Marks of a CSUF graduate from the College of Natural Sciences and Mathematics

GRADUATES FROM THE COLLEGE OF NATURAL SCIENCES AND MATHEMATICS:

- Understand the basic concepts and principles of science and mathematics.

- Are experienced in working collectively and collaborating to solve problems.

- Communicate both orally and in writing with clarity, precision, and confidence.

- Are adept at using computers to do word processing, prepare spreadsheets and graphs, and use presentation software.

- Possess skills in information retrieval using library resources and the internet.

- Have extensive laboratory, workshop, and field experience where they utilize the scientific method to ask questions, formulate hypotheses, design and conduct experiments, and analyze data.

- Appreciate diverse cultures as a result of working side by side with many people in collaborative efforts in the classroom, laboratory, and on research projects.

- Have had the opportunity to work individually with faculty members in conducting research and independent projects, often leading to the generation of original data and contributing to the research knowledge base.

- Are capable of working with modern equipment, instrumentation, and techniques.

DIMENSIONS

About the Cover

Mathematics and art have long been intertwined: The ancient Greeks believed that "perfect" mathematical proportions governed what was aesthetically pleasing, and this belief carried over to the Renaissance where artists used precise mathematical measurements to replicate the natural world in man-made creations. This was possible because math and nature are also inherently related: every complex structure and elaborate pattern in nature can usually be broken down into a series of equations and theorems. I sought to illustrate that triangular relation between math, art, and nature in this cover, by highlighting the persistence of the Golden Ratio/Golden Spiral, a mathematical understanding of proportions based on the Fibonacci sequence, in the natural geometry of objects studied in various fields. From the swirl of the celestial bodies in a galaxy to the spiral structures formed by complex molecular interactions, all the sciences have a beauty and art to them in isolation, but nothing truly exists in a vacuum, and the marvels of each field of study are further amplified when we celebrate the way all of STEM is simultaneously diverse and wonderfully interconnected.

The cover itself is designed to look like traditional 16th/17th century scientific illustration, which originally accompanied naturalist findings and were used to convey exciting and remarkable research, much like the manuscripts and abstracts that fill these pages.

Special Thanks

DIMENSIONS Editorial Staff

Editor-In-Chief
Sonali Vyas - Mathematics

Advisor
Tatiana Pedroza - Assistant Dean for Student Affairs

Layout Editor and Cover Designer
Sherif Ibraheem - Layout Editor
Shannon Chou - Cover Designer

Editors
Jessica Sherman - Biological Science
Mayra Silva - Biological Science
Melissa Fernandez - Biological Science
Krisha Marie Uy - Biological Science
Mykayla Miller- Chemistry and Biochemistry
Maria Valdovinos- Geological Sciences
Luis Gonzalez - Mathematics
Madeline Ceccia - Mathematics
Erick Engelby- Physics
Anny Antunovich - Physics

College of Natural Sciences & Mathematics

Dr. Marie Johnson - Dean
Dr. Sean E. Walker - Associate Dean
Dr. Merri Lynn Casem - Chair, Department of Biological Science
Dr. Peter de Lijser - Chair, Department of Chemistry and Biochemistry
Dr. Adam Woods - Chair, Department of Geological Sciences
Dr. Alfonso Agnew - Chair, Department of Mathematics
Dr. Ionel Tifrea - Chair, Department of Physics

Dear Titan NSMers,

Even in a pandemic, the college and our students still pursued research, designed experiments, collected data, wrote draft upon draft, and ultimately published their findings to share with other mathematicians and scientists. We could not be prouder of the effort that these pages represent, both for the students themselves and for the faculty mentors who stand behind these students and enable this work. On behalf of all of us here in NSM, please enjoy this issue of Dimensions as we celebrate the discovery of new knowledge and the original research of our students.

Marie Johnson
Dean, College of Natural Sciences and Mathematics

Table of Contents

Diet of Coyotes in Rural Agricultural Areas of California: A Comparative Analysis of Stable Isotope and Stomach Contents Analysis

Jesi Nonora
Advisor: Paul Stapp, Ph.D.
California State University, Fullerton
Department of Biological Science

ABSTRACT

Coyotes (Canis latrans) are opportunistic, generalist predators that, due to their flexible diet and tolerance of people, occupy a wide range of natural and human-modified ecosystems in California. As a result of their omnivorous habits, the trophic position of coyotes in food webs is highly variable depending on habitat type, the availability of wild vs. anthropogenic prey, and the degree of urbanization or agricultural development. Traditional methods of diet analysis such as stomach or scat content analysis (SCA) provide a snapshot of coyote diet over time but can provide precise information on the species and numbers of prey individuals eaten if there are undigested, identifiable remains present. Stable carbon and nitrogen isotope analysis (SIA) provides an integrated measure of assimilated foods over longer time periods, depending on the type of tissue analyzed, although it may be difficult to determine specific prey types or quantities consumed. We used both SCA and SIA to estimate diet of coyotes from rural Fresno County, California, the most productive agricultural county in the state, with nearly half its land area dedicated to farms and rangeland. Stomachs of 55
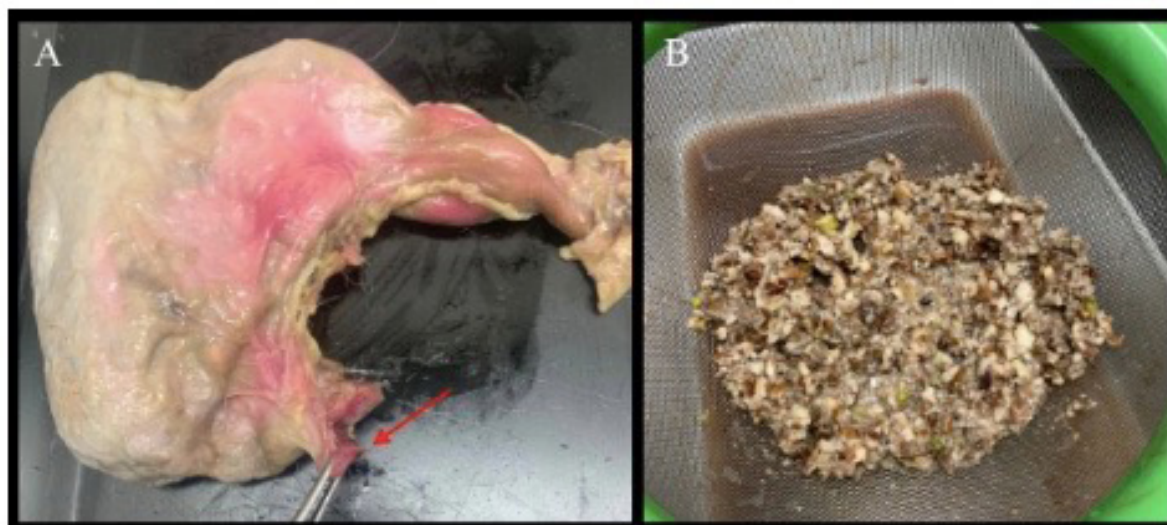
Figure 1. *Initial stages of stomach content analysis for a rural Fresno coyote, where a piece of stomach muscle tissue that connects to the small intestine is sampled for stable isotope analysis (A) before cutting and emptying contents of the stomach into a wire sieve (B). In this example, the coyote appears to have eaten many almonds.*

coyotes euthanized as nuisance animals were dissected to identify the types of foods eaten, based on bones, hair, scales, feathers, and plant parts. We also removed muscle tissue (stomach or jaw) from coyote carcasses and sent them to the UC Davis Stable Isotope Facility to determine their stable carbon and nitrogen isotope signatures. Preliminary analysis of stomach contents confirm the highly omnivorous diet of rural coyotes,

Figure 2. *Underbelly and dorsal skin and scales of a pacific gopher snake (Pituophis catenifer catenifer), identified through stomach content analysis of a rural Fresno coyote and viewed under a dissection microscope.*

which included small mammals, snakes, and birds, as well as large numbers of insects and agricultural crops, such as grapes and almonds. We will attempt to identify the species of mammals prey further by microscopic analysis of hair anatomy. When SIA results are received, we will compare them to diet estimates based on SCA and also to SCA and SIA-based estimates of food habits of urban coyotes from southern California.

Figure 3. *Dorsal view of the European earwig (Forficula auricularia), identified through stomach content analysis of a rural Fresno coyote and viewed under a dissection microscope.*
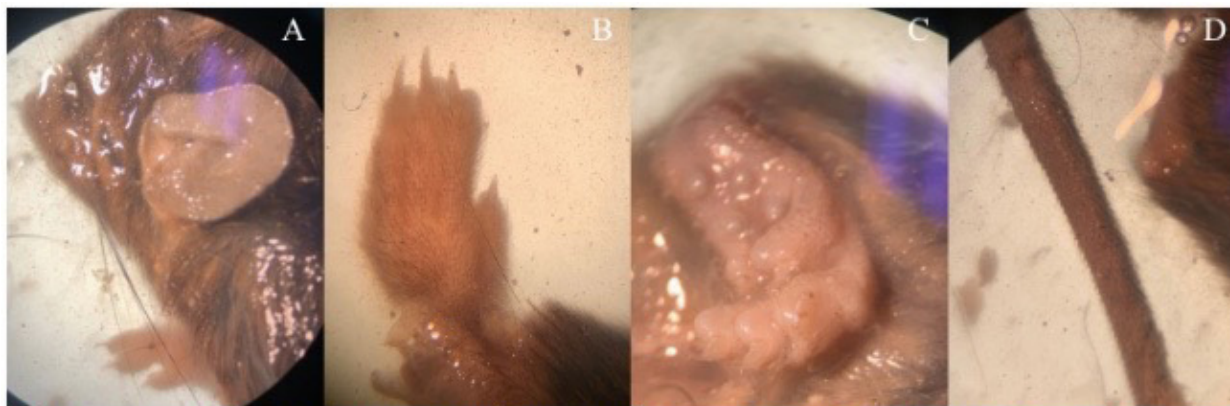
Figure 4. *Lateral view of the head (A), dorsal view of the front paw (B), ventral view of the hind paw (C), and tail (D) of a deer mouse of the genus Peromyscus, which was identified through stomach content analysis of a rural Fresno coyote and viewed under a dissection microscope.*

Green Waste Perceptions at California State University, Fullerton

Elijah Slaven
Advisor: Joel Abraham, PhD.
California State University, Fullerton
Department of Biological Science

Abstract

Food waste in landfills is a major contributor to climate change due to the release of methane into the atmosphere as organic materials decompose. Due to the large volume of food waste in landfills, highly concentrated pockets of methane build up can occur. One of the major contributors to green waste in landfills is higher education institutions due to the large volume of resources consumed on campus. A possible solution to reduce green waste in the landfills would be to implement composting bins around campus so that green waste could be diverted. A survey to determine student perceptions of green waste as well as perceived barriers and support for such a program was created and distributed to students at California State University, Fullerton. The survey results displayed a high positive attitude toward the environment was present among students as well as support for a composting program. However, gaps in green waste knowledge were also present with students having difficulty correctly identifying green waste form other types of waste. Overall, student support for green waste programs were present and the program has potential for being implanted across the campus.

