labeled with the lowest number of the clique, the element rightmost intersection of each clique will be labeled with the highest number of the clique. The other elements within a clique will be labeled with arbitrary, but non-repeating, numbers such that each element is labeled sequentially.

We have defined the median voter's approval set as $I_{\frac{n+1}{2}}$ when n is odd and the union of $I_{\frac{n}{2}}$ and $I_{\frac{n}{2}+1}$ when n is even. These new labeling conditions will allow for more consistency in location of the median voter interval in regard to the approval winner.

Define C as a collection of pairwise intersecting sets that have the highest agreement number on the spectrum. Thus the size of C, or $|C|$, is equivalent to the agreement number.

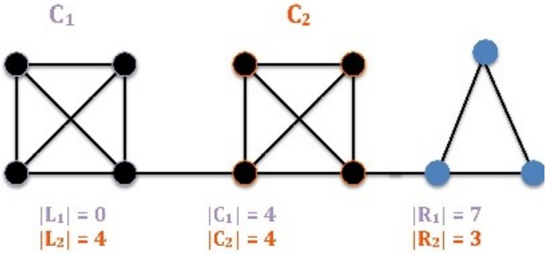


Figure 11: A Society S with two approval winners

Theorem 5.1. *Choose C to be the approval winner with the minimum difference between $|R|$ and $|L|$. The median voter interval contains the winner if and only if $\big||L| - |R|\big| \leq |C|$.*

Proof. We will first show: if $\big||L| - |R|\big| \leq |C|$ then the median voter interval contains the winner.

Assume $\big||L| - |R|\big| \leq |C|$ and the median voter interval does not belong to C. Thus the median voter interval set belongs to L or R. Assume the median voter interval belongs to L. By definition of the median voter interval, there must be an equal amount of sets on both of its sides. Every set to the left of a set contained in L must also be contained in L. Since L contains all of the sets to the left of the median voter interval and the median voter interval itself, $|L| > |C| + |R|$ but $|L| \leq |C| + |R|$, thus a contradiction exists and the median voter interval can not be contained in L. If we assume the median voter interval belongs to R, then without loss of generality we can assume the median voter interval can not be contained in R. Then the median voter interval must contain be contained in C and therefore contains the approval winner.

We next show: if the median voter interval contains the winner then $\big||L| - |R|\big| \leq |C|$.

Assume that $\big||L| - |R|\big| > |C|$ and assume the median voter interval belongs to C. By definition of the median voter interval, there must be an equal number of sets on either side of the median voter interval. Thus $|C| + |R|$ is greater than half of the sets. This contradicts that $|L| > |C| + |R|$. Therefore, if the median voter interval contains the winner then $\big||L| - |R|\big| \leq |C|$.

□

6 Conclusion

We found when the pairwise agreement proportion is $\frac{3}{4}$ the median voter interval will always contain the approval winner when the approval sets are equal length. Also for equal length approval sets, if the number of voter intervals containing the approval winner is high enough, then the median voter interval will contain the approval winner. Unlike the Median Voter Theorem, we cannot say the candidate who places himself towards the middle of the political spectrum will win the election. This is an expected result because approval voting is intended to be strategy-proof [2]. However, we were able to define conditions where the median voter interval will contain the approval winner therefore there exists an analogue of the Median Voter Theorem.

7 Acknowlegdements

We would like to thank our advisors Dr. Duane Cooper, Dr. Francis Edward Su, and Mr. Daniel Eckhardt. We would also like to thank The Mathematical Sciences Research Institute, NSA, and NSF for sponsoring us.

References

[1] Deborah E. Berg, Serguei Norine, Francis Edward Su, Robin Thomas, and Paul Wollan, *Voting in agreeable societies*, Amer. Math. Monthly **117** (2010), no. 1, 27–39. MR 2599465 (2011i:91065)

[2] Steven J. Brams and Peter C. Fishburn, *Approval voting*, American Political Science Review **72** (1978), 831–847.

[3] Tyler Cowen, *Why politics is stuck in the middle*, New York Times (2010).

A Matroid Generalization of Sperner's Lemma

Gabriel Passamani Andrade
University of Massachusetts, Amherst

Andrés Rodríguez Rey
Universidad de los Andes

Alberto J. Ruiz Sandoval
University of Puerto Rico, Río Piedras

July 25, 2015

This work was conducted during the 2015 Mathematical Sciences Research Institute Undergraduate Program (MSRI-UP), supported by grants from the National Science Foundation (DMS 1156499) and the National Security Agency (H98230-15-1-0039). We also thank supervisors Duane Cooper, Francis Edward Su, and Mutiara Sondjaja.

Abstract

In a 1980 paper, Lovász generalized Sperner's lemma for matroids. He claimed that a triangulation of a d-simplex labeled with elements of a matroid M must contain at least one "basis simplex". We present a counterexample to Lovász's claim when the matroid contains loops and provide a necessary condition such that Lovász's generalization holds. Furthermore, we show that under some conditions on the matroids, there is an improved lower bound on the number of basis simplices. We present further work to sharpen this lower bound by looking at M's lattice of flats and by proving that there exists a group action on the simplex labeled by M with $\mathcal{S}_n$.

1 Introduction

Sperner's lemma is a claim about the triangulations of simplices, which is noted for its equivalence to the Brouwer Fixed Point theorem. It states that given a triangulation T of a d-simplex S and a *Sperner labeling* on T, there must exist at least one *fully labeled Sperner simplex.* In [1], Lovász extends Sperner's lemma for *matroids*, a construct that generalizes the concept of linear independence. His extension states the following:

Theorem 1.1 (Lovász, 1980). *Let S be a d-simplex, K a simplicial subdivision of S and assume that a matroid of rank $d+1$ is defined on the vertices of K. Assume furthermore that the vertex-set $V(S)$ of S is independent in the matroid and that for each $A \subseteq V(S)$, those vertices of K on the face spanned by A are contained in the flat of the matroid spanned by A. Then K has a simplex whose vertices form a basis.*

This theorem asserts that there must exist at least one *basis simplex*. We found counterexamples to this theorem when the matroids contain *loops*, i.e., singleton dependent sets. We show that if we add to the hypothesis of Lovász's theorem the condition that the matroid used in the labeling is *loopless*, then the conclusion of the theorem holds.

In addition to understanding Lovasz's result, the main motivation of our project is to improve the lower bound on the number of basis simplices that we can guarantee in a matroid-labeled triangulation. That is, under what conditions on the matroids can we assure the existence of more than one basis simplex.

We give the necessary background on Sperner's lemma and matroids in Section 2. In Section 3, we formalize our corrections to Lovász's paper and provide an improved lower bound for the one dimensional case. Additionally, we provide different approaches to solve this problem in higher dimensions. Section 4 of this paper returns to Lovász's constructions and highlights a group action on the vertices of a triangulation labeled by a matroid. Finally, Section 5 is dedicated to remarks, conjectures, and future work.

2 Background

At the heart of our question is Sperner's lemma and a number of constructs from matroid theory. In this section we will define and discuss the necessary notions. Furthermore, we introduce Lovász's results that bridges these concepts, provide a correction to his paper, and prove this new claim.

2.1 Sperner's Lemma

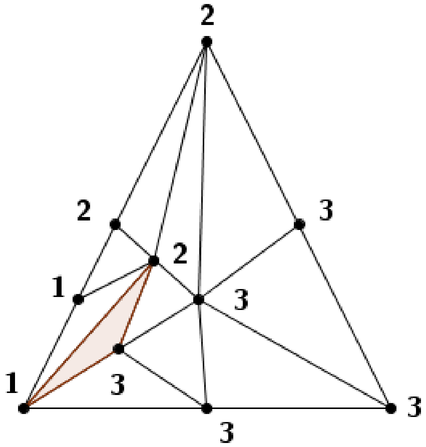


Figure 1: A Sperner labeling on a 2-simplex

We start our introduction of Sperner's lemma with a motivating example. Consider the example shown in Figure 1. For this triangulation, we start by labeling the three main vertices of the triangle distinctly by $1, 2, 3$. Then, for any vertices on an edge of the main triangle we impose the label of one of the vertices at the endpoints of the edge. For example, on the edge labeled by $1, 2$ in the main triangle we have the two vertices in between labeled arbitrarily by either 1 or 2. The vertex in between the edge labeled by 2 and 3 on the main triangle is labeled by 3 and the vertex on the edge $1, 3$ is labeled by 3; although 1 is a valid labeling. Any vertex inside the main triangle can be labeled by any element in $\{1, 2, 3\}$.

This type of labeling is what is known as a Sperner labeling for a 2-simplex. What Sperner's lemma asserts is that we have an odd number of fully labeled simplices and that there exists at least one. Fully labeled triangles are triangles labeled distinctly by elements in $\{1, 2, \ldots, d + 1\}$ for a d-simplex; in this case, a triangle labeled by $1, 2$ and 3. Going back to Figure 1, the shaded simplex is the only fully labeled Sperner simplex in this triangulation.

In general, a Sperner labeling on a d-dimensional simplex S with a triangulation T is a labeling that satisfies the following rules:

- The vertices of the main simplex S are distinctly labeled by all the elements in $\{1, 2, \ldots, d + 1\}$.
- The vertices located on any k-dimensional face $\{a_1, a_2, \ldots, a_{k+1}\}$ of the main simplex are labeled by any element in $\{a_1, a_2, \ldots, a_{k+1}\}$.

Then, Sperner's lemma states the following:

Lemma 2.1 (Sperner's Lemma)**.** *Any Sperner-labeled triangulation of a d-simplex must contain an odd number of fully labeled elementary d-simplices. In particular, there is at least one.*

2.2 Matroid Theory

Matroids are mathematical objects that capture the notion of linear independence in vector spaces. We follow Oxley's [2] definition and notations for our introduction to matroids.

Definition 2.2. A matroid is a pair $M = (E, \mathcal{I})$ consisting of a finite set E called the ground set and a collection of subsets $\mathcal{I}$ from E that satisfy the following conditions:

- $\emptyset \in \mathcal{I}$
- If $I \in \mathcal{I}$ and $I' \subseteq I$, then $I' \in \mathcal{I}$
- If I_1 and I_2 are in $\mathcal{I}$, and $|I_1| < |I_2|$ then there is an element $e \in I_2 - I_1$ such that $I_1 \cup e \in \mathcal{I}$.

It is useful to consider some examples. In the first example below, we have a matroid of vectors in $\mathbb{R}^2$ where the independent sets are sets of linearly independent vectors. The second example involves a *graphic matroid*, a type of matroids which we will use in further examples throughout the paper.

Example 1. Consider the following matrix whose columns are vectors in $\mathbb{R}^2$:

$$\begin{array}{ccccc} e_1 & e_2 & e_3 & e_4 & e_5 \\ \end{array}$$
$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Let $E = \{e_1, \ldots, e_5\}$ denote the set consisting of the five column vectors of the above matrix and let $\mathcal{I}$ denote the collection of all subsets of E which forms linearly independent sets of the vectors in $\mathbb{R}^2$.

Then, it is not hard to see that $M = (E, \mathcal{I})$ is a matroid. The empty set of vectors is defined to be linearly independent, so $\emptyset$ is in $\mathcal{I}$. For $I \in \mathcal{I}$, a set of linearly independent vectors, then any subset $I' \subseteq I$ is also a linearly independent set of vectors, so I' must be an element of $\mathcal{I}$ as well. We leave it to the reader to verify that the third condition in Definition 2.2 also holds for M.

Example 2. Consider the graph in Figure 2 below. Let E denote the set of edges in the graph and let $\mathcal{I}$ denote the collection of all sets of edges that do not form a cycle. That is, a set $I \subseteq E$ of edges is an independent set if it does not form a cycle. Otherwise, if $X \subseteq E$ forms a cycle, we say that it is a dependent set.

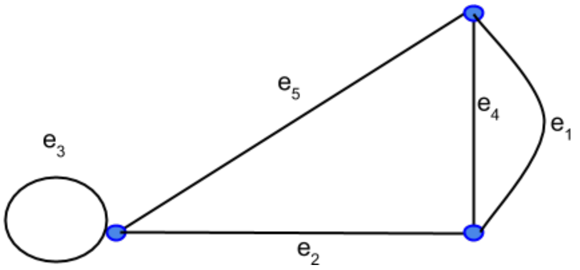


Figure 2: Example of a graphic matroid.

We claim that $M = (E, \mathcal{I})$ is a matroid. The empty set of vectors does not form a cycle, so $\emptyset$ is in $\mathcal{I}$. If I is a set of edges that do not form a cycle, then any subset $I' \subseteq I$ must not form a cycle, which means that $I' \in \mathcal{I}$. Again, we leave it to the reader to verify that the third condition also holds for M, thereby showing that M is a matroid. Such a matroid, whose ground set is the set of edges in a given graph and whose independent sets are the sets of edges with no cycles, is called a *graphic matroid*.

A minimal dependent set in an arbitrary matroid M will be called a *circuit* of M and we shall denote the set of circuits of M by $\mathcal{C}$. If a two-element set $\{m_1, m_2\}$ form a circuit in M, then m_1 and m_2 are *parallel* in M. The *parallel class* of an element $m \in E$ is then the set of all elements in E that are parallel to m.

We know from linear algebra that any set of n linearly independent vectors in $\mathbb{R}^n$ will span all of $\mathbb{R}^n$ and we call this set a *basis* of $\mathbb{R}^n$. Another useful concept is the *rank* of a set of vectors. We know that any basis in $\mathbb{R}^n$ will be of rank n, we also know that adding any other vector to a basis will make the set dependent but the rank will remain the same. This suggests a generalization of basis and rank for matroids:

- A basis of a matroid $M = (E, \mathcal{I})$ is a maximal independent set of E.
- The rank $r(M)$ of a matroid M is the size of a basis in M. Then, the rank of a subset $X \subseteq E$ is the size of the largest independent set in X.

Formally, let $M = (E, \mathcal{I})$ be a matroid, suppose that $X \subseteq E$ and that $I|X = \{I \subseteq X : I \in \mathcal{I}\}$. We define the *rank* $r(X)$ of X to be the size of a basis B of $M|X$. That is, a function $r : 2^E \to \mathbb{Z}^+ \cup 0$ is the rank function of a matroid on E if and only if r satisfies the following conditions:

- If $X \subseteq E$, then $0 \leq r(X) \leq |X|$.
- If $X \subseteq Y \subseteq E$, then $r(X) \leq r(Y)$.
- If X and Y are subsets of E, then $r(X \cup Y) + r(X \cap Y) \leq r(X) + r(Y)$.

Two concepts of utmost importance in our paper are the ideas of closures and flats. Let cl be the function from 2^E into 2^E defined for all subsets $X \subseteq E$ by

$$\mathrm{cl}(X) = \{x \in E : r(X \cup x) = r(X)\}.$$

This function is called the *closure operator* of M. A set $X \subseteq E$ is called *flat* if $X = \mathrm{cl}(X)$. Throughout this paper we will denote the flat of any set $X \subseteq E$ as $\langle X \rangle$.

2.3 Matroid and Induced Sperner's Labelings

In his paper, Lovász described a method of labeling a triangulation by elements of a matroid in a way that emulates Sperner labeling. In the case of a triangle's triangulation, we label the main vertices of said triangle with the elements of a basis of a rank 3 matroid. Then we use the flats of the basis elements labeling the main triangle to label vertices on the edges they define. Finally we label internal vertices with any remaining elements. We see this in Figure 3.

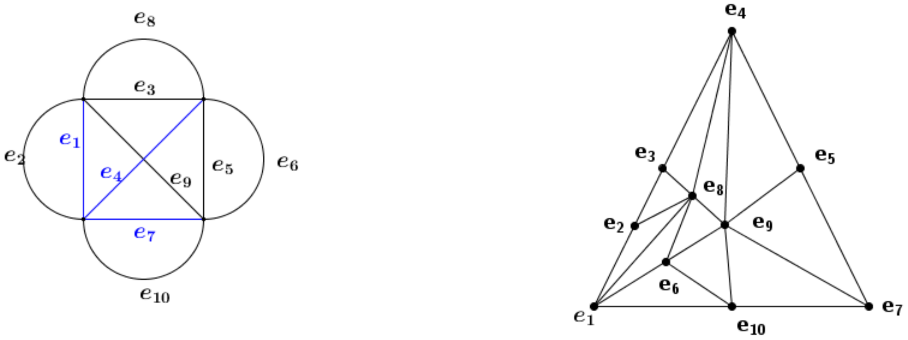


Figure 3: A matroid and its labeling on a triangulation.

In general, a labeling of this kind is defined as follows:

Definition 2.3. Let S be a d-simplex and T a triangulation of S. Assume that there exists a matroid of rank $d+1$ defined on the vertices of T. We define a *matroid labeling* on T as a labeling on the vertices of T that satisfy the following conditions:

- The vertex-set $V(S)$ of S forms a basis in the matroid.
- For each $A \subseteq V(S)$, the vertices of T that are in the face spanned by A are contained in the flat of the matroid spanned by A.

Any simplex in the triangulation whose vertices are labeled by a basis is a *basis simplex*. It is important to note that for an arbitrary matroid and triangulation there is not necessarily a proper matroid labeling. It is vital that the triangulation have the same number of vertices as the matroid has elements and that the flats of the elements labeling the main edges have enough elements to label the vertices on those edges.

Due to its similarity to Sperner labeling, matroid labeling can be converted to Sperner labeling with relative ease. We simply look at the flats of the basis elements labeling the main vertices of our simplex and impose a labeling from this. A Sperner labeling brought on from a matroid labeling is said to be induced by it.

Definition 2.4. Suppose S is a d-dimensional simplex and T corresponds to a triangulation of S. Let $M = (E, \mathcal{I})$ be a rank $(d+1)$-matroid with basis $\{a_1, ..., a_{d+1}\}$ and $F_i = \langle a_1, \ldots a_i\rangle - \langle a_1 \ldots a_{i-1}\rangle$ for all $i \in [d+1]$. A Sperner labeling on T induced by a matroid M is a labeling that satisfies the following properties:

- There is a matroid labeling on T.
- For all $v \in T$, v is labeled by some i corresponding to the flat F_i of v.

To highlight this process we shall look at the matroid and 2−simplex from earlier. Consider Figure 4.

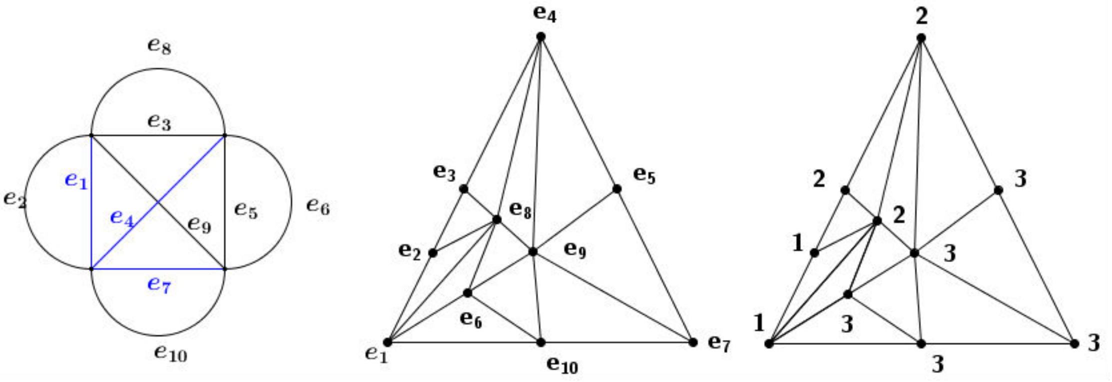


Figure 4: A matroid, its labeling on a triangulation, and its induced Sperner labeling.