Analyzing the Effect of Experimental Evolution on the Host Microbiome in Drosophila melanogaster

Author: Leah Spalding
Advisor: Dr. Parvin Shahrestani
California State University, Fullerton
Department of Biological Science

Abstract

The microbiome, or community of microbes, found within an organism affects its fitness, influencing traits such as development, immune defense, and lifespan. We studied how experimental evolution for longevity divergence in Drosophila melanogaster populations affects the associated microbiome. Moreover, we tested how flies from short- and long-lived populations are affected by manipulations to the microbiome. Quantitative analysis of bacterial abundance and composition was done by homogenizing and plating whole fly bodies. From this, we found that the associated microbiome of populations that have been evolved for prolonged lifespan largely consisted of bacteria from the phylum Proteobacteria. In contrast, the associated microbiome of populations evolved to be shorter-lived were dominated by bacteria from the phylum Firmicutes. To manipulate the microbiome of short- and long-lived flies, axenic flies were created by dechorionating D. melanogaster eggs. These eggs were subsequently inoculated with symbiotic bacterial species and abundance was surveyed. Preliminary results show that long-lived flies were more extensively colonized when inoculated with A. pomorum and A. tropicalis. Short-lived populations had overall higher microbial abundance, particularly when inoculated with L. brevis and L. plantarum. Continued experimentation is necessary to further elucidate the interactions between the Drosophila melanogaster microbiome and longevity.

Chloroplast genomes in the parasitic sandalwood order

Authors:
WEAVER, T. D.[1], LUU, L.[1], EDLUND, M[2], ROBISON, T.[3], SIMENC, M. C.[1], BONILLA, E.[1], SOSA, J.[1], SU, H[4], NEUBIG, K.[5], PETERSEN, G.[2], NICKRENT, D. L.[5], DER, J. P.[1]
Advisor: Dr. Joshua Der

[1]California State University, Fullerton, USA
[2]Stockholm University, Sweden
[3]Cornell University, USA
[4]University of Taipei, Taiwan
[5]Southern Illinois University, Carbondale, USA

Abstract

Parasitic plants exploit their hosts for water and nutrients, often resulting in a decreased reliance on photosynthetic processes. Increased levels of parasitism and relaxed evolutionary constraint often result in dramatic changes to the chloroplast genome (i.e. plastome). We sequenced plastomes in Santalales to explore the evolution of parasitic plant genomes. Santalales is an order of angiosperms containing over 2000 species in 18 families, with species representing all ranges of parasitism, from nonparasitic to hemiparasitic and holoparasitic. We hypothesize that species with increased specialization and host dependence (e.g. mistletoes and holoparasites) will have lost more plastid genes compared with non-parasitic or generalist root parasites in the order. We used Illumina shotgun sequencing to extract and assemble plastomes for 50 representative species in Santalales. Preliminary assembly and gene annotations were used to construct a phylogeny of 45 species and to examine chloroplast gene loss across the order. A secondary assembly constructed using NovoPlasty resulted in a higher quality genome sequence and was used in subsequent analyses. Genomes were aligned to existing Santalales plastomes using Mauve (implemented in Geneious) and used to manually annotate genes. Our phylogenomic analyses show strong support for major clades in Santalales and are consistent with previously published results; however, relationships among some clades remain unresolved. There is a strong pattern of ndh gene loss associated with increasing levels of parasitism. We present complete plastome assemblies and detailed gene annotations for five species in Santalales.

CHEMISTRY & BIOCHEMISTRY
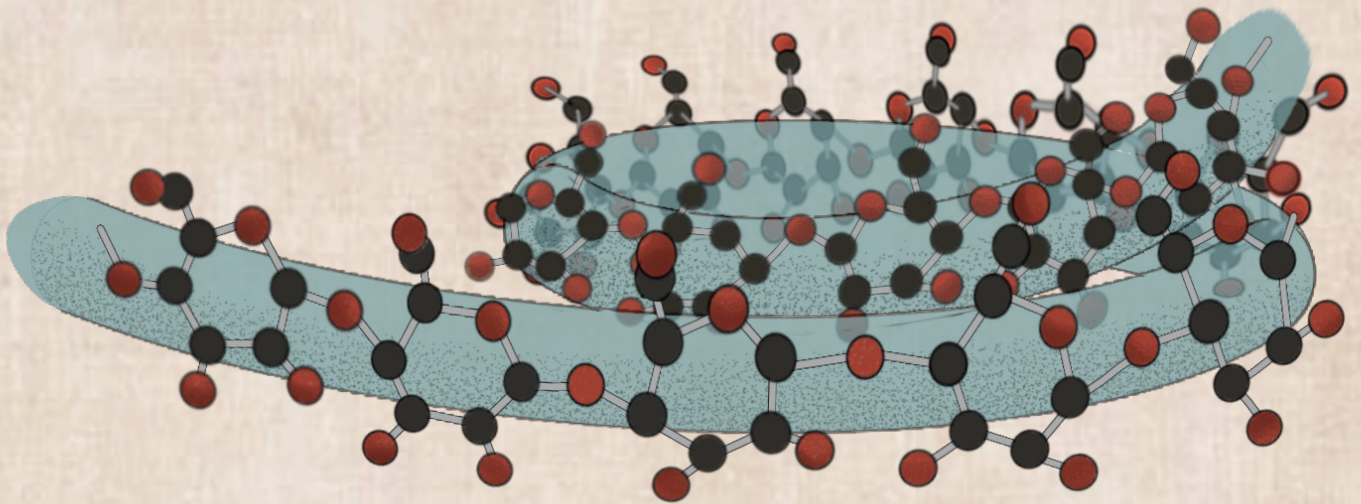
Synthesis and Study of the Pyrochlore $La_3Sb_3Co_2O_{14}$ for Use as an Oxygen Transport Membrane

Authors: Gonzalez Jimenez, Jose L. and Abad, Mark
Advisor: Dr. Allyson M. Fry-Petit
California State University, Fullerton
Department of Chemistry and Biochemistry

Abstract

Oxygen transport membranes (OTM's) are a potential replacement to current expensive cryogenic methods used to obtain pure oxygen from air. The pyrochlore structure $La_3Sb_3Co_2O_{14}$ was synthesized and studied for use as one of these materials due to the hypothesis that it is a mixed ionic-electronic conducting membrane (MIECM). Two methods were used to synthesize half gram samples, the conventional solid state method and the sol-gel method to the discovery that there was no difference in the end product. The relative simplicity of solid state synthesis led to its use for an upscale that was used to run variable oxygen experiments. When heating samples in a reducing atmosphere the structure decomposed into a multiphase structure, and when reheating in an oxygen exposed environment, the original structure was regenerated. The studies also led to the discovery of solid state synthesis' effectiveness limit, as the x-ray diffraction peaks broadened with the upscale. Consequently the sol-gel method was used for further, and current experiments to find the specific temperature and time at which oxygen motion begins in the system. Current iron doping synthetic studies are being done to compare the pyrochlore to a perovskite structure, another potential OTM.

New Approaches for the Treatment of Alzheimer's Disease: Green Chemistry Synthesis and Molecular Modeling Studies of Donepezil-Based Analogs

Author: Stephanie Mariel Salvador
Advisor: Stevan Pecic, Ph.D
California State University, Fullerton
Department of Chemistry and Biochemistry

Abstract

Alzheimer's disease (AD) is a debilitating disorder in the elderly affecting around 50 million people worldwide. No current therapies can prevent, slow, or significantly modify the progress of AD. Therefore, the discovery of new therapies is of utmost importance. A low level of acetylcholine, an important chemical in the brain, is a significant factor in AD progression. One strategy to increase acetylcholine levels in brain regions involved in memory is to inhibit, i.e. block the enzyme acetylcholinesterase (AChE), a naturally occurring biological molecule in our body that is responsible for the degradation of acetylcholine. To date, there are only three FDA approved drugs available for the treatment of AD based on this mechanism; these drugs can provide only palliative treatment of mild to moderate AD and in addition show severe side effects. Using computer software, we designed several potential therapeutics and calculated their drug properties and potential toxicity. We synthesized these analogs using environmentally friendly synthetic routes and we plan to assess their inhibition potential against enzyme AChE.

Introduction

In Alzheimer's disease (AD), the immensely destructive nature of this disease has inspired many researchers to look for a cure. Currently, research is generally focused on slowing its progression, which means that finding a cure is something that still eludes researchers today. However, there are current therapeutics to decrease the symptoms associated with Alzheimer's. According to the Cholinergic hypothesis[9], AD results from a defect in the cholinergic system. One approach (Figure 1) for AD treatment involves drugs that inhibit the activity of acetylcholinesterase (AChE), the main enzyme that metabolize the neurotransmitter acetylcholine.[6]
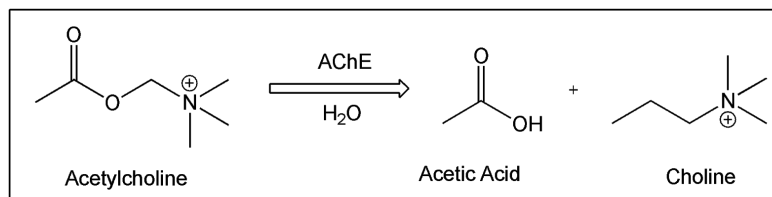
Figure 1. *Acetylcholine hydrolysis catalyzed by AChE*

AChE has two main binding domains, the catalytic site and a peripheral anionic site (PAS).[3] Donepezil or 8-((1-benzyl-piperidin-4-yl)-methyl)-3,4-dimethoxy-bicyclo[4.3.0]nona-1(6),2,4-trien-7-one (brand name Aricept®), an FDA-approved drug, inhibits the enzyme AChE and helps reduce the metabolism of acetylcholine.[4] There are many known AChE inhibitors, some of them naturally occurring compounds- natural products, e.g. galanthamine, nantenine, tacrine, and rivastigmine (Figure. 2) [8] and inhibition of the AChE will eventually increase the levels of acetylcholine in the brain. Unlike many other AChE inhibitors that pose irreversible inhibition, donepezil causes reversible inhibition, which is more valuable in drug discovery process because many enzymes, including AChE enzyme are important for general homeostasis. However, the effects of AD treatment with donepezil are modest, there are many side effects, and in addition, they do not stop full neurodegeneration in Alzheimer's disease.[7]
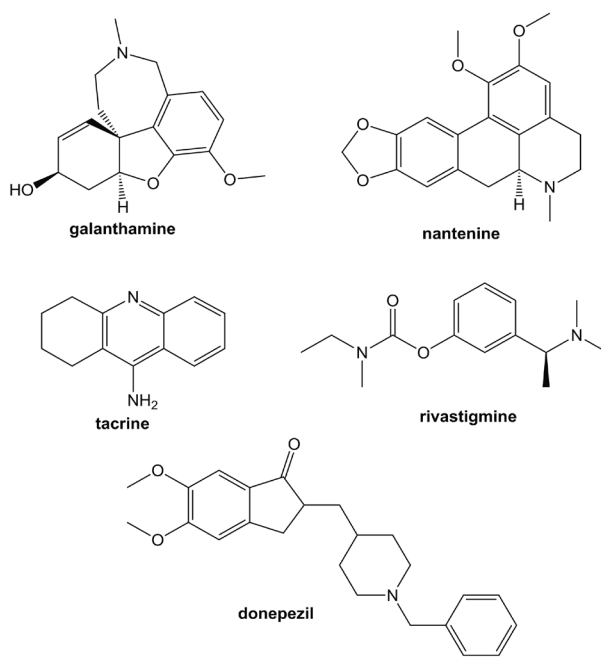
Figure 2. *Known AchE inhibitors*

Exploration of Donepezil's acetylcholinesterase inhibition has been previously published and many SAR studies have been reported.[10] Several different analogs of donepezil were synthesized with various structural changes at Site A (Figure 3), and these modifications led to the significant decrease in the inhibition against rat AChE enzyme. These SAR studies at Site A showed the importance of the initial moiety for the AChE binding and inhibition. The Site B in Donepezil contains a piperidine moiety and modifications in the piperidine ring led to an overall significant decrease in the inhibition activity. Site C has also been explored and several different analogs were synthesized with various substituents on the benzene ring. Several important SARs were observed: ortho-methyl or *para*-methyl (-CH$_3$) groups generally were very well tolerated and had IC$_{50}$s in low nanomolar range, with values of less than 40 nM. On the other hand, an addition of an *ortho*- or *para*- nitro (-NO$_2$) group led to a decrease in inhibition against rat AChE enzyme with IC$_{50}$ values of 160 and 100 nM respectively.
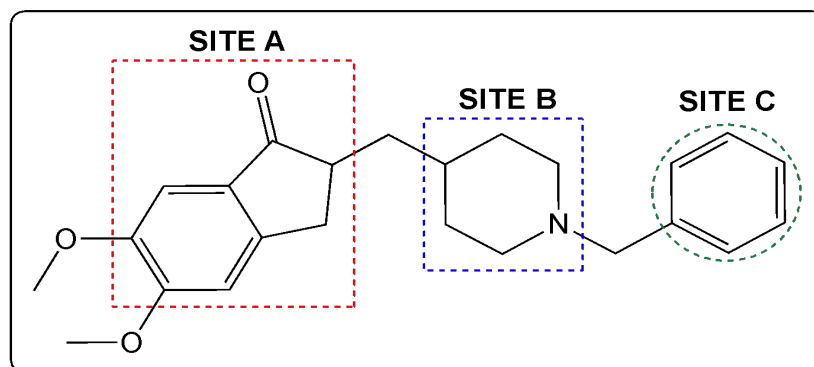
Figure 3. *Known and future SAR studies of Donepezil analogs*

Moreover, molecular modeling studies propose that donepezil and similar analogs bind through the Site C into the binding pocket of AChE and targeting this part of the molecule may lead to discoveries of novel therapeutics.[5] Site C definitely represents a promising area for further exploration since the modifications at this site has the potential to lead to identification of more potent AChE inhibitors.

Herein, we designed several analogs of donepezil at Site C, using available structure-activity relationship (SAR) information obtained from previously published studies, synthesized them using an environmentally friendly synthetic route, purified and fully characterized using proton and carbon NMR spectroscopy. These analogs will be biologically evaluated in AChE inhibition

assay according to Ellman's method.[2] Next, we performed molecular modeling studies and all synthesized compounds were docked into human acetylcholinesterase enzyme model in order to better understand which structural features of molecules are responsible for binding and inhibition properties. We also predicted several important pharmacokinetic properties of these donepezil analogs. These results will guide future synthesis and SAR studies in order to design novel and more potent inhibitors of human AChE enzyme.

Materials and Methods

Reagents and solvents were available from Sigma-Aldrich or Fisher Scientific and used as supplied. All new compounds were characterized by proton NMR and carbon NMR. NMR spectra were measured on a Bruker 400 magnetic resonance spectrometer. ^{1}H chemical shifts are reported relative to the residual solvent peak (chloroform = 7.26 ppm as follows: chemical shift (δ), proton ID, multiplicity (s = singlet, bs = broad singlet, d = doublet, bd = broad doublet, dd = doublet of doublets, t = triplet, q = quartet, m = multiplet, integration, coupling constant(s) in Hz). ^{13}C chemical shifts are reported relative to the residual deuterated solvent ^{13}C signals (chloroform = 77.2 ppm). High-resolution mass spectra were obtained using a high-resolution liquid chromatography mass spectrometer. Microwave reactions were carried out in a CEM Discover SP microwave synthesizer. Molecular modeling studies and docking experiments were performed using ICM Pro Molsoft software.[1]
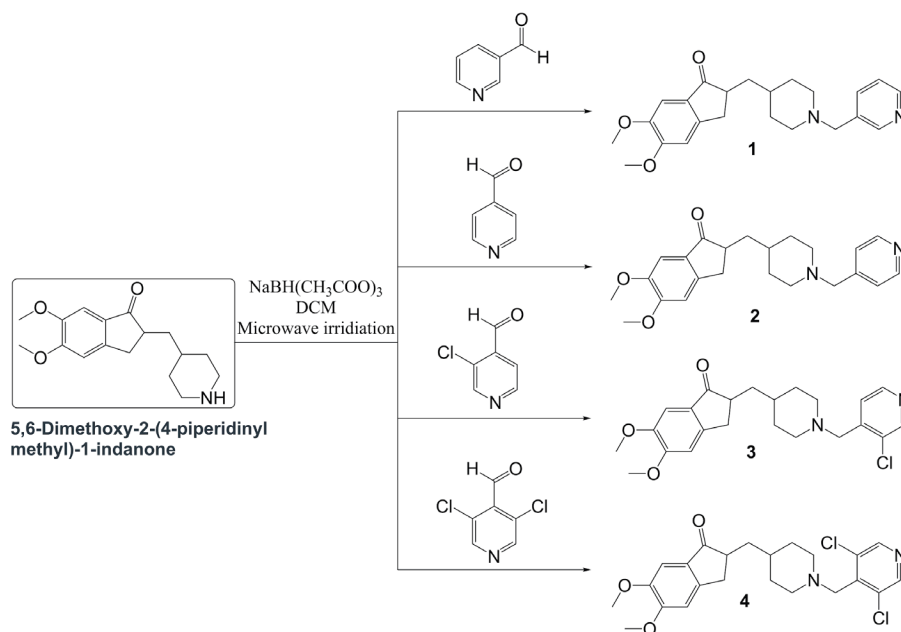
Scheme 1. *A Synthetic Route for the Preparation of Donepezil Analogs*

Results and Discussion

We designed four different donepezil analogs, 1-4 (**Scheme 1**), containing both piperidine (Site B) and the 1-indanone moiety (Site A) of donepezil, which was shown to bind to the peripheral active site (PAS) of AChE[5], and instead of phenyl group at Site C, we introduced pyrimidine ring. We expect that binding of the new analogs to the PAS of AChE will not be affected since we didn't make any alterations in the Site A and Site B, therefore the designed analogs **1**-**4** will provide us with the information whether the pyrimidine moiety in the proximity of catalytic site of AChE would increase the inhibition. Analogs **1**-**4** were synthesized in a microwave-assisted one-step reaction using reductive amination conditions with triacetoxyborohydride as outlined in **Scheme 1**. In short, starting with the commercially available 5,6-dimethoxy-2-(4-piperidinylmethyl)-1-indanone and four different aldehydes, we were able to obtain desired donepezil analogs **1**-**4** in moderate yields. All four synthesized analogs were subjected to purification using flash chromatography, and the final structures were confirmed with proton and carbon NMR spectroscopy.