Notice the vertices corresponding to the elements in the flat of e_1 are labeled with 1 in the induced Sperner labeling—namely e_1 and e_2. Then the vertices with the unlabeled elements in the flat of e_1 and e_4 are labeled with 2—namely e_3, e_4, and e_8. Lastly all of the remaining vertices are labeled with 3's in the induced Sperner labeling.

3 Results

3.1 Correction to Lovász's Results

The motivation for this paper is a result published by Lovász [1] in which he asserts the following:

Let K be a simplicial complex which is a d-dimensional manifold. Also assume that a matroid of rank $d+1$ is defined on the set of vertices of K. If K has a simplex whose vertices form a basis of the matroid, then it has at least two.

In the proof, Lovász uses the following procedure:

Assume that $(a_1, \ldots, a_{d+1})$ is the unique simplex which is a basis. Let F_i, denote the flat spanned by $\{a_1, \ldots, a_i\}$. Let Q denote the set of all sequences $(x_1, \ldots, x_d)$ of elements of the matroid such that

$$x_1 \in F_1, \qquad x_i \in F_i - F_{i-1} \tag{1}$$

Then, Lovász claims the set $\{x_1, \ldots, x_d\}$ is automatically independent in the matroid. But this is not necessarily the case.

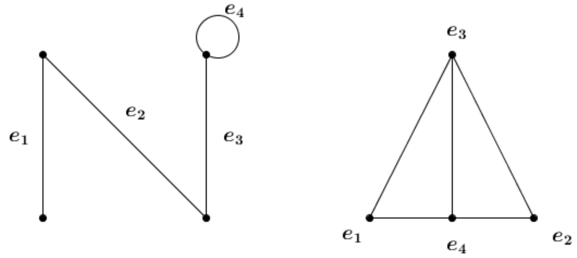


Figure 5: Lovász claims that any matroid labeling must have at least one basis simplex. Loops, which are dependent on their own, make this untrue.

Consider figure 5, in this figure the edge e_4 is a loop and is therefore dependent to itself and every other element in the matroid. Which means that in the construction in 1 every set in Q will contain a loop and will therefore be dependent. This contradicts the claim and suggests that we make the following addendum to the statement:

Theorem 3.1. *Let K be a simplicial complex which is a d-dimensional manifold. Also assume that a* ***loopless*** *matroid of rank $d+1$ is defined on the set of vertices of K. If K has a simplex whose vertices form a basis of the matroid, then it has at least two.*

We added the condition that our choices of matroids are *loopless matroids*. Loopless matroids are matroids that do not contain dependent singletons in the ground set. By restricting our choice of matroids to loopless matroids, we can be certain that the construction in 1 will yield a set Q of independent subsets of E.

Next we state the corollary presented in [1]:

Let S be a d-simplex, K a simplicial subdivision of S and assume that a matroid of rank $d+1$ is defined on the vertices of K. Assume furthermore that the vertex-set $V(S)$ of S is independent in the matroid and that for each $A \subseteq V(S)$, those vertices of K on the face spanned by A are contained in the flat of the matroid spanned by A. Then K has a simplex whose vertices form a basis.

The triangulation in figure 5 satisfies the conditions mentioned in the foregoing corollary but leads to an erroneous conclusion. There is no basis simplex in the triangulation and this is due to the fact that there is a loop in the matroid. If we then restrict our choices to loopless matroids we can prove the corollary holds.

Corollary 3.2. *Let S be a d-simplex, K a simplicial subdivision of S and assume that a **loopless** matroid of rank $d+1$ is defined on the vertices of K. Assume furthermore that the vertex-set $V(S)$ of S is independent in the matroid and that for each $A \subseteq V(S)$, those vertices of K on the face spanned by A are contained in the flat of the matroid spanned by A. Then K has a simplex whose vertices form a basis.*

To prove this corollary we first prove the following lemma:

Lemma 3.3. *Let S be a d-simplex, T a triangulation of S and $\mathcal{P}(T)$ the Sperner labeling induced by a matroid $M = (E, \mathcal{I})$ of rank $d+1$ defined on the vertices of T. If $\{v_1, v_2, \ldots, v_{d+1}\}$ are the vertices of a fully labeled Sperner simplex on $\mathcal{P}(T)$ then $\{v_1, v_2, \ldots, v_{d+1}\}$ corresponds to a basis $\{b_1, b_2, \ldots, b_{d+1}\}$ on M.*

Proof. Suppose $\{v_1, v_2, \ldots, v_{d+1}\}$ are vertices that form a fully labeled Sperner simplex on $\mathcal{P}(T)$. Let F_i denote the flat that indexes the vertex $w_i \in T$ and let $\mathcal{P}(b_i)$ denote the element in E that is labeled by $v_i \in \{v_1, v_2, \ldots, v_{d+1}\}$. Then v_i is labeled by an element b_i such that $b_i \in F_i - (F_{i-1} \cup F_{i-2} \cup \ldots \cup F_1)$, that is, b_i is independent to any element in $F_{i-1}, F_{i-2}, \ldots, F_1$. Since each element in $\{v_1, v_2, \ldots, v_{d+1}\}$ is labeled differently, then b_i is independent to any b_k such that $\mathcal{P}(b_k) = v_k \in \{v_1, v_2, \ldots, v_{d+1}\}$. This implies that the set $\{b_1, b_2, \ldots, b_{d+1}\}$ is an independent set of size $d+1$. Therefore, the set $\{v_1, v_2, \ldots, v_{d+1}\}$ corresponds to a basis $\{b_1, b_2, \ldots, b_{d+1}\}$ in M. $\square$

Proof of Corollary 3.2. By lemma 3.3 and Sperner's lemma, the corollary follows. $\square$

3.2 Lower Bound on Basis Simplices

By virtue of the conditions on labeling a triangulated simplex with a matroid, certain elements are limited on what vertices they can label. Due to this and their relative independence from the ordered basis we can see that certain matroids demand a minimum number of basis simplices. We will first explore this in the one dimensional case.

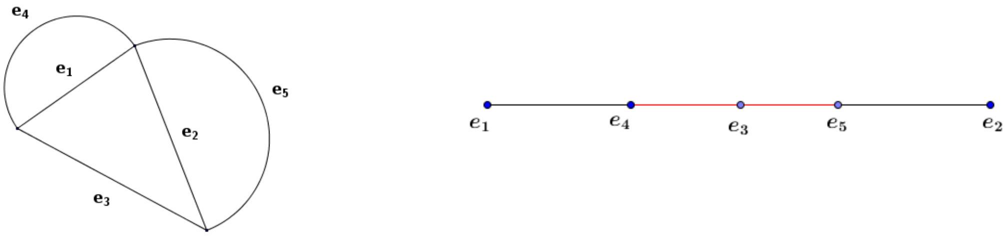


Figure 6: One dimensional example of a matroid labeling.

Consider the matroid and its corresponding triangulated 1−simplex in Figure 6. By construction, the vertex labeled by the element e_3 can be swapped with the label of any other internal vertex on the 1−simplex. It is easy to check that e_3 is dependent to both e_1 and e_2 but independent to either of those elements (and their parallel elements) individually. Therefore e_3 forms a basis with any element in the parallel classes of e_1 or e_2. Additionally, any element in the parallel class of e_1 will form a basis with any element in the parallel class of e_2 and vice versa. Thus, regardless of how we scramble the labels of the internal vertices on this 1−simplex there will be at least two basis simplices. We will now generalize this example.

Theorem 3.4. *Let M be a matroid of rank 2 that has a circuit of size 3 and let S be a 1-simplex. Then, for any triangulation T of S that is labeled by M there are at least two basis simplices.*

Proof. Without loss of generality we shall refer to elements in the parallel classes of the basis elements that label the main vertices of T as P_1 and P_2 and to P_3 as any element that forms a circuit of size 3 with elements of P_1 and P_2. We now have four cases: either an element of P_3 has a fellow P_3 element and either a P_1 or P_2 element adjacent to it, that element has only P_3 elements adjacent to it, both an element in P_1 and P_2 are adjacent to it, or that element has only elements from P_1 or P_2 adjacent to it. In the first two cases we simply note that when an element of P_3 has a fellow P_3 element adjacent to it that adjacent element must fall within the cases as well. Our triangulation is finite, so it follows that either a P_1 or P_2 element must eventually be adjacent to one of these adjacent P_3 elements. As such, we can treat these chains of P_3 elements as if they were a single element and we fall into the remaining two cases.

Case 1: Suppose the element(s) in P_3 are surrounded on both sides by elements in P_1 *and* P_2. As mentioned, said element(s) form a basis with both the elements in P_1 and P_2. Thus there are at least two basis simplices.

Case 2: By a nearly identical argument, suppose the element(s) in P_3 are surrounded on both sides by elements in either P_1 *or* P_2. Regardless of which parallel class they are adjacent to, elements in P_3 form a basis with elements in either P_1 or P_2. Therefore there are at least two basis simplices again. □

Should a matroid of rank 2 without a circuit of size 3 be used in labeling a triangulation it simply falls into a case of Sperner's lemma for 1−simplices.

When moving to higher dimensional simplices we need to be weary of overwhelmingly large parallel classes that "smother" our triangulation. To see what this means consider Figure 7.

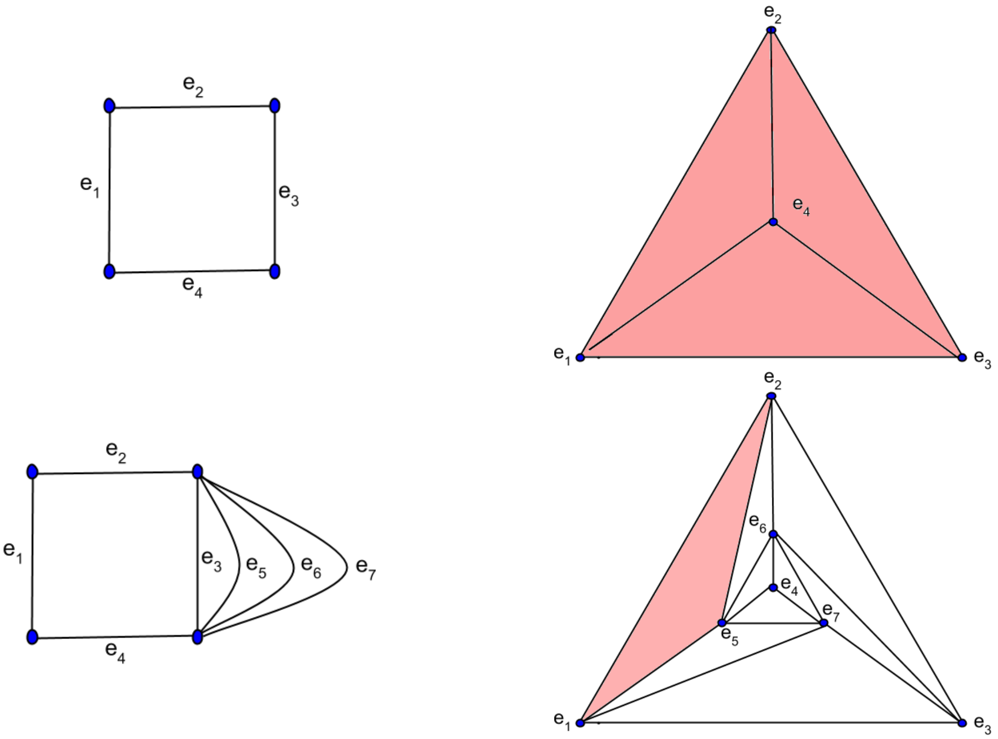


Figure 7: A matroid with a large parallel class and a "smothered" triangulation.

The huge parallel class of e_3 allows us to surround elements that would otherwise form multiple basis simplices and limit the amount of basis triangles that appear. While dealing with general matroids we have to worry about having parallel classes that run rampant.

4 Further Results

4.1 Lattice of Flats

As it was shown in the previous section, the Sperner labeling induced by a matroid M depends on the order of the basis we are fixing. This suggests that there should be an action of the symmetric group on the elements of the basis, and so, an action on the Sperner labeling. To show this, we need an auxiliary structure from algebraic combinatorics:

Definition 4.1. A *poset* P is a finite set, also denoted P, together with a binary relation denoted $\leq$ satisfying the following axioms:

- (reflexivity) $x \leq x$ for all $x \in P$.

- (antisymmetry) If $x \leq y$ and $y \leq x$, then $x = y$.
- (transitivity) If $x \leq y$ and $y \leq z$, then $x \leq z$.

For our purpose we will be interested in a particular poset:

Definition 4.2. Let $n \in \mathbb{N}$ and $E = \{1, 2, ..., n\}$ be a set. We call $\mathfrak{B}_n = (\mathcal{P}(E), \subset)$ the *boolean algebra* of n elements, where 2^E denotes the power set of E.

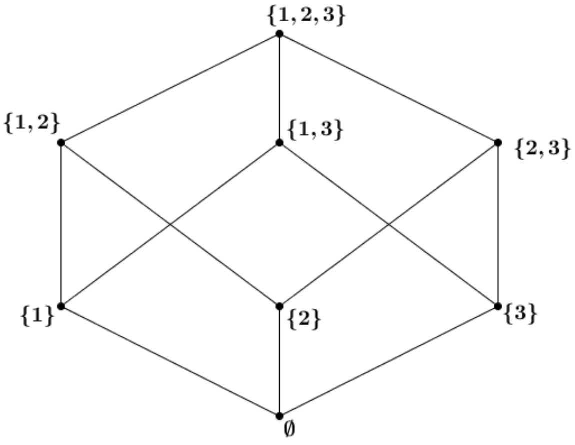


Figure 8: $\mathfrak{B}_3$.

In Figure 8 we have what is called the *Hasse-diagram* of $\mathfrak{B}_3$. Now let $\mathcal{S}_n$ be the symmetric group on n elements. For any $\sigma \in \mathcal{S}_n$ and $A = \{i_1, ..., i_m\} \in \mathfrak{B}_n$ we define $\sigma A = \{\sigma(i_1), ...\sigma(i_m)\}$ where $m \leq n$. It's easy to check that this is indeed an action of $\mathcal{S}_n$ on $\mathfrak{B}_n$. From now on we are going to drop the brackets and commas when we talk about elements of $\mathfrak{B}_n$, for example: $\{1, 2, 3\}$ will be denoted by 123.

A *lattice* is a poset $\mathcal{L}$ such that, for every pair of elements, the least upper bound and the greatest lower bound of the pair exists. Formally, if x and y are arbitrary elements of $\mathcal{L}$, then $\mathcal{L}$ contains elements $x \vee y$ and $x \wedge y$, the *join* and *meet* of x and y respectively, such that:

- $x \vee y \geq x, x \vee y \geq y,$ and if $z \geq y,$ then $z \geq x \vee y$; and
- $x \wedge y \leq x, x \wedge y \leq y,$ and if $z \leq y,$ then $z \leq x \wedge y$.

In the case of $\mathfrak{B}_3$ the operations join and meet are union and intersection, respectively. It is easy to see from the Hasse-diagram that $\mathfrak{B}_3$ is in fact a lattice.

If M is a matroid, let $\mathcal{L}(M)$ denote the sets of flats of M ordered by inclusion. It's easy to see that this is a partially ordered set. Additionally it can be endowed with the structure of a lattice, as stated by the following theorem.

Theorem 4.3. *$\mathcal{L}(M)$ is a lattice and, for all flats X and Y of M,*

$$X \wedge Y = X \cap Y \text{ and } X \vee Y = cl(X \cup Y).$$

4.2 Group Action

A Sperner labeling induced by a matroid labeling as described in **section 2.3**, is fully reliant on the order in which we consider the chosen basis. By simply changing the order of the basis we induce vastly differing Sperner labels with varying fully labeled elementary simplices. By the lemma presented in **section 4.2.2**, this means that the Sperner triangles may correspond to different basis simplices altogether. We will show that there exists a group action on the induced Sperner labelings as we permute the order of the basis elements.

4.2.1 $\mathcal{S}_3$ as a Group Action

Our goal for this section is to extend the action defined in **section 4.1** to an action on the Sperner labeling by permuting the order of the basis, first we do it for a given example and then generalize in the next subsection. Let M be the matroid depicted in Figure 9 and choose $B = \{e_1, e_4, e_7\}$, an ordered basis of M. Notice that in $\mathcal{L}(M)$, Figure 10, restricting ourselves to the flats that include $\langle e_1 \rangle$, $\langle e_4 \rangle$, or $\langle e_7 \rangle$ gives a sublattice P_B that looks as a boolean algebra (Figure 11). This suggests that there is an isomorphism between P_B and $\mathfrak{B}_3$, and since the basis B was ordered, the most natural isomorphism ϕ would be to send $\langle e_1 \rangle$ to 1, $\langle e_4 \rangle$ to 2, and so on.

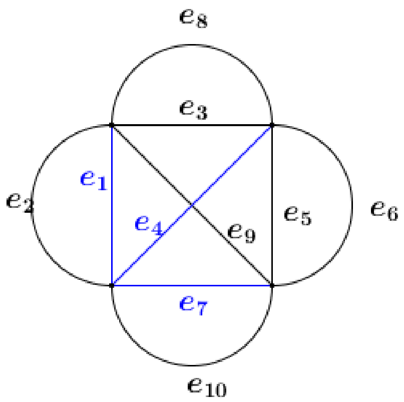


Figure 9: The matroid M.

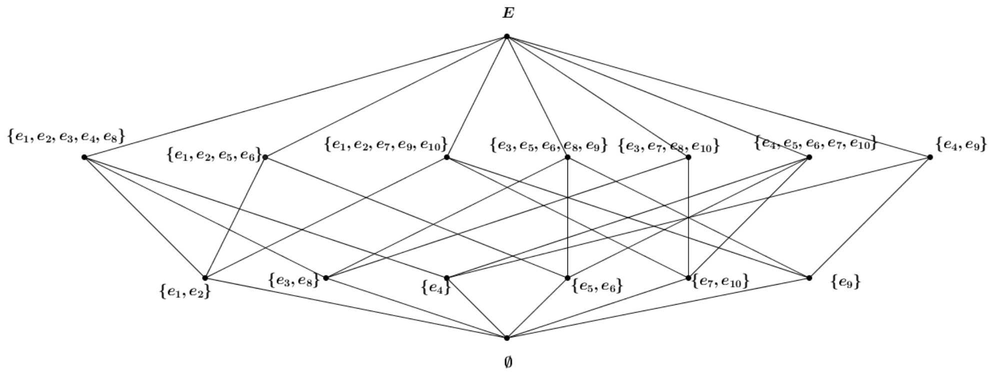


Figure 10: The lattice $\mathcal{L}(M)$ corresponding to M in Figure 9.

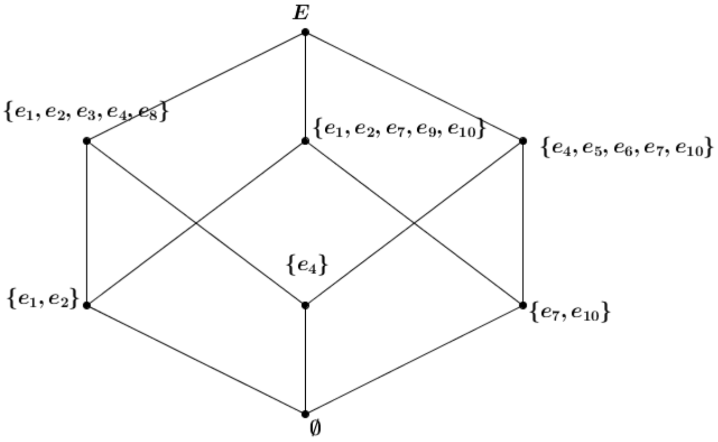


Figure 11: Restriction of $\mathcal{L}(M)$ to P_B.