Next, the binding of the donepezil analogs and AChE was examined *in silico*. All docking studies were performed using the ICM Molsoft package. The ICM score considers several free energy terms such as van der Waals, hydrogen bonding, Poisson electrostatic desolvation and entropy. The crystal structure of *human* AChE bound to donepezil was obtained from the Protein Data Bank (PDB) – PDB code 4EY7. Donepezil was removed and the AChE enzyme model for docking was optimized and generated using the modeling features of the ICM program. After energy minimization, donepezil analogs **1**-**4** were docked into the identified binding region of the enzyme model. Donepezil was also docked for comparison. The docking poses suggest that all four donepezil analogs bind preferentially to the catalytic site since the three most favorable conformations are located in this site. Many non-covalent hydrophobic interactions (**Table 1**), including a hydrogen bond with Phe295 are identified after visual inspection suggesting that these analogs may be strong inhibitors of the AchE enzyme.

Table 1. Non-covalent interactions

Compound	H-bonds	Other non-covalent interactions
1	Phe 295	Tyr 72, Trp 86, Gly 121, Tyr 124, Trp 286, Leu 289, Val 294, Phe 297, Tyr 337, Phe 338, Tyr 341
2	Phe 295	Tyr 72, Trp 86, Tyr 124, Trp 286, Leu 289, Val 294, Phe 297, Tyr 337, Phe 338, Tyr 341, His 447, Gly 448, Ile 451
3	Phe 295	Tyr 72, Asp 74, Trp 86, Tyr 124, Ser 125, Trp 286, Leu 289, Val 294, Phe 297, Tyr 337, Phe 338, Tyr 341
4	Phe 295	Tyr 72, Trp 86, Gly 121, Trp 286, Leu 289, Phe 297, Tyr 337, Phe 338, Tyr 341, His 447, Gly 448, Ile 451, Gly 120
Donepezil	Phe 295	Tyr 72, Asp 74, Trp 86, Gly 121, Tyr 124, Ser 125, Trp 286, Leu 289, Val 294, Phe 297, Tyr 337, Phe 338, Tyr 341

Finally, we calculated and performed predictions of the several pharmacokinetic parameters important for the drug development process, **Table 2**. All four donepezil analogs showed excellent predicted octanol/water partition coefficient (clogP) in the range of 3.68-4.87 (the optimal cLogP for blood-brain barrier permeability is 2-5), good predicted aqueous solubility (acceptable range for clogS is between -6.5 and -0.5 moles/liter) and no unwanted or reactive chemical functionalities (referred as "bad groups" in the table). Other predicted values that we examined were all within optimum ranges, except overall drug likeness (should be within -1 and 1 range), which was close to this range.

Table 2. Predicted ADMET properties

Compound	MW	cLogP	cLogS	Drug Like-ness	Bad Groups	CACO2	HALF- LIFE (h)	HERG Inhi-bition	Tox Score
1	380.488	3.681	-3.302	1.944	0	-5.082	6.72	0.548	0
2	380.488	3.681	-3.79	1.818	0	-5.143	6.72	0.548	0
3	414.93	4.275	-4.66	1.841	0	-5.119	9.13	0.678	0
4	449.372	4.87	-5.59	1.958	0	-5.218	13.8	0.64	0
Donepezil	379.2147	4.73	-4.43	1.557	0	-5.128	6.72	0.457	0

We noticed that even Donepezil, an FDA approved drug, had overall drug likeness above the expected range. Predicted Caco2 permeability indicates high permeability (all values above -5 suggest a high permeability) and this series of analogs had good predicted values for hERG inhibition, namely values above 0.5 indicate high probability for drug candidate of being a hERG inhibitor. All final compounds had values close, but still above 0.5, indicating that we should test these analogs for hERG inhibition if they progress into clinical trials. Our further prediction of plasma half-life showed that these compounds have an excellent predicted half-life, much better than Donepezil, ranging from 6.72 hours for analogs **1** and **2**, up to 9.13 hours and 13.8 hours for analogs **3** and **4**, respectively. In addition, all newly synthesized donepezil analogs had no unfavorable substructure/substituents, shown as a tox score of 0.

In our future work, we plan to evaluate AchE inhibitory activity of these four donepezil analogs. The inhibition potency can be measured quantitatively using microplate assay based on Ellman's method. The enzyme hydrolyzes the substrate acetylthiocholine resulting in the product thiocholine which reacts with Ellman`s reagent (DTNB) to produce 2-nitrobenzoate-5-mercaptothiocholine and 5-thio-2-nitrobenzoate which can be detected at 405 nm.

In conclusion, our work here indicates that the donepezil scaffold represents an excellent starting point to develop new AChE inhibitors.

Experimental

5,6-Dimethoxy-2-(4-piperidinylmethyl)-1-indanone, HCl salt (680.40 mg) was dissolved in 25 mL aqueous solution of saturated sodium bicarbonate and 25 mL of ethyl acetate. It was stirred for approximately 25 minutes and the organic layer was separated. Aqueous layer was washed twice with 25 mL dichloromethane, organic layers were combined, dried over anhydrous sodium sulfate and concentrated. Freebased 5,6-dimethoxy-2-(4-piperidinylmethyl)-1-indanone was obtained as a pale orange solid (350 mg).

3,4-dimethoxy-8-((1-((pyridin-3-yl)-methyl)-piperidin-4-yl)-methyl)-bicyclo[4.3.0]nona-1(6),2,4-trien-7-one, 1: 250 mg (0.86 mmol) of 5,6-dimethoxy-2-(4-piperidinylmethyl)-1-indanone, 476 mg (2.58 mmol) of sodium triacetoxyborohydride, and 73 uL of 3-Pyridinecarboxaldehyde (0.86 mmol) were dissolved in 20 mL of anhydrous dichloromethane, and the reaction mixture was subjected to microwave irradiation at 80 °C for 20 min. The mixture

was transferred to a separatory funnel and the organic layer was washed with aqueous solution of saturated NaHCO$_3$ (20 mL). The organic layer was then dried over anhydrous sodium sulfate, filtered and concentrated. The crude product was purified by column chromatography (2-5% methanol/dichloromethane) and 258.4 mg, 73.1% of **1** was obtained as a white solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.51 (d, J = 1.2 Hz, 2H), 7.31-7.27 (m, 2H), 7.17 (s, 1H), 6.86 (s, 1H), 3.96 (s, 3H), 3.90 (s, 3H), 3.50 (s, 2H), 3.24 (q, J = 1.2 Hz, 1H), 2.88-2.86, (m, 2H), 2.70 (q, J = 14 Hz, 2H), 2.05-1.99 (m, 2H), 1.95-1.88 (m, 1H), 1.77-1.67 (m, 2H), 1.53-1.51 (m, 1H), 1.40-1.33 (m, 3H). ^{13}C NMR (100 MHz, CDCl$_3$): δ 207.7, 155.5, 149.6, 149.4, 148.7, 148.0, 129.2, 123.9, 121.1, 107.4, 104.4, 62.0, 56.2, 56.1, 53.9, 45.4, 38.6, 34.2, 33.3, 32.9, 31.8 ppm.

3,4-dimethoxy-8-((1-((pyridin-4-yl)-methyl)-piperidin-4-yl)-methyl)-bicyclo[4.3.0]nona-1(6),2,4-trien-7-one, 2: 250 mg (0.86 mmol) of 5,6-dimethoxy-2-(4-piperidinylmethyl)-1-indanone, 476 mg (2.58 mmol) of sodium triacetoxyborohydride, and 74 uL of 4-Pyridinecarboxaldehyde (0.86 mmol) were dissolved in 20 mL of anhydrous dichloromethane, and the reaction mixture was subjected to microwave irradiation at 80 °C for 20 min. The mixture was transferred to a separatory funnel and the organic layer was washed with aqueous solution of saturated NaHCO$_3$ (20 mL). The organic layer was then dried over anhydrous sodium sulfate, filtered and concentrated. The crude product was purified by column chromatography (2-5% methanol/dichloromethane) and 181.1 mg, 69.1% of **2** was obtained as a white solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.53-8.50 (m, 2H), 7.24 (d, J=8 Hz, 1H), 7.29-7.26 (m, 1H), 7.17 (s, 1H), 6.86 (s, 1H), 3.96 (s, 3H), 3.90 (s, 3H), 3.55 (s, 2H), 3.24 (q, J =8 Hz, 1H), 2.93-2.89, (m, 2H), 2.72-2.67 (m, 2H), 2.07-2.01 (m, 3H), 1.94-1.88 (m, 1H), 1.77-1.68 (m, 2H), 1.58-1.48 (m, 1H), 1.40-1.30 (m, 3H). ^{13}C NMR (100 MHz, CDCl$_3$): δ 155.6, 150.4, 149.5, 148.8, 148.5, 137.2, 133.4, 129.3, 123.5 107.4, 104.5, 60.3, 56.3, 56.2, 53.6, 45.4, 38.7, 34.2, 33.4, 32.7, 31.6 ppm.

8-((1-((3-chloro-pyridin-4-yl)-methyl)-piperidin-4-yl)-methyl)-3,4-dimethoxy-bicyclo[4.3.0]nona-1(6),2,4-trien-7-one, 3: 250 mg (0.86 mmol) of 5,6-dimethoxy-2-(4-piperidinylmethyl)-1-indanone, 546 mg (2.58 mmol) of sodium triacetoxyborohydride, and 123 mg of 3-Chloro-4-pyridinecarboxaldehyde (0.86 mmol) were dissolved in 20 mL of anhydrous dichloromethane, and the reaction mixture was subjected to microwave irradiation at 80 °C for 20 min. The mixture was transferred to a separatory funnel and the organic layer was washed with aqueous solution of saturated NaHCO$_3$ (20 mL). The organic layer was then dried over anhydrous sodium sulfate,

filtered and concentrated. The crude product was purified by column chromatography (2% methanol/dichloromethane) and 100 mg, 30% of **3** was obtained as a white solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.50 (s, 1H), 8.43 (d, J = 4.8 Hz, 1H), 7.48 (d, J = 5 Hz, 1H), 6.86 (s, 1H), 7.16 (s, 1H), 6.85 (s, 1H), 3.95 (s, 3H), 3.90 (s, 3H), 3.58, (s, 2H), 3.24 (q, J =8.4 Hz, 1H), 2.89-2.84 (m, 2H), 2.70 (d, J = 14.4 Hz, 2H), 2.11 (t, J = 11.2 Hz 2H), 1.96-1.89 (m, 1H), 1.78-1.67 (m, 2H), 1.51 (s, 1H), 1.43-1.31 (m, 3H). ^{13}C NMR (100 MHz, CDCl$_3$): δ 207.8, 155.7, 149.7, 149.3, 148.9, 147.9, 146.0, 131.9, 129.5, 124.5, 107.5, 104.6, 58.7, 56.4, 56.3, 54.3, 45.6, 38.9, 34.4, 33.5, 33.3, 32.1 ppm.

8-((1-((3,5-dichloro-pyridin-4-yl)-methyl)-piperidin-4-yl)-methyl)-3,4-dimethoxy-bicyclo[4.3.0]nona-1(6),2,4-trien-7-one, 4: 250 mg (0.86 mmol) of 5,6-dimethoxy-2-(4-piperidinylmethyl)-1-indanone, 476 mg (2.58 mmol) of sodium triacetoxyborohydride, and 151.35 mg of 3,5- Dichloro-4-pyridinecarboxaldehyde (0.86 mmol) were dissolved in 20 mL of anhydrous dichloromethane, and the reaction mixture was subjected to microwave irradiation at 80 °C for 30 min. The mixture was transferred to a separatory funnel and the organic layer was washed with aqueous solution of saturated NaHCO$_3$ (20 mL). The organic layer was then dried over anhydrous sodium sulfate, filtered and concentrated. The crude product was purified by column chromatography (0-2% methanol/dichloromethane) and 70 mg, 16.9% of **4** was obtained as a white solid. ^{1}H NMR (400 MHz, CDCl$_3$) δ 8.45 (s, 2H), 7.16 (s, 1H), 7.48 (d, J = 5 Hz, 1H), 6.84 (s, 1H), 3.95 (s, 3H), 3.89 (s, 3H), 3.69, (s, 2H), 3.22 (q, J =8.4 Hz, 1H), 2.89-2.85 (m, 2H), 2.71-2.66 (m, 2H), 2.26-2.19 (m, 2H), 1.91-1.85 (m, 1H), 1.72-1.62 (m, 3H), 1.55-1.44 (s, 1H), 1.33-1.20 (m, 4H). ^{13}C NMR (100 MHz, CDCl$_3$): δ 207.8, 155.6, 149.5, 148.8, 147.8, 143.3, 134.1, 129.4, 107.4, 104.5, 100.1, 56.3, 56.2, 56.1, 54.1, 45.5, 38.7, 34.4, 33.4, 33.1, 31.9 ppm.

References

1.   Abagyan, R.; Totrov, M., Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J Mol Biol* **1994,** *235* (3), 983-1002.

2.   Adsersen, A.; Kjolbye, A.; Dall, O.; Jager, A. K., Acetylcholinesterase and butyrylcholinesterase inhibitory compounds from Corydalis cava Schweigg. & Kort. *J Ethnopharmacol* **2007,** *113* (1), 179-82.

3.   Bolognesi, M. L.; Minarini, A.; Rosini, M.; Tumiatti, V.; Melchiorre, C., From dual binding site acetylcholinesterase inhibitors to multi-target-directed ligands (MTDLs): a step forward in the treatment of Alzheimer's disease. *Mini Rev Med Chem* **2008,** *8* (10), 960-7.
4.   Cacabelos, R., Donepezil in Alzheimer's disease: From conventional trials to pharmacogenetics. *Neuropsychiatr Dis Treat* **2007,** *3* (3), 303-33.

5.   Cavalli, A.; Bottegoni, G.; Raco, C.; De Vivo, M.; Recanatini, M., A computational study of the binding of propidium to the peripheral anionic site of human acetylcholinesterase. *J Med Chem* **2004,** *47* (16), 3991-9.

6.   Colovic, M. B.; Krstic, D. Z.; Lazarevic-Pasti, T. D.; Bondzic, A. M.; Vasic, V. M., Acetylcholinesterase inhibitors: pharmacology and toxicology. *Curr Neuropharmacol* **2013,** *11* (3), 315-35.

7.   Dunn, N. R.; Pearce, G. L.; Shakir, S. A., Adverse effects associated with the use of donepezil in general practice in England. *J Psychopharmacol* **2000,** *14* (4), 406-8.

8.   Pecic, S.; McAnuff, M. A.; Harding, W. W., Nantenine as an acetylcholinesterase inhibitor: SAR, enzyme kinetics and molecular modeling investigations. *J Enzyme Inhib Med Chem* **2011,** *26* (1), 46-55.

9.   Sivaprakasam, K., Towards a unifying hypothesis of Alzheimer's disease: cholinergic system linked to plaques, tangles and neuroinflammation. *Curr Med Chem* **2006,** *13* (18), 2179-88.

10.   Sugimoto, H.; Iimura, Y.; Yamanishi, Y.; Yamatsu, K., Synthesis and structure-activity relationships of acetylcholinesterase inhibitors: 1-benzyl-4-[(5,6-dimethoxy-1-oxoindan-2-yl) methyl]piperidine hydrochloride and related compounds. *J Med Chem* **1995,** *38* (24), 4821-9.

Human Soluble Epoxide Hydrolase Enzyme - Molecular Modeling Studies and Docking Experiments

Author: Aleksei Solomatov
Advisor: Stevan Pecic, Ph.D
California State University, Fullerton
Department of Chemistry and Biochemistry

Abstract

The objective of this experiment is to validate *in silico* model of human soluble epoxide hydrolase enzyme (sEH) for future docking experiments. The model was built using a crystal structure of the enzyme provided by a protein database (RCSB Protein Database) and ICM Pro molecular modeling software, and validation of the model was obtained using an analysis verification server from UCLA. After observations of the five validation criteria given from the server, the model was used for the docking experiment. A known sEH inhibitor AUDA was docked using ICM Pro software, and energy scores were retrieved, and docking poses were analyzed. The docking results show that our *in silico* model for sEH is suitable as a template for future docking experiments and structural-based drug discovery.

Introduction

Human soluble epoxide hydrolase (sEH) enzyme is a type of enzyme that converts epoxide groups of naturally occurring fatty acids to corresponding alcohol groups by the addition of water. This enzyme is widely distributed in mammalian tissues where the highest concentrations are located in the liver, kidney, intestine, and vascular tissue[1]. Epoxyeicosatrienic acids (EETs) are one of three derivatives of arachidonic acid, which is synthesized from cytochrome P450 epoxygenase (CYP) shown in *Figure 1*.
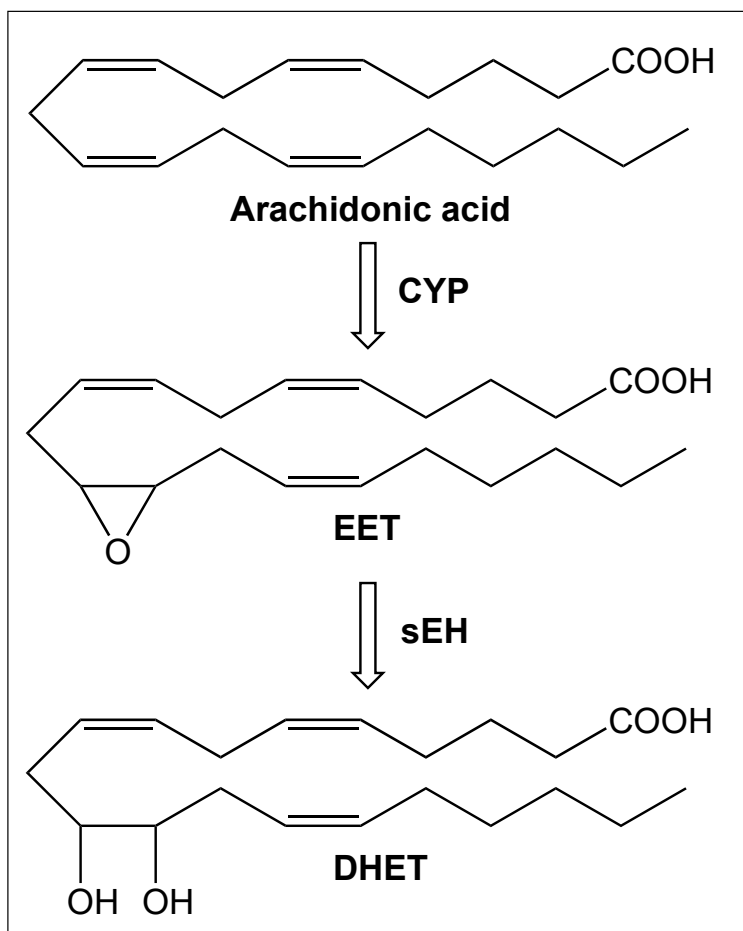
Figure 1. *Metabolic degradation of Arachidonic acid*