Now suppose that we have a triangulation T on a triangle that is matroid-labeled by a matroid M (Figure 12). We are now going to show a different way to construct the induced Sperner labeling from B. First, take the path in P_B whose vertices are the flats for the induced Sperner labeling and label each of them by $1, 2$ and 3 in order of appearance from bottom to top(Figure 13). For any other nonempty element X in P_B, we label it so that it has the same label of the first element on the path that contains it. If we label each $x \in M$ with the label of the first set that contains x we will get an Sperner labeling on T.

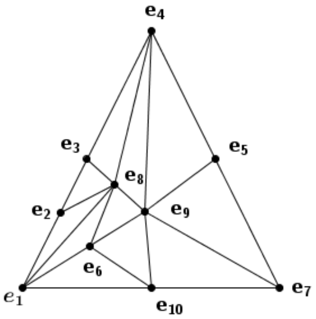


Figure 12: A triangle triangulated and matroid-labeled by M.

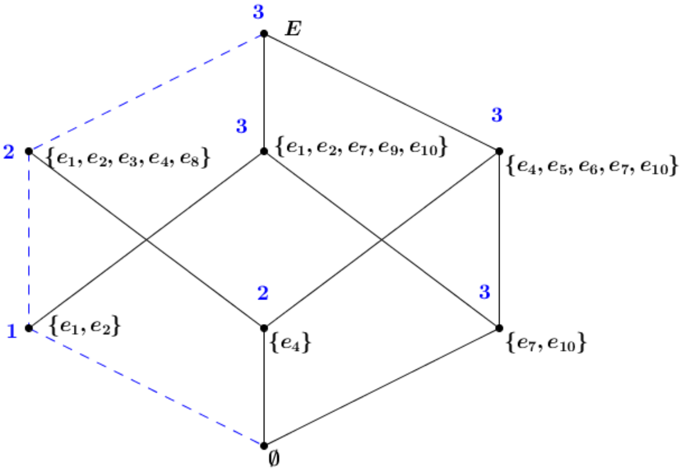


Figure 13: Sperner labeling in terms of P_B.

A natural question arises; for any order of the basis, can its respective induced Sperner labeling be constructed like this? It is indeed the case and we will give an example in Figure 14 that shows the induced Sperner labeling for the basis $B' = \{e_4, e_7, e_1\}$. This can be viewed as the basis B permuted by (132).

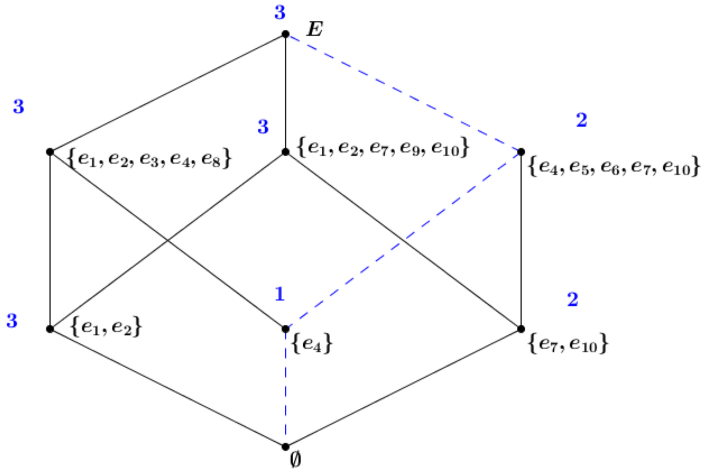
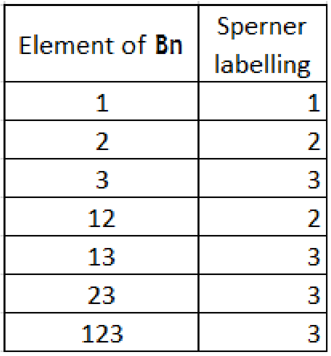


Figure 14: Sperner labeling for basis B'.

Notice that the given Sperner labeling induced by B can be assigned to the boolean algebra $\mathfrak{B}_3$ through the isomorphism ϕ. Now, if we let $\mathcal{R}$ be the set of pairs (A, b) where $A \neq \emptyset \in \mathfrak{B}_3$ and b is the respective Sperner labeling of A, we can extend the action of $\mathcal{S}_3$ on $\mathfrak{B}_3$ to $\mathcal{R}$ by letting $\sigma(A, b) = (\sigma(A), c)$ where $\sigma \in \mathcal{S}_3$ and c is the respective Sperner labeling of $\sigma(A)$. Notice that this is well defined because every element of $\mathfrak{B}_3$ is paired uniquely to one of the elements of $\{1, 2, 3\}$. In Figure 15, we show all the different pairs that live in $\mathcal{R}$.

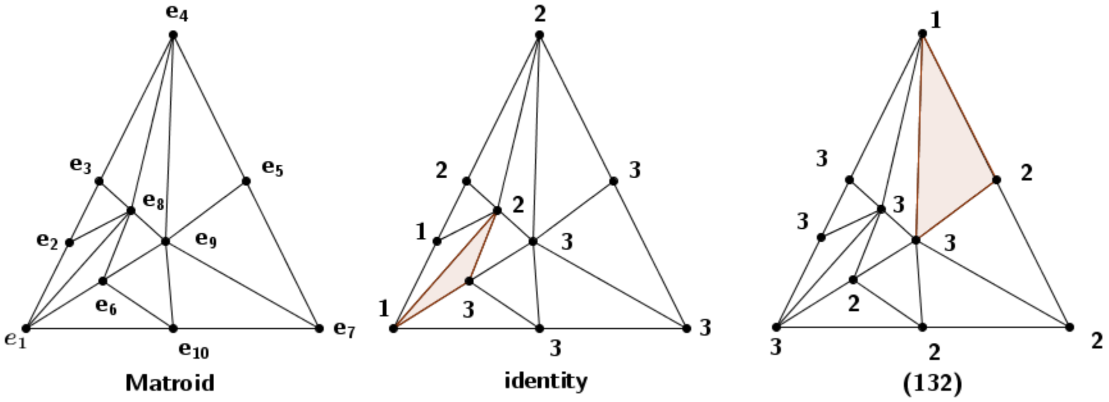


Element of **Bn**	Sperner labelling
1	1
2	2
3	3
12	2
13	3
23	3
123	3

Figure 15: $\mathcal{R}$.

So far the action acts in some auxiliary set and it is not so clear how permuting objects in the basis will relate to those tuples in $\mathcal{R}$. Since P_B is isomorphic to $\mathfrak{B}_3$ we can think of each flat of P_B as an element of $\mathfrak{B}_3$, and since every element in the matroid appears in some element of the lattice P_B we can "label" each element of M with the tuple in $\mathcal{R}$ corresponding to the first "flat" in $\mathfrak{B}_3$ that contains it.

As an example consider $e_3 \in M$, its label would be $(12, 2)$ since the first flat that contains e_3 is $\{e_1, e_2, e_3, e_4, e_8\}$ which corresponds to 12 in the isomorphism. And, if we were going to change the order of the basis by permuting it with (132), the new labeling of e_3 would be $(132)(12,2) = (13,3)$ since $(123)(\{1,2\}) = \{1,3\}$ and 13 is paired with 3. Although the action is defined on the labelings of the elements of M it is really an action on the elements of the basis B, since the elements of B are those who are labeled by tuples of the form (i,i) with $i \in \{1,2,3\}$. To see that the action indeed corresponds to changing the order of the basis and then finding the induced Sperner labeling the reader should note that the action is permuting maximal paths on the boolean algebra. As an example of how the action would look in a triangulation we give Figure 16.



Matroid **identity** **(132)**

Figure 16: Example of a triangulation labeled by M, the induced Sperner labeling and by permuting the elements of the basis by (132).

4.2.2 $\mathcal{S}_n$ as a Group Action

We will now generalize this concept to any dimension. Let M be a matroid on the vertices of a $(d-1)$-simplex T such that M induces a matroid labeling. Let the ordered basis corresponding to the main vertices of T be $B = \{b_1, ..., b_d\}$. For $H_i = \langle b_i \rangle$ with $i \in [d]$, we denote P_B as the poset generated by the H_i under the $\vee$ and $\wedge$ operations. We call P_B the *boolean algebra* induced by B.

A map f such that $f(H_i) = \{i\}$ easily constructs a poset $\mathfrak{B}_d$. An isomorphism between the posets P_B and $\mathfrak{B}_d$ follows trivially.

The reader should note that inducing a Sperner labeling from B', an ordered basis using the elements of B, is equivalent to:

- Taking a "maximal" path[1], $w = w_1 ... w_{d+1}$, in P_B and —beginning at level 1 then moving up —labeling $f(w_i)$ and its unlabeled subsets in $\mathfrak{B}_d$ as i.
- Label each element x of M with the label of $f(A)$, where A is the smallest set in P_B that contains x.

---

[1] We mean maximal without the empty set.

Note that by the previous method each element of $\mathfrak{B}_d$ is uniquely represented by a Sperner label. Hence, every Sperner labeling induced by B' is uniquely determined by a path in P_B (we are taking the path so that we may label). Moreover, when we take $B' = B$, because our construction of the isomorphism depends on the order the basis, the path that corresponds to B is the leftmost path in $\mathfrak{B}_d$.

When fixing an ordered basis B, define $\mathcal{R}_B$ as the set of tuples $(A, l(A))$ where $A \in \mathfrak{B}_d$ and $l(A)$ is the unique Sperner label attached to A as described above. Given $\sigma \in \mathcal{S}_d$, the action on $\mathfrak{B}_d$ induces an action on $\mathcal{R}_B$ where $\sigma(A, l(A)) = (\sigma(A), l(\sigma(A))$. Here $l(\sigma(A))$ represents the Sperner label that accompanies $\sigma(A)$. Formally stated we have the following:

Theorem 4.4. *Suppose T is a triangulation of a $(d-1)$-simplex that is matroid-labeled by M and let P_B be the boolean algebra induced by B, the ordered basis labeling the main vertices of T. The action on $\mathcal{R}_B$ corresponds to an action on the basis B and, therefore, an action in the uniquely determined induced Sperner labelings.*

Proof. To show that the action described before is in fact an action on the induced Sperner labelings we need to show two things:

- That the action on $\mathcal{R}_B$ is in fact an action.
- The action permutes paths in P_B while permuting the elements on the basis.

First we will show that the action is indeed an action: Suppose that $A \in \mathfrak{B}_d$ and $l(A)$ is the Sperner label such that $(A, l(A)) \in \mathcal{R}_B$. It follows that for the identity, $e \in \mathcal{S}_d$, $e(A, l(A)) = (e(A), l(e(A))) = (A, l(A))$. Furthermore, if $\sigma, \tau \in \mathcal{S}_d$ we have that:

$$\sigma(\tau(A, l(A))) = \sigma(\tau(A), l(\tau(A))) = (\sigma\tau(A), l(\sigma\tau(A))).$$

Then it follows that $\mathcal{S}_d$ acts on $\mathcal{R}_B$.

Now we will show that the action permutes paths in P_B while permuting the order of the basis. First, label each element $x \in M$ by $(A, l(A)) \in \mathcal{R}_B$, where $A = f(X)$ and X is the smallest set in P_B that contains x. We want to show that for a given $\sigma \in \mathcal{S}_d$ the induced Sperner labeling by B', the basis B reordered with σ, will correspond to the one of $\sigma\mathcal{R}_B$.

Let $x \in M$ and X be the flat such that X is the minimum element in P_B that contains x. By construction, x is labeled by $(f(X), i) \in \mathcal{R}_b$ for some i so that $f(X) \subseteq \{1, ..., i\}$. We know by definition of the action that $\sigma(f(X), i) = (\sigma f(X), m)$ where $m = \max\limits_{i \in f(X)} \sigma(i)$. Moreover we know that:

$$\sigma f(X) \subset \{\sigma(1), \ldots, \sigma(i)\} \subset \{1, \ldots, m\}.$$

So, when we apply σ^{-1} to the sequence of sets we will have that

$$f(X) \subset \{1, \ldots, i\} \subset \{\sigma^{-1}(1), \ldots, \sigma^{-1}(m)\}$$

because σ is an automorphism of $\mathfrak{B}_d$. Moreover, by using the fact that f is an isomorphism we have that:

$$X \subset \langle b_1, \ldots, b_i \rangle \subset \langle b_{\sigma^{-1}(1)}, \ldots, b_{\sigma^{-1}(m)} \rangle.$$

Notice that $b_{\sigma^{-1}(i)}$ is the element of the basis B that is in the "ith" position in B'. Now by the discussion before the theorem, we know that inducing a Sperner labeling by B' depends uniquely on a path $w = w_1 \ldots w_{d+1}$ where $w_i = \langle b'_1, ..., b'_i \rangle$ and b'_j is the element of B' in the position j. Therefore by definition: $w_m = \langle b_{\sigma^{-1}(1)}, ..., b_{\sigma^{-1}(m)} \rangle$. By construction of $\langle b_{\sigma^{-1}(1)}, ..., b_{\sigma^{-1}(m)} \rangle$ we know that w_m is the smallest element of the path that contains x. Thus, x is labeled by m in the induced Sperner labeling. We conclude that both the induced Sperner labeling and the labeling from $\mathcal{R}_B$ coincide. Hence the action on $\mathcal{R}_B$ is an action on the induced Sperner labelings.

□

With this group action we can now revisit the concepts from Section 2 in a new light. Specifically we will revisit the lemma we used to prove Lovász's corollary.

Lemma 4.5. *Suppose $\sigma \in \mathcal{S}_d$. Let S be a $d-1$-simplex, T a triangulation on S, and $\mathcal{P}(T)$ the Sperner labeling induced by a matroid $M = (E, \mathcal{I})$ of rank d defined on the vertices of T as previously described. Let $\sigma(\mathcal{P}(T))$ be the Sperner labeling induced by applying σ to the poset of the ordered basis on the main vertices of T. If there is a fully labeled Sperner triangle in $\sigma(\mathcal{P}(T))$ then its vertices correspond to a basis in M.*

Proof. This follows by construction. Suppose the ordered basis that that induced $\mathcal{P}(T)$ is $B = \{b_1, ..., b_d\}$. It follows that $\sigma(\mathcal{P}(T))$ has the ordered basis $B = \{b_{\sigma^{-1}(1)}, \ldots, b_{\sigma^{-1}(d)}\}$ and that if a vertex, x, is labeled by i then $x \in F_{\sigma^{-1}(i)} - (F_{\sigma^{-1}(i-1)} \cup ... \cup F_{\sigma^{-1}(1)})$. As such x must be independent to any element in $F_{\sigma^{-1}(i-1)} \cup ... \cup F_{\sigma^{-1}(1)}$ and has a different Sperner label than any of them. Clearly any fully labeled triangle must correspond to d independent elements. Thus, by the equicardinality of basis, a fully labeled triangle must correspond to a basis in M. □

5 Conclusion

Finding a lower bound on the number of basis simplices is heavily reliant on the triangulation being labeled and on the matroid being used to label it. We are currently working to find the necessary conditions for smothering. We believe the following to be true.

Conjecture 5.1. *Let M be a matroid such that there exists a parallel class with $r(M) + 1$ elements. Also suppose S is an $(r(M) - 1)-$simplex. Then there exists a triangulation T of S and a matroid labeling of M on T such that there is only* 1 *basis simplex.*

If we manage to prove this, we would know that finding a lowerbound different from 1 is impossible for certain matroids. With the knowledge of what matroids will not cause smothering, we can proceed with our main concern of finding a sharpened lower bound. We have some conjectures as to how this can be accomplished by looking at the lattice of flats.

If an improved lower bound is found for general d-simplices, we believe this problem can be generalized further to a version analogous to Sperner's lemma for polytopes.

References

[1] L Lovász. Matroids and sperner's lemma. *European Journal of Combinatorics*, 1(1):65–66, 1980.

[2] James G Oxley. *Matroid theory*, volume 3. Oxford university press, 2006.

[3] Richard P Stanley. Topics in algebraic combinatorics. *Course notes for Mathematics*, 192, 2012.

Committee Selection With Approval Voting and Hypercubes

Caleb Bugg **Gabriel Elvin**
Morehouse College UCLA

MSRI-UP, 7/24/15

Abstract

In this paper we will examine elections of the following form: a committee of size k is to be elected, with two candidates running for each position. Each voter submits a ballot with his or her ideal committee, which generates their *approval set*. The approval sets of voters consist of committees that are "close" to their ideal preference. We define this notion of closeness with Hamming distance in a hypercube: the number of candidates by which a particular committee differs from a voter's ideal preference. We establish a tight lower bound for the popularity of the most approved committee, and consider restrictions on voter preferences that may increase that popularity. Our approach considers both the combinatorial and geometric aspects of these elections.

This work was conducted during the 2015 Mathematical Sciences Research Institute Undergraduate Program (MSRI-UP), supported by grants from the National Science Foundation (DMS 1156499) and the National Security Agency (H98230-15-1-0039). We also thank supervisors Duane Cooper, Francis Edward Su, and Pam Urresta.

Contents

1 Background

We find it both interesting and useful to study committee selection under approval voting, which allows voters to specify a set of outcomes that would satisfy them. This system contrasts with the classic "one person, one vote" method. Ratliff [4] describes what can happen in certain elections if approval voting is not used. At Wheaton College in Massachusetts, a faculty election was held to select a committee with three seats. It was narrowed down to a runoff: two candidates running for each slot. At this stage, voters submitted a ballot of their ideal committee, and majority voting was used to decide the winner of each spot on the committee individually. As described in the paper, Wheaton College has a history of diversity and especially gender equality, and most of the electorate voted for a diverse committee. Unfortunately, the majority voting system ignored people's committee preferences as a whole. This led to an elected committee of all white males, a group most people were unhappy with [4]. Outcomes such as this suggest that, at least in certain cases, approval voting can lead to a more accurate representation of what voters actually want.

In [3], authors consider elections with a pool of n candidates, who are each competing for any of k positions on a committee, where $k < n$. This is contrary to candidates running for a particular spot, as in the scenario at Wheaton College. The authors allow each voter to select j candidates that they would prefer to be elected, with $j \leq k$. Therefore, candidates are not competing for a particular seat. Furthermore, the authors are able to guarantee a certain fraction of the voters will be satisfied with the outcome under these conditions. Then they provide further characteristics of voter data that they believe will increase the fraction of satisfied voters, such as making the voter's preferences intersect. Both the methods and proof techniques utilized in this paper prove useful to our current research.

2 Introduction

Here we consider elections in which there are a finite number of voters, k spots on the committee, and two candidates running for each spot on the committee. With this format, we will represent a voter's preference for the first candidate with a 0, and his or her preference for the second candidate with a 1. Also, we will write the k-tuples as binary strings (i.e. for $(1, 0, 1)$ we will simply write 101.)

We define C_k as the set of all committees of size k, which is equivalent to the set of binary k-tuples (i.e. the set of all k-tuples whose entries are 0 or 1.) Note that this is a finite set with $|C_k| = 2^k$. Geometrically, this set is a k-dimensional cube where each element of C_k represents the coordinates of a vertex on the cube. This will be of great use to us throughout the paper.