EETs regulate ion transport and gene expression by producing anti-inflammatory, antihypertensive, antinociceptive, and pro-fibrinolytic effects.[1] However, sEH converts EETs to dihydroxyeicosatrieoic acids (DHETs), thereby diminishing the biological activity of EETs. Inhibition of sEH, therefore, represents an attractive approach to increase EET levels, which can be beneficial for cardiovascular diseases, pain, inflammation, diabetes, and kidney disease.[2] The X-ray crystallographic structure of human sEH and a known inhibitor (AUDA), shown in *Figure 2*, complex revealed the catalytic pocket and key structural features required to inhibit the sEH enzyme. Two tyrosines (Tyr383and Tyr466) and one aspartic acid (Asp335) residues, located in the hydrolase catalytic pocket of sEH, are involved in the degradation of EET - tyrosine residues act as hydrogen bond donors to promote the epoxide ring-opening by Asp335.[2]
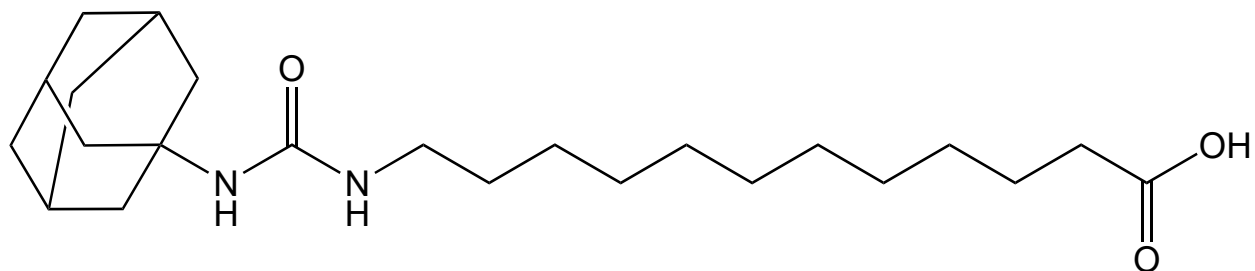
Figure 2. *12-3(3-adamantan-1ylureido) dodecanoic acid (AUDA)*

As seen in *Figure 2*, AUDA[3] has a urea moiety connected to the hydrophobic adamantly ring. The urea group allows the inhibitor to bind to sEH and is a vital part of the compound. However, the urea-based inhibitors often suffer from poor solubility and stability, which hinders their pharmacological use *in vivo*.[4]

Our research interest is to design new, non-urea compounds that will inhibit enzyme sEH, which will eventually be used as therapeutics for various medical conditions, including pain and inflammation. The goal of this project is to prepare and test a valid *in silico* model for human sEH enzyme, which will be used for future docking experiments and will help us to guide the future design of sEH inhibitors.

METHODS

***In silico* Model Building, Model Evaluation, and Docking Experiments**

X-ray crystallographic structure of human sEH enzyme was retrieved from RCSB Protein Data Bank- PDB code: 4HAI.[5] The PDB was then brought into ICM Pro and converted into an ICM file. Conversions to the ICM model were performed by optimization of hydrogens, certain amino acids (H, P, N, C, Q), and ligand tautomer, assignment of the secondary structure, and approximations of the formal charges of the ligands. The original PDB file had an inhibitor named **ai23**, n-cycloheptyl-1-[(2,4,6-trimethylphenyl) sulfonyl] piperidine-4-carboxamide, which was removed. The *in silico* model of sEH was then converted back to the PDB file and uploaded to the verification server. Evaluations were provided via a server of the UCLA DOE Institute for Genomic and Proteomics. Verify3D[6], ERRAT[7], and PROCHECK[8], WHAT-IF[9], and PROVE[10]

were used to structurally verify and provide additional data on the *in silico* model.[11] Once verification scores were retrieved, the model was brought back into ICM Pro, where the model's binding pockets were identified and listed. An amino acid sequence of Met418, Tyr381, Asp333, and His337[2] was used as the key amino acids for defining the binding pocket. The structure of the known sEH inhibitor AUDA was drawn in ChemDraw and energy minimized using ICM Pro and saved. The AUDA was then docked into *in silico* model, which brought the E-DOCK scores, and a 2D model was obtained to display the inhibitor's interactions with the *in silico* model's binding pocket.

DISCUSSION

Model Evaluation:

The *in silico* model was evaluated first with the program Procheck. PROCHECK is used to check the detailed residue-by-residue stereochemical quality of a model structure. The Ramachandran plot gives us an analysis of the geometrical aspects of the model by using a plot of φ-ψ torsion angles for all residues seen. The Ramachandran plot can be found in **Figure 3.** As shown in the plot statistics, the residues in most favored regions of [A, B, L] display that our model has a high amount of allowed regions (>90% residues). The Goodness factor or G-Factor shows the quality of the model by its dihedral, covalent, and bond angles, which should stay above the threshold of -0.5 to be a dependable model. In our case, the overall G-factor parameter value is at 0.2, which shows that our bond angles are right. Upon additional observation, the Ramachandran plot also has 89.1% of residues being in the most favorable position, while another 10.1% of residues are in additional allowed regions. These percentages are suitable enough to show the Procheck evaluation as a notable pass for the *in silico* sEH model.
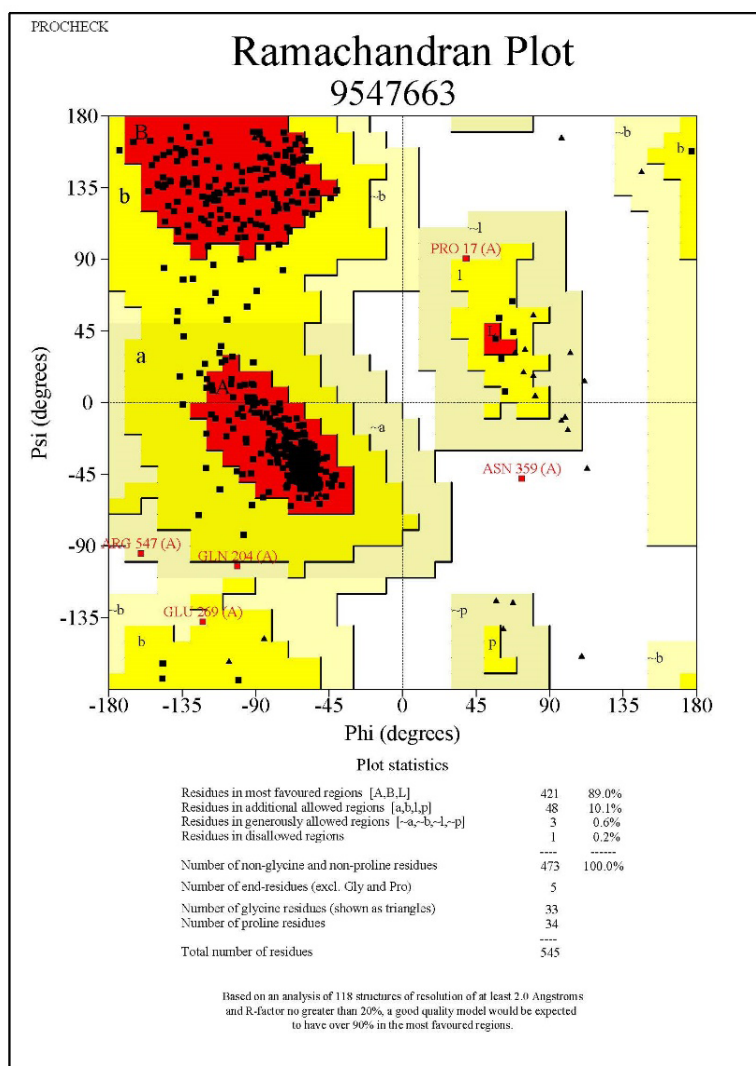
Figure 3. *Ramachandran Plot of sEH in silico model*

The ERRAT evaluation results are shown in **Figure 4.** ERRAT is a so-called "overall quality factor" for nonbonded atomic interactions (CC, CN, CO, NN, NO, and OO), and higher scores mean higher quality, which is displayed by both thresholds of 95% warning and 99% error. Structures that present value above 90% are deemed suitable by the ERRAT test. The in silico model of human sEH, PDB: 4HAI, has an overall quality factor of 96.9298, which is an excellent score to deem the structure of this model as valid.
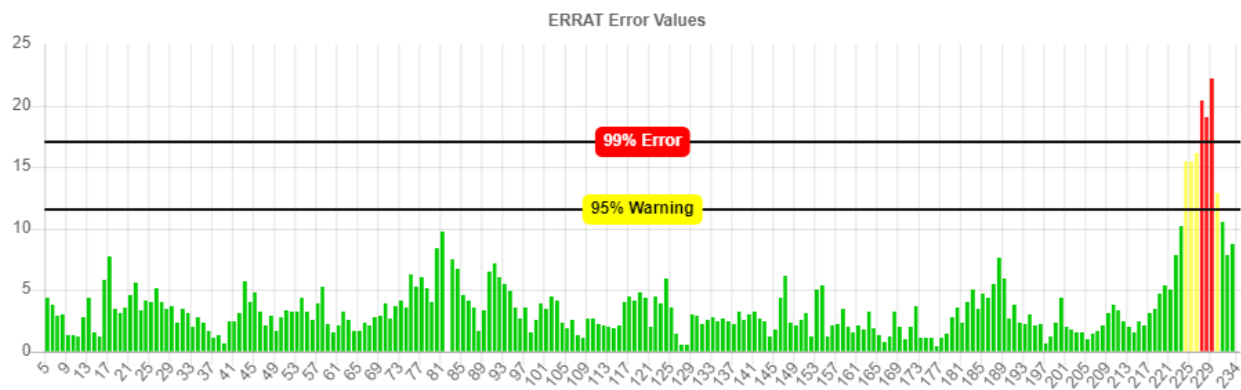
Figure 4. *ERRAT error values of sEH in silico model*

The Prove evaluation results are shown in **Figure 5.** The graphs give an average volume Z-score of all the atoms in the *in silico* sEH structure. Z-score is the calculated difference between the volume of the atom and the average atomic volume of the corresponding atom type, which is then divided by the standard deviation of the appropriate distribution. Z-score standard deviation (Zstd) scores the atoms in the test proteins, which describe the standard deviation of their volume to the average (Zsm) for the type of atom. At the same time, a structural Z-score RMS (Zrms) determines the structural quality of the model. Zrms is calculated by taking the square root of the total scores. The boundary for Zrms scores is 1.2 with anything above, resulting in a poor resolution. In order to have a successful Prove test, our model must be around the 1.0 area. As shown in **Figure 5.** The Zsm score is near 1.0 with some deviation. Also, our model does have 3.5% total buried outlier atoms, which validates the success of the model since the caution range is 1-5% with an error range of anything above 5%.
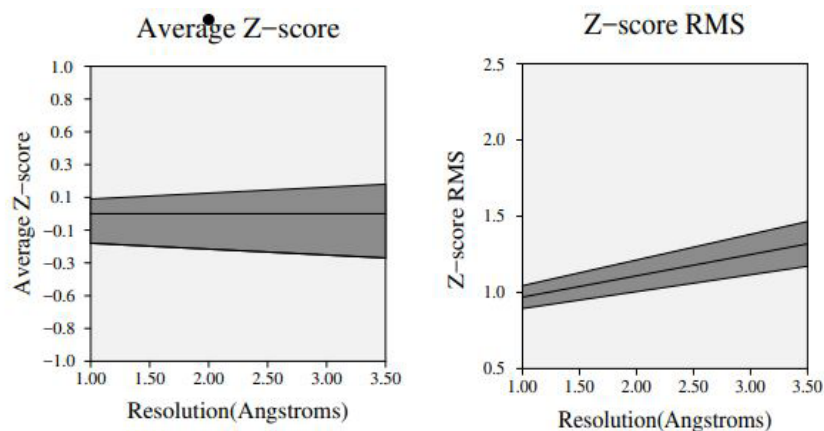
Figure 5. *Z-score results from Prove test of in silico sEH model*

The results of the Verify 3D evaluation is shown in **Figure 6.** VERIFY-3D uses energetic and empirical methods to produce averaged data points for each residue to evaluate the quality of protein structures. Using this scoring function, if more than 80% of the residue has a score of >0.2, then the protein structure is considered high quality. The program works best on proteins with at least 100 residues and measures the score of the compatibility of a model and its sequence. The results obtained represent the average data score of the *in silico* sEH model in comparison with the template of Verify 3D. The data in **Figure 6** show that 84.40% of residues have an average 3D-1D score of >= 0.2, meaning that at least 80% of amino acids have scored less than or equal to 0.2 percent.
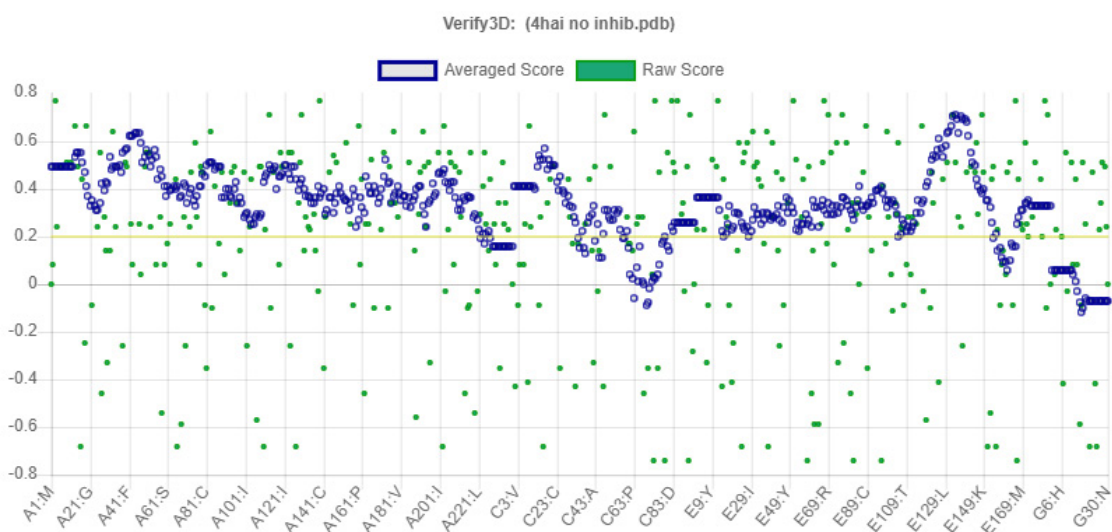
Figure 6. *Verify3D of in silico sEH model*

WHAT CHECK is a program that evaluates: bond lengths, bond angles, omega angle restraints, side-chain polarity, improper dihedral distribution, and inside/outside distribution. The summary gives a useful report for proper structure analysis with careful calculations. **Table 1.** Shows an organized table of all the scores found in the WHAT CHECK analysis. The first part of the table from 1st generation packing quality to RMS Z-scores should be positive; however, having a negative score doesn't mean the structure fails the analysis. Positive score only shows that the score is above average, while negative can just deem the structure scores as average. The second part (ranges from Bond lengths to side-chain planarity) displays how well the *in silico* sEH model adjusts to common refinement constraint values, which should have a number close to 1.0 and

as shown by the results significant of the scores stay relatively close to 1.0. Lastly, there is one abnormality in the B-factor distribution and backbone conformation. It can be suggested that the backbone conformation can't be aligned with any other backbone structures in the database, which is the same possibility for the B-factor distribution. However, with these scores, the overall summary report of the WHAT CHECK results in a pass, validating this *in silico* model.

Table 1. *WHAT CHECK for in silico model.*

Overall Summary Report: PASS	
1st generation packing quality	-0.869
Ramachandran plot appearance	-2.609
Chi-1/Chi-2 rotamer normality	-2.066
Backbone conformation	-32.335
Bond Lengths	0.515
Bond Angles	0.677
Omega angle restraints	1.314
Side Chain Planarity	0.462
B-factor distribution	54.326

Docking Experiments:

To further validate the *in silico* model of the sEH enzyme, molecular docking experiments were performed with the program ICM Pro. The known inhibitor AUDA was interactively docked, and the E-dock score of -33.19 was obtained after this procedure. The 2D diagram in **Figure 7.** shows the amino acid interactions with the AUDA inhibitor in the catalytic site of the *in silico* model of the sEH enzyme. The green and grey colors show the noncovalent hydrophobic interactions between the surrounding amino acids and the inhibitor AUDA. Inspection of the top, binding pose in the catalytic site indicates that AUDA interacts with the enzyme via noncovalent interactions between adamantyl moiety of AUDA, and residues Proline 229 and Arginine 103. Two hydrogen bonds were observed: the one between the oxygen of the urea functional group of the inhibitor AUDA and Threonine 230, and the second one between the carboxylic acid moiety of AUDA and Glutamine 138. Other hydrophobic interactions help to stabilize the AUDA in the catalytic site. All interactions are listed in **Table 2.** with the distances in Angstroms.
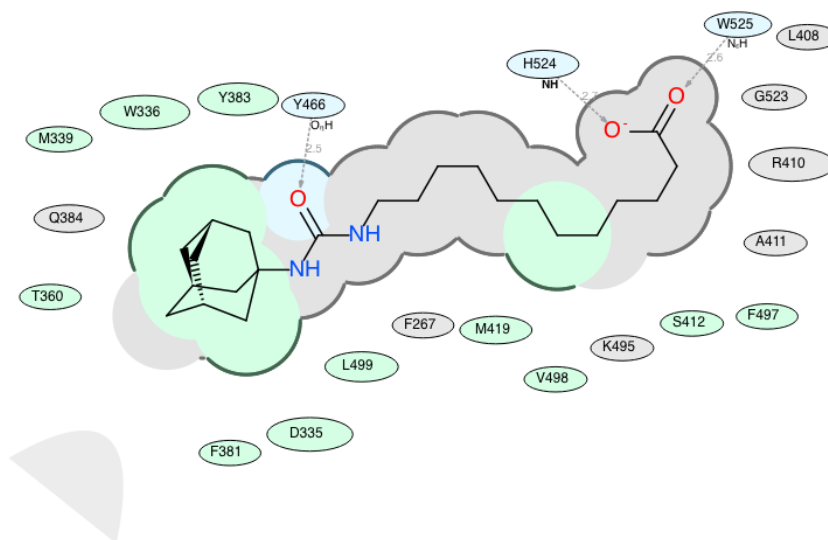
Figure 7. *The binding pose of the inhibitor AUDA in the catalytic site of the human sEH model*

Table 2. *Non-covalent interactions of AUDA inhibitor within the catalytic pocket.*

Resi-due	Size	Type	at	Distance
W525	1.033264	hbond	o1	2.564312
H524	1.337716	hbond	o2	2.729384
Y466	0.937354	hbond	o3	2.548728
D335	2.199965	hydrophob	c21	4.383119
W336	2.611098	hydrophob	c14	3.220187
M339	0.692859	hydrophob	c20	4.452448
T360	0.686042	hydrophob	c22	3.889638
F381	0.748289	hydrophob	c17	3.974089
Y383	1.532805	hydrophob	c12	4.476076
S412	1.02566	hydrophob	c3	3.868343
M419	1.315517	hydrophob	c6	2.692708
F497	0.709106	hydrophob	c3	4.08035
V498	0.726371	hydrophob	c6	3.842652
L499	1.427013	hydrophob	c17	3.578437

Conclusion

In the present investigation, we have simulated and evaluated *in silico* model of the sEH enzyme, based on the crystal structure of human soluble epoxide hydrolase complexed with N-cycloheptyl-1-(methylsulfonyl)piperidine-4-carboxamide. The visual inspection of docked AUDA and evaluation of docking results in **Figure 7.** and **Table 2.** indicate that our model has significant H-bonds interaction with Tyr 466, and the urea moiety is in the proximity of Asp 355 and Tyr 383 residues. Also, two additional hydrogen bonds are stabilizing this complex. Overall, AUDA has been perfectly docked, which helps represent an excellent template for future docking experiments. The docking experiments of known sEH inhibitor AUDA, however, didn't correlate with the previously published data. Namely, the obtained docking score (E-dock) is in the acceptable range. Still, we observed that the fundamental interactions important for potent sEH inhibition, Tyrosine 383, Tyrosine 466, and Aspartic acid 335 residues are not directly involved in the binding with the AUDA in our model. Still, these key residues are in the proximity, and our future studies will be focused on improving our *in silico* model and adjusting the binding pocket around the catalytic site. The detailed structural understanding will help the structure-based design of novel sEH inhibitors.

References

1.   Spector, A. A.; Fang, X.; Snyder, G. D.; Weintraub, N. L., Epoxyeicosatrienoic acids (EETs): metabolism and biochemical function. *Prog Lipid Res* **2004,** *43* (1), 55-90.