Definition 2.1. Let V be a set of voters. Then we define a function

$$\sigma : V \to C_k,$$

where $\sigma(v)$ is the ideal committee for a voter $v \in V$.

Definition 2.2. The *distance* between any two committees c_1, c_2 is a nonnegative integer given by the number of candidates by which the committees differ, denoted $d(c_1, c_2)$. We also define the distance between any two voters to be the distance between those voters' ideal preferences.

When viewing committees as vertices on a cube, one can see distance as the length of the shortest path along edges between two vertices. Here we solidify these concepts with a few examples.

Example 1. The distances between committee 101 and all committees of size 3 are given in Table 1.

Table 1: Distances relative to 101.

d = 0	d = 1	d = 2	d = 3
101	001	011	010
	111	000	
	100	110	

Definition 2.3. Given a nonnegative integer r, the *Hamming neighborhood* of a binary k-tuple x is the set of all binary k-tuples that are a distance of at most r away from x, i.e.

$$H_r(x) = \{c \in C_k \mid d(c, x) \leq r\}.$$

We call r the *radius* of a given Hamming neighborhood H_r. We consider societies where a voter's approval set is $H_r(\sigma(v))$. We will denote approval sets as $A_r(v)$ and call r the *approval radius.*

Example 2. Consider a voter $v \in V$ such that $\sigma(v) = 101$. Then

$$A_1(v) = \{101, 001, 111, 100\},$$

i.e. the first two columns of the table in Example 1. See Figure 1 for a geometric interpretation.

In general, approval voting defines a set of outcomes for each voter that he or she would deem acceptable. For example, suppose candidates a through f are running in some election. A voter may feel: "I approve of candidates a, b, and c, but not d, e, or f." Then a, b, and c go on the voter's ballot, and the candidate with the most votes at the end wins. In most literature on the subject of approval voting, the term *society* refers to a set of voters and their approval sets. Here we have a slightly altered definition, as our approval sets are easily defined once we have the committee size (k), a voter's ideal preference ($\sigma(v)$), and a radius for the election (r).

Definition 2.4. Consider a *society* in which each voter's approval set is a Hamming neighborhood of a fixed radius r centered around their ideal preference. We denote such a society $S = S(k, V, r)$, where k is the size of the committee to be elected, V is a finite set of voters, and r is the radius used for the election.

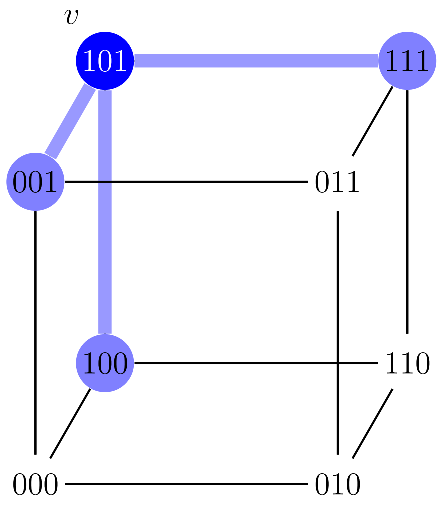


Figure 1: A geometric representation of $A_1(v)$.

Remark. When electing a committee of size k, realistically it only makes sense to consider elections with radii that are strictly between 0 and k. For if a society had radius 0, then we would be moving back to something that is not in the true spirit of approval voting, since every voter would only approve of their ideal committee. Also, if a society had radius k, then every voter would approve of every possible committee, and elections would not even matter. Therefore, our elections make practical sense only if $0 < r < k$.

We borrow the following definition from Berg et al. [1].

Definition 2.5. The *agreement number* of a committee c, denoted $a(c)$, is the number of voters who approve of c. The *agreement number* of a society S, denoted $a(S)$, is the maximum agreement number over all committees, i.e.

$$a(S) = \max_{c \in C_k} a(c).$$

Letting n denote the number of voters, we will also define *agreeement proportion*, both of a committee and a society, as $a(c)/n$ and $a(S)/n$, respectively. We denote the agreement proportion of a society S with $P(S)$. The agreement proportion of a committee can be thought of as the fraction of voters who approve of a particular committee, while the agreement proportion of a society is the largest fraction of people who can be made "happy" by electing most approved committee.

Example 3. Consider an election to be held in a society S with $k = 3$, $V = \{v_1, v_2, v_3, v_4, v_5\}$, and $r = 1$, and let $\sigma(v_1) = 000$, $\sigma(v_2) = 100$, $\sigma(v_3) = 101$, $\sigma(v_4) = 101$, and $\sigma(v_5) = 111$. Then the committee approved of by the most voters is 101 (actually a tie with 100), which is approved of by 4 voters, and hence the agreement proportion is 4/5. See Figure 2 for a geometric perspective.

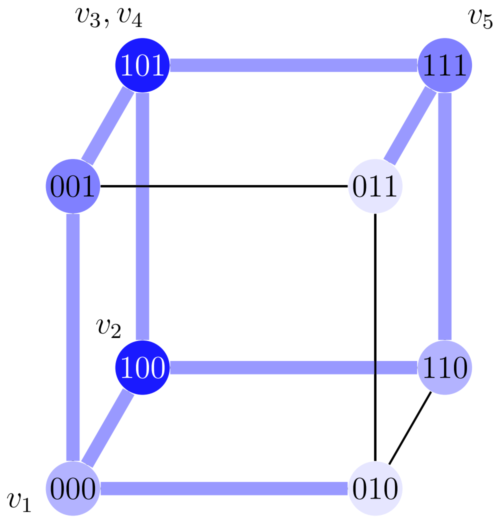


Figure 2: Each vertex represents a possible election outcome. For the election in Example 3, the darker the vertex, the more voters approve of that committee. Here 101 and 100 both have the highest agreement proportion.

Throughout the paper, we will refer to things combinatorially and geometrically, whichever works best for the given situation. For instance, we may refer to "a committee of size k," but we may also refer to the same concept as "a cube of dimension k."

Remark. For practical reasons, perhaps in a real election, if committees are tied for the highest agreement proportion we can use a fallback bargaining technique similar to that described by Brams et al. [2], which minimizes the maximum distance from voters to a committee. For instance, in our previous example, this technique would declare 101 the winning committee, since more voters share the ideal preference 101.

3 Balanced Societies

Now we will look at some results that indicate lower bounds on agreement proportions. First, we define a specific type of society that gives insights into agreement proportions of societies in general.

Definition 3.1. We define a *balanced society* $\tilde{S} = (k, V, r)$, where $|V| = 2^k$, by a bijective function

$$\beta : V_i \to \sigma(v_i) \qquad \text{for} \qquad i = 1, 2, \ldots, 2^k$$

that assigns each of the 2^k voters a unique ideal preference.

For example, a balanced society voting for a committee of size 3 would mean that there are 8 voters, where each voter represents one of the binary triples (see Figure 3.) This definition may be an unrealistic scenario in an actual election, but it is a useful case to consider, as we will see later.

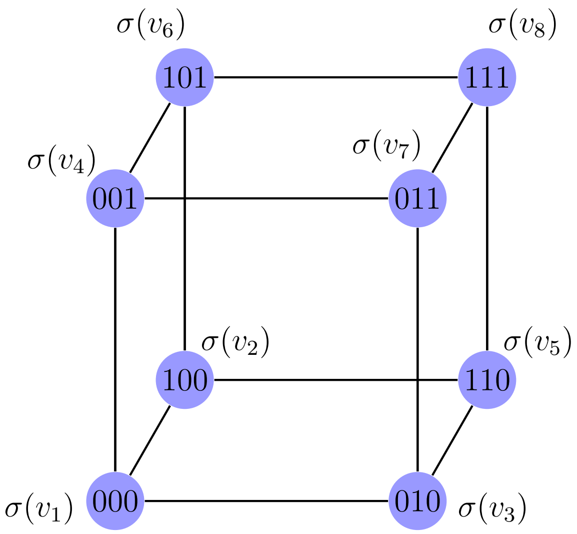


Figure 3: A balanced society.

Theorem 3.2. *If $\tilde{S}(k,V,r)$ is a balanced society, then the agreement proportion $P(\tilde{S})$ is given by*

$$P(\tilde{S}) = \frac{1}{2^k}\sum_{i=0}^{r}\binom{k}{i}. \tag{1}$$

Proof. First we must note that every possible committee will have the same agreement proportion by symmetry. Consider a cube of k dimensions, and pick any vertex c (which also represents a committee.) Choosing paths along edges is equivalent to changing coordinates. Hence, for each i, $0 \le i \le r$, there are $\binom{k}{i}$ possible vertices to arrive at by traveling along paths of length i. This means that for any committee, there are

$$\sum_{i=0}^{r}\binom{k}{i} \tag{2}$$

other committees within a Hamming neighborhood of radius r, and since there is exactly one voter for each committee in a balanced society, we can just divide (2) by the number of voters to obtain the desired result. □

Corollary 3.3. *For any voter v, $\left|A_r(v)\right| = \sum_{i=0}^{r}\binom{k}{i}$.*

Proof. For each $i = 0,\dots,r$, by definition of $A_r(v)$, for every committee $c \in A_r(v)$ we are choosing i coordinates to change out of k possible coordinates. Looking at all possible i gives us the desired result. From a geometric standpoint, we are basically traveling from $\sigma(v)$ along paths of length i arriving at other vertices, which are the committees in the approval set of v. □

Theorem 3.4. *If $P(\tilde{S})$ is as in (1), then for any society S, $P(S) \ge P(\tilde{S})$.*

Intuitively, this result appears reasonable: in a balanced society, no voters share an ideal preference and these preferences are spread evenly across C_k. To prove this theorem we use an averaging method similar to that of Davis et al. [3].

Proof. Recall that for a society with n voters $a(c)/n$ is the fraction of voters who approve of committee c. Define

$$D_v(c) = \begin{cases} 1 & \text{if } A_r(v) \ni c \\ 0 & \text{else} \end{cases},$$

i.e. D just says whether or not voter v approves of committee c. This gives us another way to define the agreement proportion of a committee:

$$\frac{a(c)}{n} = \sum_{v \in V} \frac{D_v(c)}{n}.$$

To calculate the average agreement proportion of a committee, first sum over all possible committees and then divide by the total number of committees. Here we first manipulate the sum and explain below:

$$\begin{aligned} \sum_{c \in C_k} \frac{a(c)}{n} &= \sum_{c \in C_k} \left(\sum_{v \in V} \frac{D_v(c)}{n} \right) \\ &= \sum_{v \in V} \left(\frac{1}{n} \sum_{i=0}^{r} \binom{k}{i} \right) \\ &= \sum_{i=0}^{r} \binom{k}{i}. \end{aligned}$$

The first equality follows by our definition. The second equality follows by summing over committees first. We can see this from Corollary 3.3 because for each voter $v \in V$, there will always be $\sum_{i=0}^{r} \binom{k}{i}$ committees v approves of. For the final equality, we simply note that n represents the number of voters, and we are summing over all n voters in V. Dividing by the total number of committees gives us the average: $\frac{1}{2^k} \sum_{i=0}^{r} \binom{k}{i}$. So, for any election there must be a committee whose agreement proportion is at least that of the average. If any committee had an agreement proportion that was less than the average, then there must exist a committee with a higher agreement proportion. Hence, for any society S,

$$P(S) \geq \frac{1}{2^k} \sum_{i=0}^{r} \binom{k}{i}. \tag{3}$$

□

Corollary 3.5. *In a society $S(k, V, r)$ with $k = 2r + 1$, $P(S) \geq 1/2$.*

Proof. Here we simply plug in $k = 2r + 1$ into (3), and with some algebraic manipulation and binomial coefficient identities we get:

$$\begin{aligned}
P(S) &\geq \frac{1}{2^{(2r+1)}} \sum_{i=0}^{r} \binom{(2r+1)}{i} \\
&= \frac{1}{2^{2r+2}} \left[\sum_{i=0}^{r} \binom{2r+1}{i} + \sum_{j=r+1}^{2r+1} \binom{2r+1}{j} \right] \\
&= \frac{1}{2^{2r+2}} \left(2^{2r+1}\right) \\
&= \frac{1}{2}.
\end{aligned}$$

□

From Theorem 3.2, we know that equality in (3) can always be achieved with a balanced society. This makes sense because in a balanced society, every committees' agreement proportion will be exactly the same, hence they will all be equal to the average.

4 Pairwise Intersecting Societies

Now let us restrict our voter preferences. After all, when considering real-world elections, there is rarely a uniform distribution, and many voters may have similar preferences. We consider a society in which any two voters will always share some committee in their approval sets, the same *super-agreeable* definition used by Berg et al [1]. Here we use our notation to define this type of society rigorously.

Definition 4.1. We call a society $S(k, V, r)$ a *pairwise intersecting society* if for every $v_i, v_j \in V$, $A_r(v_i) \cap A_r(v_j) \neq \emptyset$.

By definition, in pairwise intersecting societies, any two voters can differ on no more than $2r$ coordinates. Therefore, looking at societies $S(k, V, r)$ where $k \leq 2r$, all voters would pairwise intersect automatically and our previous results would hold. In this section we will focus on societies with $k = 2r + 1$, where the committee size is just large enough so that we can see some interesting results in pairwise intersecting societies. First we state and prove some general facts about such societies.

Proposition 4.2. *If $k < 2r + 1$, then $A_r(v_1) \cap A_r(v_2) \neq \emptyset$.*

Proof. Assume $k < 2r + 1$, then $d(v_1, v_2) \leq 2r - m$ where $m \geq 0$. Let $\sigma(v_1) = (b_1, b_2, \ldots, b_k)$ and $\sigma(v_2) = (c_1, c_2, \ldots, c_k)$ Then there exists a committee in C_k that is approved of by both voter v_1 and v_2. Paricularly,

$$C_k = \{b_1, b_2, \ldots, b_r, c_{r+1}, c_{r+2}, \ldots, c_k\}$$

is equivalent to changing the first r coordinates of $\sigma(v_2)$ to match the first r coordinates of $\sigma(v_1)$, then changing the last $r-m$ coordinates of $\sigma(v_1)$ to match the last $r-m$ coordinates of $\sigma(v_2)$, where $r-m \leq r$. Both voters have changed at most r coordinates of their ideal preference, so they both approve of this committee choice. Then, this committee lies in both $A_r(v_1)$ and $A_r(v_2)$, such that the intersection is not the empty set. Therefore, if $k < 2r+1$, then $A_r(v_1) \cap A_r(v_2) \neq \emptyset$. □

We will utilize the contrapositive of this statement, namely "If $A_r(v_1) \cap A_r(v_2) = \emptyset$, then $k \geq 2r+1$". If two voters do not agree on any committees, then k must be at least $2r+1$.

Here is an interesting question: does a society need to meet our definition of *balanced* in order for its agreement proportion to be as low as possible? We know we can reach equality in (3) with a balanced society. However, that is not the only way. Consider a society $S(3, \{v_1, v_2\}, 1)$ where $\sigma(v_1) = 000$ and $\sigma(v_2) = 111$. $P(S) = 1/2$ (equality in (3)), yet S does not meet the exact criteria of being *balanced.* Here we generalize what a society must look like in order to have the smallest agreement proportion possible. To help us, we will define the notion of an "anti-committee."

Definition 4.3. Let $c \in C_k$. The *anti-committee of* c, denoted $\bar{c}$, is the committee $\bar{c} \in C_k$ such that $d(c, \bar{c}) = \text{k}$. Hence, every member of a committee's anti-committee is different. We can similarly define voter and anti-voter, where these voters have the exact opposite ideal preference.

Here are some small propositions regarding committees and anti-committees.

Proposition 4.4. *If* $k = 2r+1$*, then for any* $c \in S(k, V, r)$*,* $\dfrac{a(c)}{n} + \dfrac{a(\bar{c})}{n} = 1$

Proof. Assume $k = 2r+1$ and $a(c) + a(\bar{c}) \neq 1$. Then there exists a voter, v, such that neither c nor $\bar{c}$ is in $A_r v$. This implies that $d(c, \sigma(v)) > r$. Then, v must have distance at least $r+1$ away from c. So, $A_r v$ contains all committees with distance at most r from $\sigma(v)$, which already has distance at least $r+1$ from c. Since $r + (r+1) = 2r+1$, then $A_r(v)$ contains a point with distance $2r+1$ from c. But, if $k = 2r+1$, then there is only one committee with distance $2r+1$ from c, namely $\bar{c}$. So $\bar{c} \in A_r(v)$. This is a contradiction. Therefore, if $k = 2r+1$, then for any $c \in S(k, V, r)$, $\frac{a(c)}{n} + \frac{a(\bar{c})}{n} = 1$. □

Proposition 4.5. *For any society* $S(k, V, r)$ *with* $k = 2r+1$*, for all* $v \in V$*,* $\sigma(v) \in H_r(c) \cup H_r(\bar{c})$*, and* $H_r(c) \cap H_r(\bar{c}) = \emptyset$*, i.e every voter* v *will approve of committee* c *or committee* $\bar{c}$*, but not both. Recall* H *is the Hamming neighborhood defined in 2.3.*

Proof. Assume there is a voter $v \in V$ such that $\sigma(v) \notin H_r(c) \cup H_r(\bar{c})$. Then $d(\sigma(v), c) \geq r+1$ and $d(\sigma(v), \bar{c}) \geq r+1$. This would mean $d(c, \bar{c}) = 2r+2$, which is a contradiction. Also, without loss of generality, suppose there is some $v \in V$ such that $\sigma(v) \in H_r(c)$. Then $d(\sigma(v), c) < r$, and hence, $d(\sigma(v), \bar{c}) > r$. Therefore, $H_r(c) \cap H_r(\bar{c}) = \emptyset$. □

Proposition 4.6. *Consider a society* $S(3, V, 1)$ $(k = 3, r = 1)$ *such that equality is reached in (3). Then for every* $c \in C_3$,

$$\big|\{v \in V \mid \sigma(v) = c\}\big| = \big|\{v \in V \mid \sigma(v) = \bar{c}\}\big|,$$

i.e. for any committee c*, the number of voters whose ideal preferences lie on* c *must equal the number of voters whose ideal preferences lie on* c*'s anti-committee.*

Proof. For this proof, when refer to a voter v "at" or "lying on" a committee c, we mean $\sigma(v) = c$. Throughout the proof we will refer to Figure 4, where $A, B, C, D, E, F, G, A + j$ represent the number of voters who lie at each vertex. Assume there exists a committee and anti-committee with a different number of voters at each of them. In Figure 4 this pair is 000 and 111 and without loss of generality we assume there are more voters at 000 than at 111. Now since $P(S)$ reaches equality, we know the agreement number of ever committee the same, which means that if we add up the agreement numbers of all committees exactly distance 1 away from 000, it should equal the sum of the agreement numbers of all committees exactly 1 away from 111 (since there are three committees in each of the mentioned groups.) We get:

$$a(001)+a(100)+a(010) = (C+B+G+A+j)+(D+B+F+A+j)+(E+F+G+A+j),$$

and

$$a(101) + a(011) + a(110) = (B + C + D + A) + (G + C + E + A) + (F + E + D + A).$$

If we set these equations equal to each other and cancel terms that occur on both sides of the equation, we get:

$$3j + B + G + F = C + D + E,$$

which means $C + D + E > B + G + F$. However, since $a(000) = a(111)$, $C + D + E < B + G + F$. This yields a contradiction. $\square$

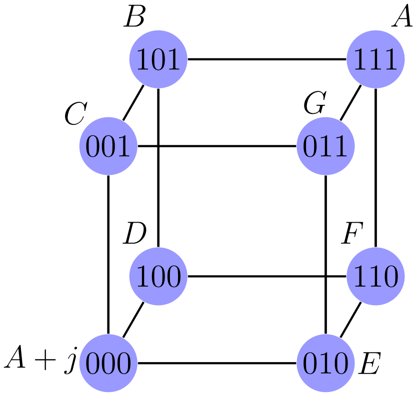


Figure 4: A society S corresponding to Proposition 4.6.

Theorem 4.7. *For any pairwise intersecting society $S(k, V, r)$ with $k = 3$ and $r = 1$, $P(S) > 1/2$.*

Proof. Corollary 3.5 immediately gives the non-strict inequality, and we know equality can never be reached because Proposition 4.6 states that there must always be voters with opposite preferences. If this were the case, there would exist voters $v, w \in V$ such that $d(v, w) = 3$, which would mean $A_1(v) \cap A_1(w) = \emptyset$. Then v, W would not pairwise intersect. Hence, we must have a strict inequality. $\square$

5 Discussion and Future Work

We have successfully shown that there exists a tight lower bound for the agreement proportion of any society under this form of committee selection. We know that this bound is tight because the exact proportion can always be achieved with a balanced society. We also looked at agreement proportions in societies where every pair of voters shared some common element in their approval sets. Specifically, we considered societies where the size of the committee k depended on the radius r for the election: $k = 2r + 1$. There were many interesting symmetric properties in this specific case. We also wanted to show more generally what a society had to look like in order for it to achieve the lower bound of its agreement proportion, as it would help us prove a bound for the agreement proportion of pairwise intersecting societies. We were able to prove claims in a specific case ($r = 1, k = 3$), and we noticed similar patterns as we moved to a higher dimension. As we move ahead in further research, we hope to generalize the proof strategy used for the specific case.

We also investigated what a pairwise intersecting society would look like geometrically if it were to have the smallest agreement proportion possible. We believe we found a construction for such societies. It involves "balancing" the voters, in a sense. We define voters such that each voters' ideal preference lies on a distinct vertex of a facet (i.e. a balanced society, but only on one facet), and then stretch half the voters out to another facet while maintaining a pairwise intersecting construction.

We created a computer program to generate specific societies, run elections, among other things. Here is a table of agreement proportions for what we believe are minimal for pairwise intersecting societies.

$S^*(k, V, r)$	n	$P(S^*)$
$(3, V_1, 1)$	4	$3/4 = 0.750$
$(5, V_2, 2)$	16	$9/16 = 0.625$
$(7, V_3, 3)$	64	$34/64 = 0.531\ldots$
$(9, V_4, 4)$	256	$133/256 = 0.519\ldots$
$(11, V_5, 5)$	1024	$526/1024 = 0.513\ldots$
$(13, V_6, 6)$	4096	$2090/4096 = 0.510\ldots$

It appears that the agreement proportions are decreasing towards 1/2, but never actually reaching it. In future work, we hope to prove this is actually a construction that minimizes the agreement proportion for pairwise intersecting societies.

Furthermore, we want to look at fractional pairwise agreement: a certain fraction of pairs of voters have overlapping approval sets. This could also be thought of probabilistically, i.e. looking at a pair of voters at random, there is a certain chance that their approval sets will intersect.

6 References

[1] Deborah E. Berg, Serguei Norine, Francis Edward Su, Robin Thomas, and Paul Wollan. Voting in agreeable societies. *Amer. Math. Monthly*, 117(1):27–39, 2010.

[2] Steven J. Brams, D. Marc Kilgour, and M. Remzi Sanver. A minimax procedure for electing committees. 2006.

[3] Matt Davis, Michael E. Orrison, and Francis Edward Su. Voting for committees in agreeable societies. In *The mathematics of decisions, elections, and games*, volume 624 of *Contemp. Math.*, pages 147–157. Amer. Math. Soc., Providence, RI, 2014.

[4] Thomas C. Ratliff. Selecting committees. *Public Choice*, 126:343–355, 2006.

7 Appendix

We created a computer program in C++ that does a variety of tasks, including simulating elections. We have many helper functions, but the central function of the program simulates an election with approval voting for a committee of size k, a radius r, and a number of voters n. We can also restrict voter preferences, e.g. make all voters pairwise intersecting. Here is the pseudocode for some of the functions, with the program code included for the first item. If there is any further interest in the programming code, please contact the researchers.

Algorithm 1 Hamming Distance

```
procedure HAMMING_DISTANCE(string $s1$, string $s2$)
    create integer $distance = 0$.
    for each index $i$ of $s1, s2$ (assume same length):
        if $s1$ at character $i$ does not equal $s2$ at character $i$ then:
            increment $distance$.
    return $distance$.
```

```
/**
 * calculates the hamming distance between two strings;
 * hamming distance is the number of places in the
 * strings at which they differ;
 * @param s1 the first string
 * @param s2 the second string
 * @return int the hamming distance between s1, s2
 * */
int hamming_distance(string s1, string s2) {
    int d = 0;
    for (int i = 0 ; i < s1.length() ; i++) {
        if (s1.at(i) != s2.at(i))
            d++;
    }
    return d;
}
```

Algorithm 2 Tally Votes

```
procedure TALLY_APPROVAL_VOTES(vector<string> $possibleComtyOutcomes$, vector<string> $voters$, int $r$)
    create a vector<int> $tally$ (to keep track of how many votes each committee gets).
    for each possible committee $i$:
        for each voter $j$:
            if hamming distance between $i$ and $j$ is less than $2r$ then
                increment $tally[i]$.
    return $tally$.
```

Algorithm 3 Election

```
procedure APPROVAL_ELECTION(int $r$, vector<string> $possibleCommittees$, vector<string> $voters$, bool $display$)
    get the $tally$ of all approval votes with
        tally_approval_votes(possibleCommittees, voters, r).
    find max of all tallies:
        create int $maxIndex = 0$.
        for each index $i$ in $tally$:
            if $tally[i] > tally[maxIndex]$ then
                set $maxIndex$ to $i$.
    if $display$ is true then
        display all election information (committees, votes for each committee, winner).
    return winner.
```

For the full code, contact Gabriel Elvin at gabrielelvin@gmail.com.

The Banquet Seating Problem

Michelle Rosado Pérez
University of Puerto Rico in Mayagüez

Ashley Scruse
Clark Atlanta University

A.J. Torre
University of Arizona

July 24, 2015

This work was conducted during the 2015 Mathematical Sciences Research Institute Undergraduate Program (MSRI-UP), supported by grants from the National Science Foundation (DMS 1156499) and the National Security Agency (H98230-15-1-0039). We also thank supervisors Duane Cooper, Francis Edward Su, and Lyda Urresta.

Abstract

Suppose you want to seat $n = mk$ people around k tables with m people at each table. Each person gives you a list of j people next to whom they would enjoy sitting. What is the smallest j for which you can always make a seating arrangement that would seat each person next to one of the people on their list? In this paper we show that j must be strictly more than half of n,the total number of people. Our key tool is a particular 'blue-green-red' lemma that helps us construct 'worst-case scenario' seating arrangements. We consider cases with two tables and more than two tables.

Contents

1 Introduction to the Problem

Our research focuses on a very concerning and pressing issue that affects people all across the world: unhappy seating arrangements at banquets and weddings. Many have fallen victim to bad seating arrangements, forced to sit between two people that they do not like. In order to prevent further unhappiness and awkwardness at banquets and weddings, our research team has set out to find the relationship between the number of people each guest would be happy sitting next to and happy seating arrangements.

Suppose you are planning a banquet, and have $n = mk$ people that you want to seat around k tables with m people at each table. What is the 'best' way to do this, and what information do you need from people about their preferences? Suppose each person gave you a list of j people that they would enjoy sitting next to. What is the smallest j for which you can always make a seating arrangement that would seat each person next to one of the people on their list? Joining the ranks of many fearless banquet and wedding planners, we have set out to find a seating arrangement that makes everyone happy. Our research looks at the relationship between the size of j and happiness.

1.1 Background research

In order to get an idea on how to approach the problem, we attempted to find previously published literature on the topic. However, in our search for papers, we were not able to find much work done on this question or similar questions. We read through articles on committee selecting as selecting a committee is similar to selecting a table as the goal to assemble a group of people of a certain size. We looked through different methods of committee selection, such as Klamler, Pferschy, and Ruzika's method of using weight constraints [3] and Bock, Day, and McMorris' using consensus rules to vote for a committee [2]. Ratliff's paper on selecting committees also gave an overview of different methods of committee selection, such as the drawbacks of looking at candidates as individuals rather than units [5]. While these papers helped us understand the difficulty in selecting a group of people as we have to view people as units rather than individuals, we found that the methods of voting for a committee were not analogous to looking at preference set size to assemble a happy table. One key difference is that outside voters choose committees while our research question called for invitees to list out their own preferences.

We also referenced similar questions to ours, such as how we can split a group of school girls into different subgroups to walk to school so that we have no two girls in the same group within a week's span [4]. While this overview of how mathematicians have approached similar problems in the past helped us understand the combinatorial aspect of this problem, we still felt that the idea of preference set size made this problem unique from past research.

Lastly, we looked for research papers that discussed happy seating arrangements. In Bellows and Peterson's paper on finding an optimal seating arrangement, they use their own wedding guest list to show how they can use ranks in matrices and computer programming to seat dates and family members at the same table [1]. Though this paper helped us understand one approach to getting preferences to sit together, we wanted to guests to sit directly next to each other rather than simply at the same table. We also wanted to approach this problem in a way more intuitive to us than computer programming. Overall, our literature review helped us understand the complexity of what could be a seemingly simple problem.

1.2 Definitions

Definition 1. For the purpose of this paper, we will define preference sets as a list of j people that each guest would enjoy sitting next to. Preference set size is j.

Definition 2. For the purpose of this paper, we will use the following ways to describe people, tables, and seating arrangements:

- **Happy person**: sits next to at least one person from their preference set
- **Unhappy person**: sits next to no one from their preference set
- **Happy table**: consists of all happy people
- **Unhappy table**: contains at least one unhappy person
- **Happy seating arrangement**: consists of only happy tables
- **Unhappy seating arrangement**: contains at least one unhappy table

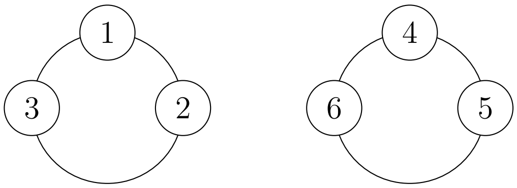
2 Two Tables ($k = 2$)

We will start with a simple case: $n = 6$ (total people), $k = 2$ (number of tables), $m = 3$ (people per table), and $j = 2$ (preference set size) to gain intuition for the problem. The following example shows a case where if the preference set size j is less than the number of people per table m, we would not find a happy seating arrangement.

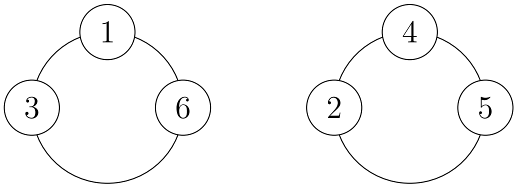
Consider this list of preference sets for people 1 to 6:

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

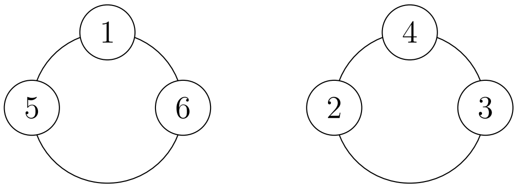
Note that people 1,2,3 all like each other but 4,5,6 only like 2 and 3. We first illustrate a few different seating arrangements for these preferences.



Here, persons 4,5, and 6 are unhappy. Let us try some switches to see if there is a seating arrangement that will satisfy everyone. First, let us try switching person 2 and person 6.



Now, everyone is satisfied but person 2, who has been isolated from the only two people he/she likes, namely 1 and 3. Let us try one more switch by switching person 3 and person 5.



Now, person 2 is happy, but we have made persons 1, 5, and 6 unhappy. Person 1 is now isolated from persons 2 and 3, who are the only two people that person 1 likes. These examples highlight the fact that the only way to make persons 1,2,3 happy is to keep them together as any other arrangement isolates one of them. But, then, none of persons 4, 5, and 6 can be happy.

This example has given us a sense of how to find preference sets of size $j < m$ such that every seating arrangement is unhappy. Our next theorem and lemma will formalize this notion.

Theorem 1. *Suppose there are* $n = 2m$ *people sitting at* $k = 2$ *tables with* $m = \frac{n}{2}$ *at each table and preference set of size* j*. If* $j < m$*, meaning the size of the preference set is less than the number of people at each table, then we can always find preference sets for which every seating arrangement is unhappy.*

Note that our theorem states no conditions on the preference set other than size. This means that some people may appear multiple times in preference sets while others may never appear. We will prove this theorem by first proving a lemma about how to construct preference sets of size $j = m - 1$ so that no happy seating arrangement is possible. To prove our next lemma and our further ideas, we have defined a specific case of preference sets below.

2.1 Blue Red Green (BRG) Preference Sets

Definition 3. Suppose we have n people, and there are blue, red, and green people amongst them who are defined by their preference sets as follows:

- Blue people only want to sit with green people.
- Green people only want to sit with each other and blue.
- Red people only want to sit with green people.

We will denote these types of preference sets as BRG preferences.

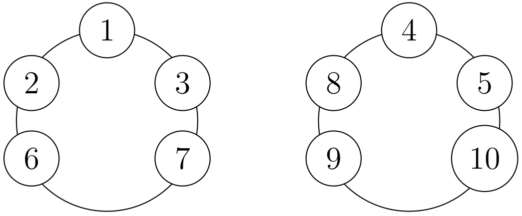
Our claim is that we can use BRG preference sets to find preference sets of size $j < m$ where no happy seating arrangement is possible. Consider the following example:

Example 1. Suppose $n = 10$ (people), $m = 5$ (per table), $k = 2$ (tables), and $j = 4$ (list size).

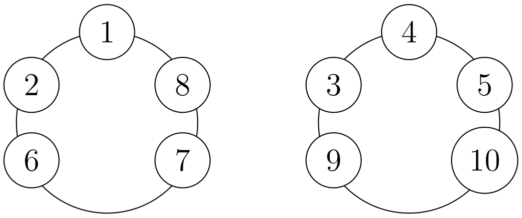
- 1: 2,3,4,5
- 2: 1,3,4,5
- 3: 1,2,4,5
- 4: 1,2,3,5
- 5: 1,2,3,4
- 6: 2,3,4,5
- 7: 2,3,4,5
- 8: 2,3,4,5
- 9: 2,3,4,5
- 10: 2,3,4,5

Note that these are BRG preferences: person 1 is blue, persons 2-5 are green, and persons 6-10 are red.

A possible arrangement:



In the above arrangement, everyone is happy except person 9, because he/ she is sitting next to only red people.



In this arrangement, we switched person 8 and 3 to make person 9 happy, but now person 7 is unhappy. We will formalize why no switch is able to make a happy seating arrangement below.

Lemma 1 (Blue Red Green Lemma)**.** *Suppose we have n people with BRG preference sets. If we have more red people than green people (at any table or overall), then every seating arrangement is unhappy.*

Proof. Suppose there are n people total seated, and among these n people are blue, red, and green people. Recall the BRG Preference Sets:

- Blue people only want to sit with green people.
- Green people only want to sit with each other and blue.
- Red people only want to sit with green people.

Suppose that we have created a happy seating arrangement with these preferences. In a happy seating arrangement, each red would need to be next to at least one green as red people only prefer green. Our two cases describe when a red is next to one green and when a red is next to two greens.

Case 1: A red person is seated next only one green person. Thus, this red person will be assigned to this green person.

Case 2: A red person is seated between two green people. Then, the red person should be assigned to the green person that is in the clockwise direction from them. We know that this assignment is unique because we are assuming that we have a happy seating arrangement, so in order for greens to be happy, they must be next to at least one green or the blue person. Thus, each green is next to at most one red.

These two cases show that a happy seating arrangement would have each red person assigned to a unique green person, which means that there would need to be at least as many green people as red at any table or overall. So, if we have less greens than reds, then every seating arrangement will be unhappy

□

2.2 Constructing Preference Sets to Prove Theorem 1

Proof. To finish the proof of our Theorem 1, let us look at $n = 2m$ people with m people at each table, and preference set size of $j = m - 1$. Following from our Blue Red Green Lemma, below is one way to construct preference sets so that there is no seating arrangement that will satisfy everyone:

- Person 1 will be considered a blue person. Their preference set consists of only the green people, $\{2, 3, ..., m\}$.
- Persons $2, 3, ..., m$ are called the green people. They only want to sit with blue and green people. For example, Person 2's preference set is $\{1, 3, ..., m\}$.
- Persons $m + 1, m + 2, ..., n$ are considered red people. They only want green people. So, all the red people have the same preference set: $\{2, 3, ...m\}$.

So, we have $m-1$ greens and m reds. From our BRG Lemma, we know that if we have more reds than greens, we are unable to find a happy seating arrangement.

Therefore, we cannot find a happy seating arrangement when we have less greens than reds. □

Note that we can again use a similar argument to handle $j < m-1$ using the BRG preference sets for those preference set sizes. The smaller that j is, the more people that can be left out of preference sets, which leaves the possibility of more red people. Thus, when $j = m-1$ is the most interesting case to examine when we have $j < m$.

Using BRG preference sets, we have proved Theorem 1. This means that we must have $j \geq m$, where m is the number of people at one table, in order to guarantee a happy seating arrangement when we have $k = 2$ tables. We will next examine the case where $j = m$ to sharpen the bound on the preference set size j that always guarantees a happy seating arrangement.

2.3 Extension of Blue Red Green to $j = m$

We showed that for $j < m$ we could find BRG preference sets where there were more red than green people, which we proved could not lead to a happy seating arrangement. For $j = m$, we find that we are always guaranteed more greens than reds, as explained by the proof below.

Lemma 2. *Suppose we have $n = 2m$ and we have preference set size of $j = m$. If people have BRG preference sets, then we have strictly more green people than red people.*

Note that this lemma states that $j = m$ will always give us only the **possibility** of happy seating arrangements, not a guarantee, as we have only proven the converse statement true.

Example 2. Suppose there are $n = 8$ people to be placed at $k = 2$ tables with $m = 4$ seats at each table and $j = m$. Define the BRG preferences as:

- 1: 2,3,4,5 (blue)
- 2: 1,3,4,5 (green)
- 3: 1,2,4,5 (green)
- 4: 1,2,3,5 (green)
- 5: 1,2,3,4 (green)
- 6: 2,3,4,5 (red)
- 7: 2,3,4,5 (red)
- 8: 2,3,4,5 (red)

Looking at these preference sets, we can see that if $j = 4$, then we are guaranteed four greens as the greens all have to like each other. We are also guaranteed one blue person as the greens must fill their fourth slot with someone who is not green. Then, we only have three people left to be red.

The idea of this lemma is that extending our preference set size to m automatically bounds our reds to be less than red. Because red people are invitees who are never mentioned in anyone's preference sets, extending the preference set limits them.

Proof. Suppose there are n total people seated at $k = 2$ tables with $m = \frac{n}{2}$ people at each table. Suppose $j = m$, and we have BRG preference sets.

Since preference sets are of size $j = m$, then there must be at least $m + 1$ blue and green people. But, then there are $m - 1$ reds. Each red person likes m greens. So, there must be at least m greens.

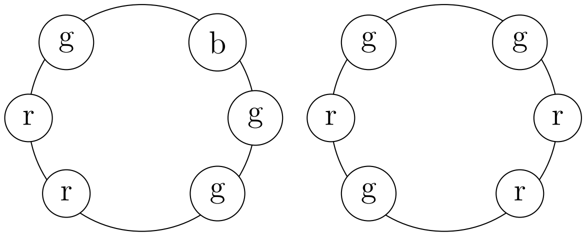
Thus, we will have strictly more greens than reds once we extend preference set size to $j = m$. □

Extending our preference set size to $j = m$ for BRG preferences always guarantees more greens than reds. However, continuing our ideas of blue, red, and green people, we found special cases where there were strictly more green than red people, but a happy seating arrangement was still not possible.

Example 3. Consider the following BRG preference sets and seating arrangement with $n = 12$, $k = 2$, $m = 6$, and $j = 6$:

- 1: 2,3,4,5,6,7
- 2: 1,3,4,5,6,7
- 3: 1,2,4,5,6,7
- 4: 1,2,3,5,6,7
- 5: 1,2,3,4,6,7
- 6: 1,2,3,4,5,7
- 7: 1,2,3,4,5,6
- 8: 1,2,3,4,5,6
- 9: 1,2,3,4,5,6
- 10: 1,2,3,4,5,6
- 11: 1,2,3,4,5,6
- 12: 1,2,3,4,5,6

Here, we have six greens: 1,2,3,4,5,6, one blue: 7, and five reds: 8,9,10,11, and 12. So, there are strictly more greens than reds. However, a happy seating arrangement is not possible. We have two possible cases for this seating arrangement: we will have three reds in a row, or we will have a green person between reds. These situations are unavoidable as we can see that we can try to make Table 2 happy by moving reds between tables, but then we will make Table 1 unhappy. There is no switch that will increase happiness.



For example, in this seating arrangement, the first table is happy. But, at the second table, we have one green unhappy as he/ she is sitting between two reds. We generalize this idea below.

Theorem 2. *Suppose we have $n = 2m$ people to sit at $k = 2$ tables with m people at each table and preference set size $j = m$. If m is even but not divisible by 4, then we can find preference sets where every seating arrangement is unhappy.*

Proof. Suppose we have $n = 2m$ people with BRG preferences to sit at $k = 2$ tables with m people at each table and preference set size $j = m$. Consider the case where m is even but not divisible by 4. We know that we must have strictly more green than red by Lemma 2. Consider the following seating arrangement:

$\frac{m}{2}$ greens	$\frac{m}{2}$ greens
1 blue	
$\frac{m}{2} - 1$ reds	$\frac{m}{2}$ reds

Note that we must set up the tables this way in order to avoid one table having strictly more reds than greens as the Blue Red Green Lemma guarantees that this seating arrangement will be unhappy. If we try to switch the blue person between the two tables, then we will simply have the same seating arrangement as we will have to add a red to Table 1 if we move the blue to Table 2.

Let us focus on Table 2 as Table 1 has strictly more greens than reds. Table 2, on the other hand, has $\frac{m}{2}$ green people and $\frac{m}{2}$ red people. We know we have an odd and equal amount of greens and reds at this table since m is even but not divisible by 4. We will think of a happy seating arrangement as strings of letters where R represents red and G represents green, where we know we must have each R next to at least one G but each G next to at most one R. Thus, since we have proven we need to assign each red to a unique green, we need to pair up our reds and greens so our string looks like:

RGRGRG...RG

Also, since $\frac{m}{2}$ is odd, we know we have an odd number of these RG pairs.

Since greens only like other greens, we will switch the order of every other RG as shown below:

$$\text{RGRG} \xrightarrow{\qquad\qquad\qquad} \text{RGGR}$$

We can now think of the strings as a unit of 4: RGGR. However, since we have an odd number of pairs of RGs, we will have one string of RG left over. This means that we will have one person unhappy as we will have a green not seated next to any greens. Thus, Table 2 is unhappy so our whole seating arrangement is unhappy.

Now, we have found that when m is even but not divisible by 4, we can find BRG preferences of set size $j = m$ such that every seating arrangement is unhappy.

□

Thus, we know that we must have preference set size $j > m$, or $j \geq m + 1$, when we have $k = 2$ tables. So, we must have preference sets that include over half the invitees in order to make people happy. Now, we will look at the cases when there are $k > 2$ tables.

3 More Than 2 Tables ($k > 2$)

We will use our ideas of BRG preference sets for more than two tables to show that we still need j to be strictly greater than half the people.

3.1 Applying BRG to $k > 2$

Theorem 3. *Suppose we have $n = mk$ people seated at k tables with m people at each table. If each person has preference set size $j \leq \frac{n}{2}$, then we can find preference sets for which every seating arrangement is unhappy.*

We will prove this theorem using two lemmas that extend our ideas from $k = 2$ tables.

Lemma 3. *Suppose there are $n = mk$ people sitting at k tables with $m = \frac{n}{k}$ at each table. Suppose everyone has BRG preference sets. If $j < \frac{n}{2}$, meaning the size of the preference set is less than half of the number of people, then we can always find preference sets for which every seating arrangement is unhappy.*

Proof. Suppose we have $n = mk$ people and we have preference set size $j < \frac{n}{2}$ with BRG preferences.

Recall that from our Blue Red Green Lemma we proved that we needed more green than red people to create a happy seating arrangement. Note that this idea is not dependent on the number of tables, meaning that we can extend this idea for a larger number of tables.

Thus, when $n = mk$ and $k > 2$, if $j < \frac{n}{2}$, we have the possibility of at least half of the people being red, i.e. left out of preference sets. Since we need to have at least one blue person, we will have less than half of the people green. Similar to our proof of Theorem 1, if $j < \frac{n}{2}$, then we can apply BRG to show that we can always find preference sets such that every seating arrangement is unhappy.

Therefore, we will need $j \geq \frac{n}{2}$ for $n = mk$ people. □

We can also extend our idea of m being even and not divisible by 4 with BRG preference sets to show that we can sharpen our previous bound to $j > \frac{n}{2}$.

Lemma 4. *Suppose we have $n = mk$ people seated at k tables with m people at each table. If m is even and not divisible by 4 and $j = \frac{n}{2}$, then we can find preference sets that always lead to unhappy seating arrangements.*

Proof. Suppose we have $n = mk$ people, where m is not divisible by 4. Suppose we have preference set size $j = \frac{n}{2}$ with BRG preferences.

Following our definitions of BRG preferences, we will have a blue person, $\frac{n}{2}$ green people, and $\frac{n}{2} - 1$ red people. Recall that the blue person likes only green people, green people only like other green people and the blue person, and reds only like greens.

We want to to distribute the greens as evenly as we can between the tables. Also, since we only have one more green than red, we will have only one table than can have strictly more greens than reds.

We will seat people in the following way:

- At Table 1, we will have the 1 blue person, $\frac{m}{2}$ green people, and $\frac{m}{2} - 1$ red people.
- At Table 2 to Table k, we will have $\frac{m}{2}$ green people and $\frac{m}{2}$ red people.

Note that any other seating arrangement would result in at least one table having strictly more reds than greens, which is automatically an unhappy seating arrangement by the Blue Red Green Lemma.

Since m is odd, we have that our numbers of greens and reds at Tables 2 through k, $\frac{m}{2}$, are odd. By Theorem 2, we know that we cannot find a happy seating arrangement when we have an odd but equal number of green and red people.

Therefore, Table 2 through Table k will have at least one person unhappy. So, we need $j > \frac{n}{2}$ to always guarantee a happy seating arrangement.

$\square$

4 Least Upper Bound on j

We have already shown that we must have $j > \frac{n}{2}$ for a happy seating arrangement. This finding suggests that j could be very large, possibly the size of the entire list of invitees, in order to satisfy everyone's preference set. By finding an upper-bound for j, we will find a range for j. First, we look at the case of $j = n-1$. This would mean that everyone picks everyone else but themselves. This is a trivial case as we could obviously find a happy seating arrangement if no one dislikes anyone.

Now, we move on to the case where $j = n - 2$. It is also impossible to have an unhappy seating arrangement with preference set size $j = n - 2$ as this would mean each person only has one person that they would not like to sit next to. Recall that being unhappy means that you are sitting next to two people that you do not like. But, if each person only dislikes one person, it is impossible for anyone to be sitting next to two people not in their preference sets. So, we know that we must have $j < n - 2$.

This strategy of finding an upper-bound gets more difficult when we hit $j = n-3$ as we have that $n \geq 6$, and we already have shown that if $n = 6$, then we must have $j \geq 3$.

Thus, we have our range of j: $\frac{n}{2} < j < n-2$. This range shows how large j needs to be in order to meet everyone's preferences. Also, since j is dependent on n, we know that the more people you invite, the more preferences you need.

5 Exploring Blue Red Green Happy Seating Arrangements

Now that we have our range for j, we will return to our ideas of blue, red, and green people to see what preference set size would create happy seating arrangements when people have BRG preference sets.

5.1 Creating Happy Tables

Lemma 5. *Suppose we have $n = mk$ people with BRG preferences sitting at k tables with m people per table. Then, the minimum number of green people per table to make a happy table, assuming there is one blue person, is $\frac{m-1}{2}$ if m is odd, and $\frac{m}{2}$ if m is even. We will call this number G_b.*

Proof. To prove it we are going to divide it in cases: when m is odd and when m is even.

Case 1: If m is an odd number, we need to prove that $G_b \geq \frac{m-1}{2}$. To prove by contradiction, suppose $G_b = \frac{m-3}{2}$. Because we also have one blue person, then $R = \frac{m+1}{2}$, where R is the number of red people per table. This mean that $R > G_b$, so we would not have enough green people to make every red person happy. Thus, $G_b \geq \frac{m-1}{2}$.

Case 2: If m is an even number, we need to prove that $G_b \geq \frac{m}{2}$. To prove by contradiction, suppose $G_b = \frac{m}{2} - 1$. Because we also have one blue person, then $R = \frac{m}{2}$. This mean that $R > G_b$, so we would not have enough green people to make every red person happy. Thus, $G_b \geq \frac{m}{2}$.

□

Lemma 6. *Suppose we have $n = mk$ people with BRG preferences sitting at k tables with m people per table.The minimum number of green people per table required to make a happy table, assuming there is not a blue person, is $\frac{m+1}{2}$ if m is odd, $\frac{m}{2}$ if m is even and divisible by 4, and $\frac{m}{2} + 1$ if m is even but not divisible by 4. We will call this number G.*

Proof. To prove it we are going to divide it in cases again: when m is odd and when m is even.

Case 1: If m is an odd number, we need to prove that $G \geq \frac{m+1}{2}$, where G is the number of green people per table required to make people happy. To prove by contradiction, suppose $G = \frac{m-1}{2}$. Then, $R = \frac{m+1}{2}$, where R is the number of red people per table. This mean that $R > G$, so we would not have enough green people to make every red person happy. Thus, $G \geq \frac{m+1}{2}$.

Case 2: m is an even number

Subcase 1: If $4 \mid m$, we need to prove that $G \geq \frac{m}{2}$. Suppose it is not; say $G = \frac{m}{2} - 1$. Then, $R = \frac{m}{2} + 1$. Again, $R > G$, which means that we would not have enough green people to make every red person happy. Thus, $G \geq \frac{m}{2}$.

Subcase 2: If $4 \nmid m$, we need to prove that $G \geq \frac{m}{2} + 1$. Suppose it is less than $\frac{m}{2} + 1$, namely suppose $G = \frac{m}{2}$. Then, $R = \frac{m}{2}$. Because $4 \nmid m$, m has the form $4t + 2$, where $t \in \mathbb{N}$, which means that we have an odd number of green and red people. By Theorem 2, we know we are unable to have a happy table when we have an odd but equal number of red and green people. Thus, we must have $G \geq \frac{m}{2} + 1$.

$\square$

5.2 More than 2 tables ($k > 2$)

We can use our previous lemmas that show how to guarantee happy tables with blue, red, and green people to discuss cases with $k > 2$ tables.

Theorem 4. *Suppose we have n people with BRG preference sets sitting at k tables with m people per table and preference set size j. The minimum preference set size j to make everyone happy is equal to $G_b + (k-1)G$, where G_b is the number of green people to make a happy table with one blue person at the table, G is the number of green people required to make a happy table with no blue person at the table, and k is the number of tables.*

Example 4. Consider the following preference sets and seating arrangement:

Suppose we have $n = 200$ people with BRG preferences seated at $k = 20$ tables with $m = 10$ people at each table. By our previous lemmas, we know that when $m = 10$ and there is one blue person, then we need five greens at the table to make a happy table since $m = 10$ is even. If $m = 10$ but we have no blue person, we would need six green people for a happy table as 10 is even but not divisible by 4. Following our definitions from the Blue Red Green Lemma, we have only one blue person. Thus, we would have 19 tables with six green people and one table with five green people, meaning we would need 119 green people total for a happy seating arrangement. Suppose that we have $j < 119$, namely $j = 118$.

Then, we would have 2 cases:

- **Case 1:** We would only be able to have 118 green people as each green person would use 117 slots to like all the other green people and one slot to like the blue person
- **Case 2:** We could try to have 119 greens, the number we need for a happy seating arrangement. But since preference sets are only of size 118, there would be at least one green that would not be liked by all of the reds, which would make them blue instead of green.

Thus, we would not be able to have as many greens as we need to satisfy everyone if $j < 119$. Now, suppose $j = 119$. Then, we would be able to have 119 greens as each green would use 118 slots to like every other green and one slot to like the blue person. The red people would also have enough slots to like every green. Therefore, we would be able to create a happy seating arrangement when j equals the minimum number of greens needed to make everyone happy.

Proof. Suppose we have $n = mk$ people with BRG preferences seated at k tables with m people per table. Then, we will have one table with the blue person as well as green and red people while the rest of the tables will only have green and red people. This means that the total number of green people at all the tables is $G_b + (k-1)G$. We need to prove that $j \geq G_b + (k-1)G$.

Suppose the contrary that $j < G_b+(k-1)G$, namely $j = G_b+(k-1)G-1$ (because if this preference set size does not work, it would not work if $j < G_b + (k-1)G - 1$). Observe that this will mean that we can only have $G_b + (k-1)G - 1$ green people, which leads us to 2 possibilities that will result in contradictions:

- Every green chooses every other green but does not have enough slots to like the blue person. Thus, this would make the blue person a red person instead, which contradicts that we should always have a blue person.
- Every green person chooses the blue person and $G_b + (k-1)G - 2$ green people, but every red person choose the same $G_b+(k-1)G-1$ green people, which would imply that we have another blue person instead of a green. This contradicts that we have $G_b + (k-1)G$ green people.

In either case, we will not have enough greens to satisfy everyone's seating preferences. Thus, we must have $j \geq G_b + (k-1)G$ in order to guarantee that we have enough greens to make each table happy. □

Thus, since we already know that we need at least as many greens as reds for a happy seating arrangement for k tables, we would need preference set size to be at least half of n, the total number of people. In many cases, we would need our preference set size strictly greater than $\frac{n}{2}$. Because the size of the preference sets must be relatively large to guarantee a happy seating arrangement, we now discuss loosening the conditions on what constitutes a happy seating arrangement in order to get a more feasible j.

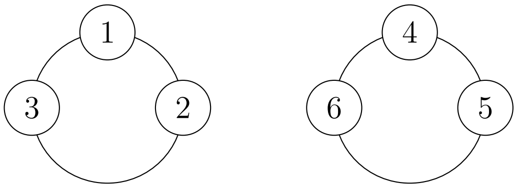
5.3 Loosening the Definitions of a Happy Seating Arrangement

From the above theorems, lemmas, and proofs, we can see the difficulty in finding a satisfactory seating arrangement when we define happiness as sitting next to at least one person in your preference set. Suppose that instead of defining a happy person as sitting directly next to the someone in their preference set, we define happiness as simply sitting at the same table with someone you like and unhappiness as not liking anyone at your table. How would this lower our j to something possibly more feasible when planning a banquet?

To get an idea for our theorem and proof below, consider what happens when you redefine happiness with $k = 2$ tables. What is the smallest j such that you can guarantee everyone has someone they like at their table?

Example 5. Consider what happens when $j = m - 1$. We can revisit one of our early examples when $n = 6$, $m = 3$, $k = 2$, and $j = 2$:

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



Here, we are unable to find a happy seating arrangement, even with our new definition of happiness. This example motivated us to continue our ideas of blue, red, and green people to find a bound on the preference set size that will always guarantee a happy seating arrangement using our new definition of happiness.

Theorem 5. *Suppose we have n total people we want to seat around k tables with m people at each table. If we redefine happiness to mean sitting at the same table as at least one person in your preference set, then we can use BRG preferences to show that we must have $j \geq 2k-1$.*

Proof. Suppose we have n people with BRG preferences that are seated at k tables with m people at each table. Recall that green people only like green and blue; blue likes green; and red only likes green.

Suppose that $j < 2k-1$, namely $j = 2k-2$. This means that we could have $2k-2$ greens, who all like one blue person in addition to liking the other greens. In order to satisfy the red people, we would want to sit a green at each table so we need k greens. But, green people only like green and blue people so we need the greens to come in pairs or with a blue person for a happy seating arrangement.

If we have $2k-2$ greens, then we can make $k-1$ pairs of greens, but then the blue person will not be able to sit next to a green person as the number of pairs of green people is less than the number of tables. Thus, the blue person will be unhappy, and we will have an unhappy seating arrangement.

If we make $k-2$ pairs of green people, then we will have two green people and one blue person to distribute amongst the $k=2$ tables left (the $k-2$ pairs of greens are able to make $k-2$ tables happy). Since red people only like green people, we would want to seat one green at each table. But, then, we only have one blue person to make one green person happy. Thus, one green person will be unhappy so we have an unhappy seating arrangement.

Therefore, we cannot have $j < 2k-1$ using the idea of BRG preferences. □

Claim: If we extend the preference set size to $j = 2k-1$, we are able to find a happy seating arrangement when our n people have BRG preferences.

Proof. Again, suppose we have n people with BRG preference sets sitting at k tables with m people at each table. Also, suppose that $j = 2k-1$.

Then, we will have $2k-1$ greens. Each green will use $2k-2$ slots to like every other green and one slot to like the blue person. So, we are able to have $k-1$ pairs of greens. Then, since we have $2k-1$ greens, we will have one green person left over; this green person can sit with the blue person. Thus, we will have $k-1$ pairs of greens, who will sit at $k-1$ tables, and the blue green pair will sit at the last table. So, if we have at least two greens or one green and one blue at a table, we can fill that table with any number of reds as the reds will be happy at the same table with one green.

Through using our ideas of blue, green, and red people, we are able to show that we need $j \geq 2k-1$ in order to create a happy seating arrangement when we redefine happiness. □

6 Conclusion and Discussion

Through constructing BRG preference sets, we have been able to find a lower bound on the preference set size j in order to make a happy seating arrangement. We have found that the range for the size of j is $(\frac{n}{2}, n-2)$.Though we have been able get both an upper and lower bound on j, we discuss our ideas and questions for future research below.

One question we have is how j could be more feasible or realistic if we put conditions on the preference set, such as limiting the number of people who could be left out or having everyone mentioned in at least one person's preference set. For further research, we could look at how the preference set size j changes if we add the constraint that every person must appear in at least one person's preference set as this would avoid the blue, red, and green scenario. Putting this condition on the preference set means that every person has at least one slot/ seat at a table.

Also, while our BRG preference set ideas have given us lower bounds for preference set size, we would like to construct an algorithm that could help us prove our conjectures of what j should be exactly equal to. We could use the types of moves that could be made in order to make an unhappy seating arrangement happier, such as switching people between tables, in order to create an algorithm to prove j.

Lastly, we would like to work more with the idea of redefining happiness. We have seen that j must be relatively large, strictly greater than half the number of people, in order to make sure everyone can sit next to at least one person in their preference set. But, if we redefine happiness to be simply sitting at the same table as someone in your preference set, we can possibly find a more feasible j in terms of real-life applications. Also, constructing an algorithm for getting people at the same table with at least one person in their preference set would be simpler than directly next to at least one person they like as there are fewer moves that need to be made in order to increase happiness.

Overall, our work on this problem has sparked more questions in our minds for future research. Because we have seen how large j needs to be to create happy seating arrangements, our ideas for future research revolve around redefining happiness or putting more conditions on the characteristics of the preference set.

7 Acknowledgements

We would like to acknowledge the following organization for their support: NSF, NSA, and MSRI. We would also like to acknowledge the following people for their tremendous support and encouragement: Dr. Cooper, Dr. Su, Dr. Tia Sondjaja, Lyda Urresta, and Dan Eckhardt!

References

[1] Megan L. Bellows and J.D. Luc Peterson, *Finding an optimal seating chart*, Annals of Improbable Research (2012), 1–7.

[2] Hans-Hermann Bock, William H. E. Day, and F. R. McMorris, *Consensus rules for committee elections*, Math. Social Sci. (1998), no. 3, 219–232.

[3] Christian Klamler, Ulrich Pferschy, and Stefan Ruzika, *Committee selection under weight constraints*, Math. Social Sci. **64** (2012), no. 1, 48–56.

[4] Erica Klarreich, *Answer to a 150-year-old math conundrum brings more mystery.*

[5] Thomas C. Ratliff, *Selecting committees*, Public Choice **126** (2006), 343–355.

A Volume Argument for Tucker's Lemma in 2-dimensions

Beauttie Kuture
Oscar Leong
Christopher Loa

July 25, 2015

Abstract

Sperner's lemma is a statement about labeled triangulations of a simplex. McLennan and Tourky (2007) provided a novel proof of Sperner's Lemma using a volume argument and a piecewise linear deformation of a triangulation. We adapt a similar argument to prove Tucker's Lemma on a triangulated cross-polytope P in the 2-dimensional case where vertices of P have different labels. The McLennan-Tourky technique would not directly apply because the natural deformation distorts the volume of P; we remedy this by inscribing P in its dual polytope, triangulating it, and considering how the volumes of deformed simplices behave. We then generalize the argument to apply to triangulated cross-polytopes whose vertices do not have different labels.

1 Introduction

1.1 Background

Sperner's Lemma is a combinatorial result that can be used to prove Brouwer's fixed point theorem and has many useful applications in economics. Sperner's Lemma states that given any triangulated n-dimensional simplex with a *Sperner labeling* [1], there exists an odd number of full simplices, or simplices whose vertices have distinct labels.

Recently, McLennan and Tourky provided a novel proof of Sperner's Lemma based on the following facts:

1. A triangulation remains a triangulation under a small perturbation.
2. The signed Lesbesgue measure of a parallelpiped is the determinant of the matrix whose columns correspond to its coordinates.

---

[1] A *Sperner labeling* associates one of the vertices on a minimal face to each vertex within the triangulation.

3. A polynomial function is constant if it is constant in a neighborhood of some point.

They constructed a continuous deformation as a function of a time parameter t that moves vertices of the triangulation linearly to the extreme points of the simplex with the same label. For small t, the deformed triangulation remains a triangulation by Fact (1). By Fact (2), the volume of each simplex can be written as a polynomial in t, and therefore the sum of the volumes is a polynomial in t. Since this sum remains constant in a neighborhood of $t = 0$, then by Fact (3), this sum remains constant for all t, and hence at $t = 1$. Note that this sum is the total volume of all deformed simplices at time 1. In order for sum to be non-zero, one of the terms must be non-zero. Therefore, some deformed simplex has non-zero volume, which can only happen if it was originally a full simplex. In fact, there must be one more full simplex of positive volume than of negative volume, which implies that the number of full simplices is odd.

Tucker's Lemma, on the other hand, is the combinatorial analogue of the Borsuk-Ulam Theorem and is stated below:

Theorem 1 (Tucker's Lemma (1946)). *Let S^n denote the n-sphere. Suppose that each vertex of a triangulated S^n is assigned a label from $\{\pm 1, \pm 2, ... \pm n\}$ in such a way that labels at antipodal vertices sum to zero. Then some pair of adjacent vertices of S^n have labels that sum to zero.*

It has been proven by several methods. For example, Freund and Todd (1981) provided a semi-constructive proof based on an algorithm of Reiser. Prescott and Su (2005) obtained another constructive proof by generalizing the hypotheses of Fan's Lemma. Baker (1970) used a combinatorial argument to prove Tucker's Lemma for the n-dimensional cube. In this paper, we develop an analogous argument to the one provided by McLennan and Tourky to prove Tucker's Lemma in 2-dimensions.

1.2 Terminology

We use the following definitions from McLennan-Tourky (2007):

An *affine combination* of points $p_1, p_2, \ldots, p_n$ in $\mathbb{R}^n$ is the sum $\sum_{i=1}^{n} \lambda_i p_i$ such that $\sum_{i=1}^{n} \lambda_i = 1$. We say that $p_1, p_2, \ldots, p_{n+1}$ are *affinely independent* if it is not possible to write one of them as an affine combination of the others. An *n-simplex* is the convex hull of affinely independent points $p_1, p_2, \ldots, p_{n+1}$.

A (finite) *simplicial complex* is a finite collection T of simplices in $\mathbb{R}^n$ such that any face (including $\emptyset$) of an element of T is an element of T and the intersection of any two elements of T is a (possibly empty) common face. The *underlying space* of T is the union of all simplices in T, i.e.

$$|T| = \bigcup_{\sigma \in T} \sigma.$$

We define the volume of each simplex $\sigma \in T$ to be

$$\mathrm{vol}(\sigma) = \frac{1}{2}\det[v_1 - v_0 \quad v_2 - v_0]$$

where $v_1 - v_0, v_2 - v_0$ is a positively oriented basis, i.e. the determinant is positive.

An *n-dimensional cross-polytope* is the convex hull of vertices whose coordinates are permutations of $(\pm 1, 0, 0, ..., 0)$. Fix a 2-dimensional cross-polytope P. Denote the extreme points of P by the set V_E which consists of the points $e_{+1} = (1, 0)$, $e_{-1} = (-1, 0)$, $e_{+2} = (0, 1)$ and $e_{-2} = (0, -1)$. We define a *facet* of P as a $(n-1)$-dimensional face of P. A *triangulation* of P is a simplicial complex T such that $|T| = P$. Fix such a T, and let V_T be the set of all vertices in T. We say that D with its set of extreme points $V_D = \{d_1, d_2, d_3, d_4\} \cup V_T$ is the *dual polytope* of P if each $e_i \in V_E$ corresponds to a face of D. In general, if P is a n-dimensional cross-polytope, then D is an n-dimensional cube.

We say that two vertices of P are *antipodal* if they are diametrically opposite to each other. We also say that an antipodally symmetric triangulation of P has a *Tucker labeling* if each vertex is assigned a label from $\{\pm 1, \pm 2\}$ and if labels at antipodal vertices sum to zero. Formally, a *Tucker labeling* is a surjective map $\ell : V_T \to \{\pm 1, \pm 2\}$ defined by $\ell(-v) = -\ell(v)$ for all $v \in \partial P$. A *complementary edge* is 1-simplex whose labels sum to zero.

In the following two sections, we prove the combinatorial version of Tucker's Lemma in 2-dimensions.

Theorem 2 (Combinatorial version of Tucker's Lemma)**.** *Let P be a 2-dimensional cross-polytope with a triangulation T. Suppose that T has a Tucker labeling. Then there exists a complementary edge in T.*

2 Main Result

Our goal is to adapt the volume argument of McLennan-Tourky to prove Tucker's Lemma in 2-dimensions. We will begin by proving Tucker's Lemma for a special case of a 2-cross polytope with distinctly labeled extreme points. In section 3, we will generalize the argument to polytopes in which the extreme points are not distinctly labeled. Here is the formal statement of our first proposition.

Proposition 1. *Let P be a 2-dimensional cross-polytope with a triangulation T. Suppose that T has a Tucker labeling and that the labels of the extreme points satisfy $\ell(e_i) = i$ for $i = \pm 1, \pm 2$. That is, the extreme points have four distinct labels. Then there exists a complementary edge in T.*

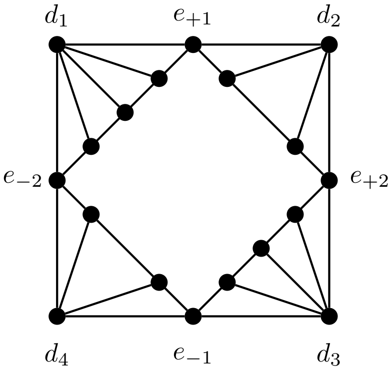
At first glance, one would deform the triangulation of the cross-polytope P using its labels, in the same way as McLennan-Tourky did with simplices. However, such a deformation, even for small time intervals, is not guaranteed to cover P, so the sum of the volumes of the simplices will not necessarily remain constant.

We can avoid this problem by inscribing P within its dual polytope D such that each $e_i \in V_E$ intersects an edge of D at its midpoint. Notice that D has side length of 2, so

$$\mathrm{vol}(P) + \mathrm{vol}(D \setminus \mathrm{int}(P)) = \mathrm{vol}(D) = 4. \tag{1}$$

We first need to triangulate $D \setminus \mathrm{int}(P)$. Note that this consists of 4 triangles glued at the extreme points of P. Each such triangle can be triangulated by "coning" the triangulation of $T \cap \partial P$ out to the unique d_i that lies in the triangle itself. Let $T \subset T_D$ where T_D is the new triangulation of D and P.

An example is given below:



Consider a perturbation function ϵ that moves vertices of this triangulation of D by a small amount δ. The lemma below, which is proved in [1], shows that if δ is small enough, and if the vertices on the boundary of D stay in their minimal face, then the perturbed triangulation remains a triangulation.

Lemma 1. *For $\epsilon : V_D \to D$ and $\tau = conv(\{v_1, \ldots, v_i\}) \in T_D$, let $\tau^\epsilon := conv(\{\epsilon(v_1), \ldots, \epsilon(v_i)\})$, and let ${T_D}^\epsilon := \{\tau^\epsilon : \tau \in T_D\}$. There exists a $\delta > 0$ such that ${T_D}^\epsilon$ is a triangulation of P whenever $\epsilon(v)$ is in the minimal face of v and $\|\epsilon(v) - v\| \leq \delta$ for all $v \in V_D$.*

We define a continuous deformation of D by first defining it on the vertices of the triangulation T_D. For $v \in V_T$ and $0 \leq t \leq 1$, let

$$\Delta_t(v) := (1-t)v + te_{\ell(v)}.$$

As a function of t, this moves every vertex $v \in V_T$ linearly to the extreme point of P with the same label. For $d_1, ..., d_4$, define $\Delta_t(d_i) = d_i$,so that these points do not move. Note that all other vertices on the boundary of D, namely e_i's also do not move because $\ell(e_i) = i$.

Now extend the definition of Δ_t linearly across the simplices of D. For any simplex σ spanned by vertices v_i, v_j, and v_k, notice $\sigma = \mathrm{conv}(\{\Delta_0(v_i), \Delta_0(v_j), \Delta_0(v_k)\})$. Let $\sigma' := \mathrm{conv}(\{\Delta_1(v_i), \Delta_1(v_j), \Delta_1(v_k)\})$ denote the simplex after the complete deformation at time $t = 1$. Suppose that we let

$$\mathcal{I} = \bigcup_{\sigma \in T} \sigma \quad \text{and} \quad \mathcal{K} = \bigcup_{\sigma \in T_D \setminus T} \sigma.$$

Then we can define

$$\mathcal{I}' = \bigcup_{\sigma' \in T} \sigma' \quad \text{and} \quad \mathcal{K}' = \bigcup_{\sigma' \in T_D \setminus T} \sigma'$$

to be the union of simplices after the complete deformation at time $t = 1$ for each triangulation T and $T_D \setminus T$, respectively.

The following lemma is self-evident.

Lemma 2. *For any triangulation T of D we have $vol(D) = \sum_{\sigma \in T} vol(\sigma)$.*

The volume of each simplex $\sigma \in T$ for any time $t \in [0, 1]$ is:

$$\mathrm{vol}(\Delta_t(\sigma)) = \frac{1}{2}\det[\Delta_t(v_2) - \Delta_t(v_0) \ \ \Delta_t(v_1) - \Delta_t(v_0)]$$

where $\Delta_0(v_2) - \Delta_0(v_0),\ \Delta_0(v_1) - \Delta_0(v_0)$ is a positively oriented basis. Then, define $p_D(t) := \sum_{\sigma \in T_D} \mathrm{vol}(\Delta_t(\sigma))$, which is a polynomial of time t by construction. For $\mathcal{I}$ and $\mathcal{K}$, let

$$\mathrm{vol}(\mathcal{I}) := \sum_{\sigma \in T} \mathrm{vol}(\Delta_0(\sigma)) \ \text{ and } \mathrm{vol}(\mathcal{K}) := \sum_{\sigma \in T_D \setminus T} \mathrm{vol}(\Delta_0(\sigma)).$$

Said more explicitly, $\mathrm{vol}(\mathcal{I})$ is the summation of the volumes of each simplex in the triangulation T of P while $\mathrm{vol}(\mathcal{K})$ is the summation of the volumes of each simplex in the triangulation T_D of D minus the simplices included in T prior to the deformation.

By Lemma 1, T_D will remain a triangulation in some δ-neighborhood of $t = 0$ because Δ_t is continuous. Therefore, Lemma 2 implies $p_D(t) = \mathrm{vol}(D)$ for the same δ-neighborhood of $t = 0$. Since p_D will remain constant around some small neighborhood around 0, we may conclude by Fact (3),

$$\sum_{\sigma \in T_D} \mathrm{vol}(\Delta_1(\sigma)) = \mathrm{vol}(D). \tag{2}$$

Notice that

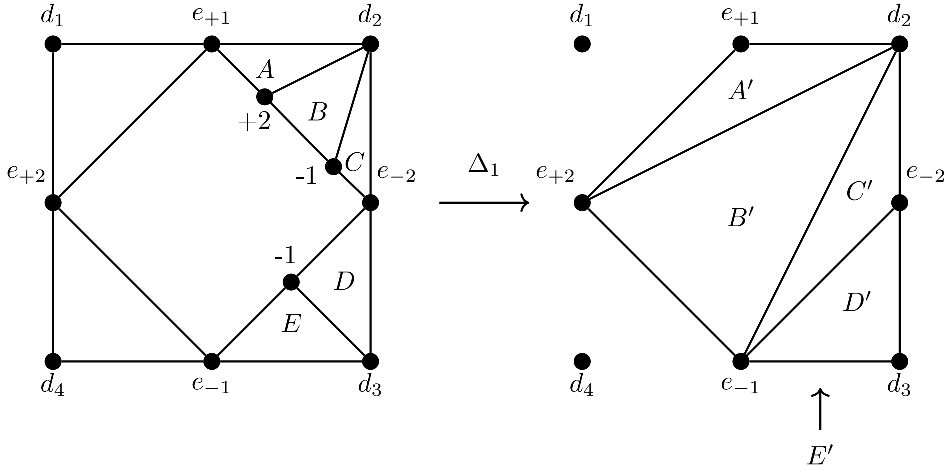
$$\mathrm{vol}(\mathcal{I}') := \sum_{\sigma \in T} \mathrm{vol}(\Delta_1(\sigma)) \ \text{ and } \mathrm{vol}(\mathcal{K}') := \sum_{\sigma \in T_D \setminus T} \mathrm{vol}(\Delta_1(\sigma)).$$

Therefore, by equations (1) and (2), we have

$$\mathrm{vol}(\mathcal{I}') + \mathrm{vol}(\mathcal{K}') = 4.$$

It will be useful to consider some examples of the image of $\mathcal{K}$ under the deformation Δ_1 and its volume.

Example 1. The first example denotes the simplices A, B, C, D and E on one half of $\mathcal{K}$, maps each $v \in V_T \setminus V_E$ to an extreme point of the same label and illustrates the resulting simplices under Δ_1. Note that since any Tucker labeling is antipodally symmetric, we may compute the volume of one half of the triangulation and the total volume will be twice that value. For the example below, we may compute the image of the volume under Δ_1 as $\mathrm{vol}(\mathcal{K}') = (\mathrm{vol}(\Delta_1(A)) + \mathrm{vol}(\Delta_1(B)) + \mathrm{vol}(\Delta_1(C)) + \mathrm{vol}(\Delta_1(D)) + \mathrm{vol}(\Delta_1(E))) = 2(\frac{1}{2} + \frac{3}{2} + \frac{1}{2} + \frac{1}{2} + 0) = 2(3) = 6$. Hence, by our formula, $\mathrm{vol}(\mathcal{I}') = -2$ since $\mathrm{vol}(D) = 4$. This concludes our first example.



In a number of cases, the orientation of simplices may change upon perturbing their vertices to the extreme points of P. If the orientation of a simplex changes, then we may consider its volume to be negative. The following lemma will be useful in our further discussion regarding oppositely oriented simplices.

Lemma 3. *Given two simplices σ_i and σ_j whose vertices have the same labels but in reverse order,their images under Δ_1 will be such that $vol(\sigma_i) = -vol(\sigma_j)$.*

Proof. Consider the simplices $\sigma_i = \text{conv}(v_1, v_2, v_3)$ and $\sigma_j = \text{conv}(v_4, v_5, v_6)$. The vertices of each simplex will have the same label so let $\ell(v_1) = \ell(v_4) = \ell(e_x)$, $\ell(v_2) = \ell(v_6) = \ell(e_y)$, and $\ell(v_3) = \ell(v_5) = \ell(e_z)$ where $x, y, z \in \{\pm 1, \pm 2\}$. Then the image of the volume of each simplex under Δ_1 is defined as

$$\text{vol}(\Delta_1(\sigma_i)) = \frac{1}{2}\text{det}[\Delta_1(v_3) - \Delta_1(v_1) \;\; \Delta_1(v_2) - \Delta_1(v_1)]$$

and

$$\text{vol}(\Delta_1(\sigma_j)) = \frac{1}{2}\text{det}[\Delta_1(v_6) - \Delta_1(v_4) \;\; \Delta_1(v_5) - \Delta_1(v_4)].$$

Notice that upon completion of the deformation, each vertex has moved to an extreme point of the same label so $\Delta_1(v_1) = \Delta_1(v_4) = e_x$, $\Delta_1(v_2) = \Delta_1(v_6) = e_y$, and $\Delta_1(v_3) = \Delta_1(v_5) = e_z$. Thus

$$\text{vol}(\Delta_1(\sigma_i)) = \frac{1}{2}\text{det}[e_z - e_x \;\; e_y - e_x]$$

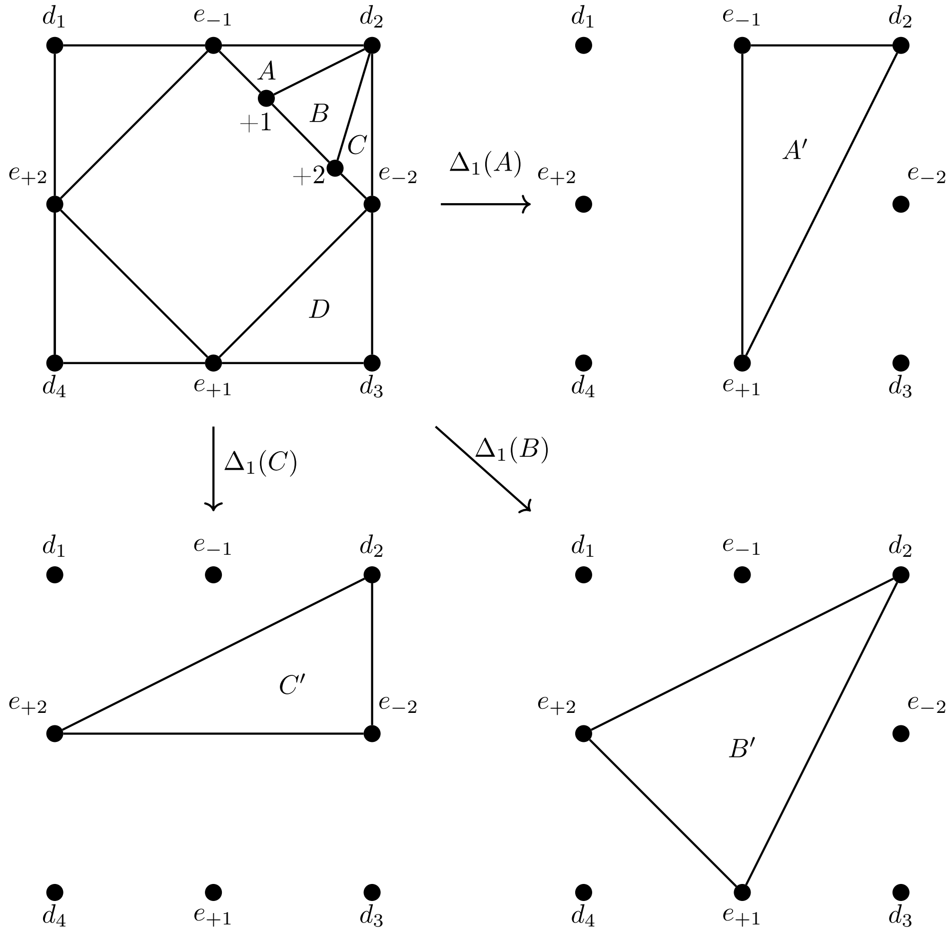
and

$$\text{vol}(\Delta_1(\sigma_j)) = \frac{1}{2}\text{det}[e_y - e_x \;\; e_z - e_x].$$

Since the columns of each matrix are switched, $\text{vol}(\Delta_1(\sigma_i)) = -\text{vol}(\Delta_1(\sigma_j))$ by nature of the determinant. ☐

The following example illustrates Lemma 3.

Example 2. In the next figure, note that the orientation of simplex B has changed so that $\text{vol}(B') = -\frac{3}{2}$. For the other simplices, $\text{vol}(A') = 1$, $\text{vol}(C') = 1$, and $\text{vol}(D) = \text{vol}(D') = \frac{1}{2}$ since its main vertices already constituted extreme points. Then $\text{vol}(\mathcal{K}') = 2(1 - \frac{3}{2} + 1 + \frac{1}{2}) = 2(1) = 2$.



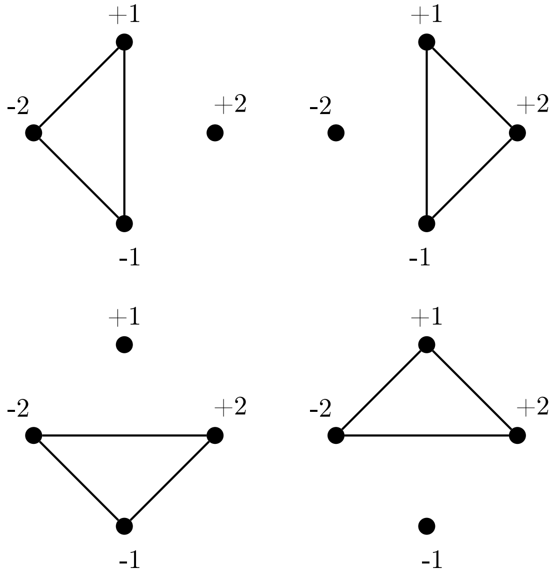
We return to the proof of Tucker's lemma. Recall the formula:

$$\text{vol}(\mathcal{I}') + \text{vol}(\mathcal{K}') = 4.$$

We will prove the existence of a complementary edge depending upon possible values of $\text{vol}(\mathcal{K}')$.

Proposition 2. *If $\text{vol}(\mathcal{K}') \neq 4$, then there exists a complementary edge in P.*

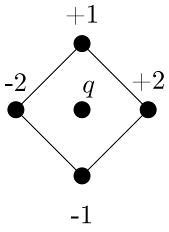
Proof. The assumption implies $\mathrm{vol}(\mathcal{I}') \neq 0$, meaning there must exist a simplex $\sigma \in T$ such that $\mathrm{vol}(\Delta_1(\sigma)) \neq 0$ since $\mathrm{vol}(\mathcal{I}') = \sum_{\sigma \in T} \mathrm{vol}(\Delta_1(\sigma))$. In order for $\sigma$to have nonzero volume, it must have three distinct labels. With only four possible labels, any combination of three distinct labels will result in σ having a complementary edge. The possible simplices containing nonzero volume are illustrated in the following figure.



□

Proposition 3. *If vol($\mathcal{K}'$) $= 4$, then there must exist a complementary edge on ∂P.*

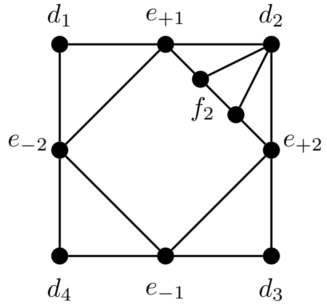
We will prove the contrapositive of Proposition 3. That is, if ∂P does not contain a complementary edge, then $\mathrm{vol}(\mathcal{K}') \neq 4$. To proceed with the proof, we will define the notion of a winding number for a sequence of labels using the diagram Q below.



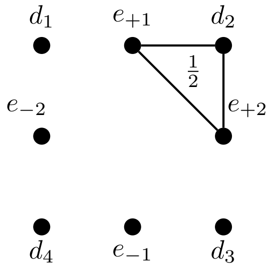
Given any sequence of labels $l_1, l_2, ..., l_n$ chosen from $\pm 1, \pm 2$ with no consecutive labels that sum to zero, there is a corresponding path along the edges of Q that starts at the vertex of Q labeled l_1 and tracing out the sequence $l_1, l_2, \ldots, l_n$. Note that if no consecutive labels sum to zero, then this path traces out a well-defined number of revolutions about its center q. Define the *winding number* $w(l_1, l_2, ..., l_n)$ to be the integer representing the net number of *full* counterclockwise revolutions of the path around q. Clockwise revolutions are represented by negative integers. For example, consider the sequence of labels $(+1, -2, -1, +2, +1, -2)$. Its winding number is equal to 1 because it completes one full counterclockwise revolution from the vertex labeled $+1$. As another example, consider $(+1, +2, -1, -2, +1, +2, -1, +2)$. Its winding number is equal to -1 because it completes one full clockwise revolution.

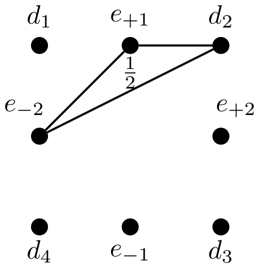
Proof of Proposition 2. Suppose there does not exist a complementary edge on any facet of P. Consider the sequence of labels on one of these facets. Because there are no consecutive labels summing to zero, then this sequence of labels corresponds to a path along Q with a well-defined winding number.

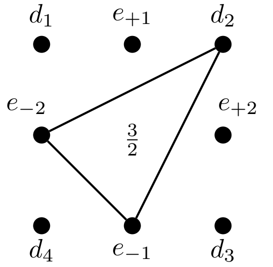
In particular, consider the collection of simplices containing d_2 within the triangulation of $\mathcal{K}$. Let f_2 denote the facet of P with extreme points labeled $+1$ and $+2$. We traverse along f_2 in a clockwise direction and obtain a sequence L of labels of vertices on f_2 that starts with $+1$ and ends with $+2$.

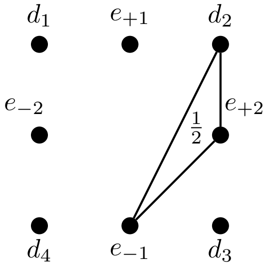


Any pair of adjacent vertices in L will be an edge of a simplex containing d_2 and will have a specific volume under Δ_1.Since we assumed that no complementary edges may exist, the following are four out of the eight possible ordered pairs of adjacent vertices in any simplex of f_2: $\{+1, +2\}, \{+1, -2\}, \{-2, -1\}$, and $\{-1, +2\}$. Recall that by Lemma 3, two simplices with the same labels but of opposite orientation will have volume of the same magnitude but of opposite sign. Hence, the volume of the pairs of adjacent vertices labeled $\{+2, +1\}$, $\{-2, +1\}$, $\{-1, -2\}$, and $\{+2, -1\}$ under Δ_1 will have the same magnitude as the simplices containing the aforementioned pairs but of opposite sign.The image of each simplex containing d_2 under Δ_1 is portrayed in the following figure along with their respective volumes in the center. Thus, the values of all possible simplices containing d_2 on f_2 are $\pm\frac{1}{2}$ and $\pm\frac{3}{2}$.









Then we can calculate the volume of the deformed simplices containing d_2 by inspecting the sequence of labels L. For example, if

$$L = \{+1, -2, -1, +2, +1, +2, -1, -2, +1, +2\},$$

then the volume of the simplices containing d_2 under Δ_1 is

$$\frac{1}{2} + \frac{3}{2} + \frac{1}{2} - \frac{1}{2} + \frac{1}{2} - \frac{1}{2} - \frac{3}{2} - \frac{1}{2} + \frac{1}{2} = \frac{1}{2}.$$

We claim that by computing $w(L)$ we can compute the associated volume of $\mathcal{K}'$. Notice that when $w(L) \neq 0$, $w(L)$ has a unique associated volume depending upon the ordering the sequence L. For $w(L) \in \mathbb{Z}$, if $w(L) \geq 1$, the ordering of L will be a counterclockwise path of the form

$$L = \{+1, -2, -1, +2, \ldots, +2\}$$

with $w(L)$ being the number of full counterclockwise revolutions. Then the associated volume will be $2w(L) + \frac{5}{2}$. Similarly, if $w(L) \leq -1$, the ordering of L will be a clockwise path of the form

$$L = \{+1, +2, -1, -2, \ldots, +2\}$$

where $|w(L)|$ is the number of full clockwise revolutions so the associated volume will be $2w(L) + \frac{1}{2}$. If $w(L) = 0$, note that the associated volume is either $\frac{1}{2}$ or $\frac{5}{2}$ depending if the order of the sequence L is clockwise or counterclockwise, respectively.

We may recreate the argument for the other three facets of P. Given the antipodal symmetry of the Tucker labeling of ∂P, each of the volume equations can appear on 0, 2, or 4 of the facets of P. Then we can create the following equation that describes all cases,

$$\mathrm{vol}(\mathcal{K}') = \gamma_1\left(2w_1+\frac{5}{2}\right)+\gamma_2\left(2w_2+\frac{5}{2}\right)+\gamma_3\left(2w_3+\frac{1}{2}\right)+\gamma_4\left(2w_4+\frac{1}{2}\right) \quad (3)$$

where $w_1, w_2, w_3, w_4 \in \mathbb{Z}$, $w_1, w_2 \geq 0$, $w_3, w_4 \leq 0$, and for $1 \leq i \leq 4$, $\gamma_i \in \{0, 2, 4\}$ and $\sum_{i=1}^{4}\gamma_i = 4$. Here γ_i represents the number of facets of P each volume equation holds for. For example, if $\gamma_1 = \gamma_3 = 2$, and $\gamma_2 = \gamma_4 = 0$ then

$$\begin{aligned}\mathrm{vol}(\mathcal{K}') &= 2\left(2w_1+\frac{5}{2}\right)+0\left(2w_2+\frac{5}{2}\right)+2\left(2w_3+\frac{1}{2}\right)+0\left(2w_4+\frac{1}{2}\right)\\ &= 4w_1+4w_3+6\end{aligned}$$

Notice that there does not exist $w_1 \geq 0$ and $w_3 \leq 0$ that can satisfy the equation $4w_1+4w_3+6 = 4$. It is easily checked that for all remaining cases, equation (3) cannot satisfy $\mathrm{vol}(\mathcal{K}') = 4$. Therefore, we have proven the contrapositive of Proposition 3.

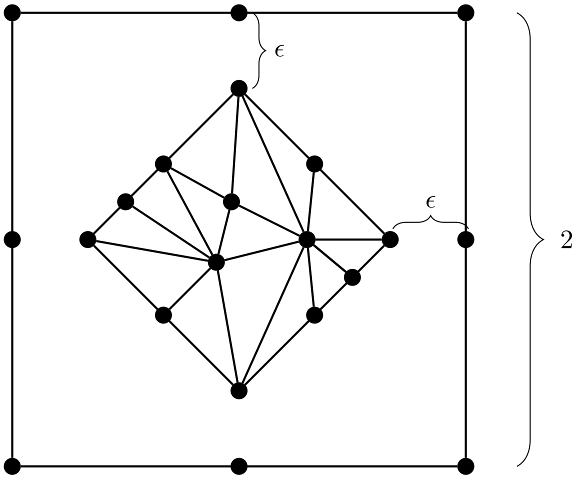
□

With the conclusion of Proposition 3, we have shown that in any there must exist a complementary edge in P. Thus Proposition 1 has been proven.

3 General Case in 2-Dimensions

We want to extend our argument for a triangulated 2-cross-polytope P with a Tucker labeling where the extreme points of P are not distinctly labeled. This will complete the proof of Theorem 2. We may assume that there exists at least one pair of ±1 and on pair of ±2 labels on ∂P. Otherwise, there must exist a complementary edge on ∂P since a ±1 or ±2 labeling of a 1-simplex is a Sperner labeling in 1 dimension, which is easily proven. So suppose the extreme points of P are not distinctly labeled but there exist at least one pair of ±1 and ±2 labels on ∂P.

Since the extreme points of P are not distinctly labeled, it is not immediately apparent as to where we want to map each vertex in the triangulation. We remedy this issue by placing P into the center of its dual polytope D whose side length is greater than the diagonal of P, as illustrated in the figure below. We can assume that D is a square with side length 2 and P is a 2-cross-polytope whose extreme points are at coordinates $(\pm(1-\epsilon), 0)$ and $(0, \pm(1-\epsilon))$ for some $0 < \epsilon < 1$.



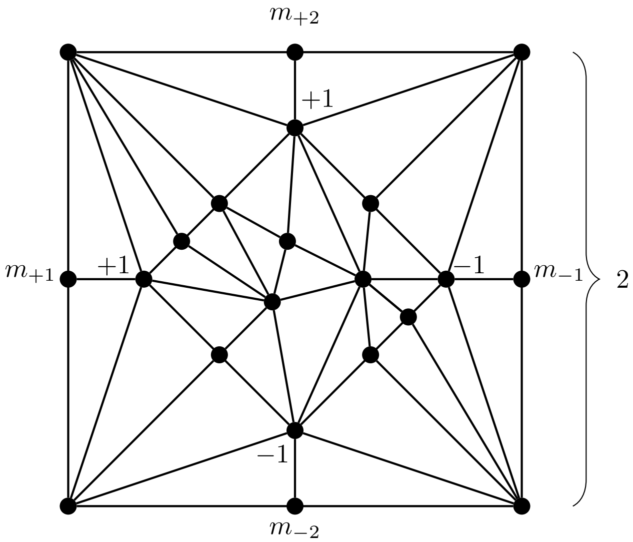
Let D_M denote the set of midpoints of the faces of D. As before, let T_D denote the triangulation of D, constructed as follows. Create an edge from each midpoint of D_M with the closest extreme point of P. Label each midpoint m_i according to these three conditions:

(i) The four midpoints are distinctly labeled.

(ii) The labels are antipodal.

(iii) There are no newly created complementary edges.

The first and second condition are clearly always possible. The third condition is also always possible, through the following procedure: without loss of generality suppose the extreme points of P have labels ± 1. Choose one antipodal pair of extreme points. Label the midpoint adjacent to the $+1$ vertex in this pair with $+1$ (let's denote this midpoint by m_{+1}). Similarly, label the midpoint adjacent to the -1 vertex with -1 (call this midpoint m_{-1}). Then, label the remaining two midpoints with $+2$ and -2 (call them m_{+2} and m_{-2}, respectively). The following figure illustrates the new triangulation along with possible labelings for each midpoint considering the given labels of the extreme points of P.



The logic of our following argument mirrors the procedure defined in the previous section. We may assert that our triangulation will remain a triangulation under small perturbations by Lemma 1. The continuous deformation, however, will change slightly. For each vertex in the triangulation T_D and $0 \leq t \leq 1$, let

$$\Delta_t(v) := (1-t)v + tm_{\ell(v)}.$$

Similarly to our previous Δ_t function, this deformation moves every vertex in the triangulation of P to a midpoint m_i of the same label. We may then define the volume of each simplex as a function of time t under our deformation Δ_t. Then $\mathrm{vol}(D)$ will be a polynomial in time t by our previous construction. Hence $\mathrm{vol}(D)$ will remain constant throughout our deformation since it is constant through a small δ-neighborhood and we reach the same equation

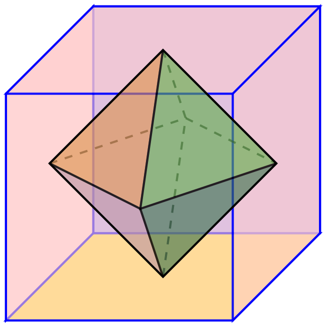
$$\mathrm{vol}(\mathcal{I}') + \mathrm{vol}(\mathcal{K}') = \mathrm{vol}(D) = 4$$

where $\mathcal{I}, \mathcal{I}', \mathcal{K}$, and $\mathcal{K}'$ are defined as they were previously.

Then we arrive to our previous cases. If $\mathrm{vol}(\mathcal{K}') \neq 4$ then $\mathrm{vol}(\mathcal{I}') \neq 0$ so there must exist a simplex with nonzero volume. Hence this simplex is distinctly labeled and there must exist a complementary edge. If $\mathrm{vol}(\mathcal{K}') = 0$ then we may still prove that there must exist a complementary edge on ∂P by proving the contrapositive of the statement. In this case, let $L = \{m_1, e_1, l_3, l_4, \ldots, l_{n-2}, e_2, m_2\}$ and Case 2 will hold. With the aid of Proposition 1 and its generalization to 2-cross polytopes without distinctly labeled extreme points, we have proven Theorem 2, Tucker's Lemma in 2 dimensions.

4 Discussion and Open Question

We would like to extend our argument to higher dimensions. For example, in the 3-dimensional case, we may consider an octahedron inscribed in a cube.



The difficulty in this approach lies in the fact that our elementary methods do not translate well in higher dimensions. The topological knowledge required for tackling this question in the 3-dimensional case would require more time than allotted during our research program.

References

[1] Andrew McLennan and Rabee Tourky. Using volume to prove sperner?s lemma. *Economic Theory*, 35(3):593–597, 2008.