2.   Imig, J. D.; Hammock, B. D., Soluble epoxide hydrolase as a therapeutic target for cardiovascular diseases. *Nat Rev Drug Discov* **2009,** *8* (10), 794-805.

3.   Zhang, J.; Liu, Y. S.; Lu, Q. H., Therapeutic effects of the soluble epoxide hydrolase (sEH) inhibitor AUDA on atherosclerotic diseases. *Pharmazie* **2015,** *70* (1), 24-8.

4.   Pecic, S.; Deng, S. X.; Morisseau, C.; Hammock, B. D.; Landry, D. W., Design, synthesis and evaluation of non-urea inhibitors of soluble epoxide hydrolase. *Bioorg Med Chem Lett* **2012,** *22* (1), 601-5.

5.   Pecic, S.; Zeki, A. A.; Xu, X.; Jin, G. Y.; Zhang, S.; Kodani, S.; Halim, M.; Morisseau, C.; Hammock, B. D.; Deng, S. X., Novel piperidine-derived amide sEH inhibitors as mediators of lipid metabolism with improved stability. *Prostaglandins Other Lipid Mediat* **2018,** *136*, 90-95.

6.   Eisenberg, D.; Luthy, R.; Bowie, J. U., VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods Enzymol* **1997,** *277*, 396-404.

7.   Colovos, C.; Yeates, T. O., Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* **1993,** *2* (9), 1511-9.

8.   Laskowski, R. A.; Rullmannn, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M., AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **1996,** *8* (4), 477-86.

9.   Vriend, G.; Sander, C., Quality control of protein models: directional atomic contact analysis. *Journal of Applied Crystallography* **1993,** *26* (1), 47-60.

10.   Pontius, J.; Richelle, J.; Wodak, S. J., Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol* **1996,** *264* (1), 121-36.

11.   Bowie, J. U.; Luthy, R.; Eisenberg, D., A method to identify protein sequences that fold into a known three-dimensional structure. *Science* **1991,** *253* (5016), 164-70.

Purification and characterization of a small copper carrier in blood fluid

Authors: Theodore Roque and Heather Carrillo
Advisor: Dr. Maria Linder
California State University, Fullerton
Department of Chemistry and Biochemistry

Abstract

Copper (Cu) is an essential trace element that is maintained at nearly constant levels in the blood. In mammals, ATP7B is a Cu pump that enables the process of excreting excess Cu via the bile. Previous studies have shown that in Cu-overloaded systems where ATP7B is defective, higher than normal Cu concentrations are found in the urine, indicating that there may be an alternate exit route for Cu through the kidneys. Past investigations revealed that a small Cu carrier (SCC) is implicated in this process, and our laboratory found it was also present in the blood plasma. SCC was found to be around 1700 Da in size, but little else is known about its chemical nature. As such, current studies are focused on determining a sequence of effective purification methods for SCC with the goal of isolating large amounts of the molecule for further analysis by techniques such as mass spectrometry and solution NMR, for full structure determination. For this purpose, purification trials were performed on commercially available pig plasma, which we found had higher concentrations of SCC than in other species. Thus far, our data show that one effective method of purifying pig plasma SCC begins with ultrafiltration of the plasma using a 10 kDa cutoff filter. This is followed by another round of ultrafiltration using a 3 kDa filter. The resulting 3 kDa filtrate (containing molecules smaller than 3000 Da), typically has Cu concentrations of around 30 ng/mL. This is then dialyzed over several days in doubly-deionized water (ddH2O) with filter tubing containing 0.1 – 0.5 kDa cutoff pores (Repligen). Such filters improved Cu/A220 ratios from 6 to about 20, with a recovery of about 70% of the SCC-Cu. Efforts toward scaling up this process resulted in the ability to accommodate more than 10 times the amount of pig plasma originally used, and thousands of ng of Cu. Following this, some other separation methods were tested including native polyacrylamide gel electrophoresis (native PAGE) in 10% acrylamide gels, small pore size exclusion chromatography (SEC) using 100 mL Biogel P4 columns and ddH2O, and hydrophilic interaction liquid chromatography (HILIC). Native PAGE showed the SCC-Cu migrated slowly through the gel, indicating a low charge/

mass ratio, but offered poor recovery, while SEC has yielded more promising results with 50% recovery and average Cu/A220 ratios of about 233 across two Cu peaks. Trials using HILIC as a purification method have produced slightly lower recovery than those for SEC but still greater than those for Native PAGE. Preliminary SEC and HILIC results showing multiple peaks also suggest that SCC may be an aggregate molecule which warrants further studies. Overall, our purification results consistently show high SCC-Cu recovery and purity, suggesting it can be used as a standard protocol for SCC isolation.

Introduction

Copper (Cu) is an essential trace element that plays several different roles in the body. It is used, for example, as a cofactor in Complex IV of the electron transport chain, enabling aerobic respiration. It is also a cofactor in superoxide dismutase and thus plays a role in neutralizing reactive oxygen species. The typical amount of Cu in an average 70 kg human adult is around 110 mg and it has been observed that the concentration of Cu in blood and tissue rarely changes. This is the case in humans and most other mammals. The circulation and regulation of Cu in a healthy human adult involves mainly the gut, liver, and kidney along with Cu-carrying plasma proteins like ceruloplasmin and albumin. About 1 mg of Cu is consumed from the diet. As Cu travels down the esophagus and into the stomach, it is joined by Cu from various gastric secretions such as saliva (0.15 mg) and gastric juices (0.15 mg). In the duodenum, biliary (0.6-6.0 mg) and pancreatic (0.5 mg) Cu are secreted into the Cu pool after which some of the Cu is re-absorbed through enterocytes. Most of the Cu entering the blood at this point bind to plasma proteins and make their way either to parts of the body where Cu is most needed or to the liver where the process of excreting excess Cu begins.

Indeed, the primary mechanism for Cu excretion and thus Cu homeostasis, occurs in the liver, relying mainly on a Cu pump called ATP7B. This enzyme is found mostly in hepatocytes on the membrane of the trans-Golgi network (TGN). As Cu approaches hepatocytes by way of Cu-carrying plasma proteins, Cu is reduced and taken in to the cell by membrane enzyme copper transporter 1 (CTR1). A chaperone protein, ATOX1, brings Cu from CTR1 to the TGN where ATB7B can pump it into the TGN lumen. Vesicles budding off of the TGN membrane that contain Cu and ATP7B can find their way to the membrane of the hepatocyte that forms the bile canaliculus leading to a bile duct and ultimately the gall bladder. Cu is then released into the bile.

Notably, Cu in the bile does not re-enter the exchangeable Cu pool as it travels again through the intestinal tract. Additionally, ATP7B found in the vesicles from the TGN membrane can fuse with the hepatocyte membrane adjacent to the bile canaliculi. This allows ATOX1 carrying excess Cu from the blood to bypass the TGN and deposit Cu to these ATP7B pumps directly for excretion via the bile.

Previous studies have shown that in Cu-overloaded mammals, the typical route for Cu excretion changes. In Wilson's Disease, the most well-studied Cu-overload condition, a mutation in the ATP7B gene hinders fecal extraction of Cu via the bile. As such, patients with Wilson's Disease can suffer Cu toxicosis as Cu accumulates in the liver. Indeed, liver damage is often observed in such patients and brain and eye damage can occur as well if the condition is not properly treated, typically using Cu chelation therapy. The damage is thought to be due to increased levels of reactive oxygen species that can be produced by Cu (and Fe) ions through Fenton chemistry.

It has been observed that in Wilson's Disease mice, there are higher than normal Cu concentrations in the urine, indicating that there is an alternate route for Cu excretion when the typical mechanisms for biliary excretion are not functional. Further, this implicates a small copper carrier molecule (SCC) in the process of urinary Cu excretion as the glomerulus of the kidney filters out materials larger than 30 kDa in size (ceruloplasmin is ~130 kDa in size, albumin ~70 kDa). Previous studies using size exclusion chromatography in a small pore gel column (Superdex 30 Increase) showed that SCC has a molecular weight of about 1700 Da, which would be able to pass through the glomerulus. Additional characterization studies revealed that SCC has a negative charge at neutral pH, has a mass-to-charge ratio of 845, and contains Cu^{2+} bound to N and O ligands. However, there is still much to discover about the structure of SCC.

Thus, our primary aim is to elucidate the structure and function of SCC. Efforts currently are focused on establishing a sequence of steps for large-scale purification of SCC pig plasma. Commercially available blood plasma from pigs was chosen for this study because we have found pigs to have higher concentrations of SCC-Cu in their plasma compared to that of other species. The most pure samples will be combined and analyzed using mass spectrometry (including electrospray), EPR, and solution NMR for full structure determination.

Methods

Dialysis

Dialysis was performed on samples of 3 kDa ultrafiltered pig plasma. Volumes of 3 kDa filtrate ranged from 10 mL to 140 mL. Samples were secured inside of tubing by using two clips at each end. The tubing used, manufactured by Repligen, has a molecular weight cutoff of 0.1 - 0.5 kDa and a flat width of 31 mm. Tubes containing sample were then completely submerged in a container of ddH2O and stirred continuously. The volume of ddH2O used was around 100 times the amount of sample being dialyzed. A typical dialysis procedure lasted two days with the dialysis solution being refreshed at intervals of 30 min, 30 min, 1 hour, and overnight. The volume of dialyzed sample often grew to twice the original volume due to osmosis.

Lyophilization

Lyophilization, using a VirTis Benchtop Lyophilizer, required samples to be frozen to at least –80 degrees Celsius before being applied to the apparatus. The sample was lyophilized to complete dryness which required a few days to a week, depending on the amount of sample being frozen. The volume of sample typically used at this step ranged from 15 mL to a few hundred milliliters.

Native PAGE

Native PAGE was performed using tris-glycine as upper and lower tank buffers. The upper tank also included 1 mL of 0.01% bromophenol blue as the tracking dye. Gels were composed of 10% acrylamide and polymerized inside of plastic tubes. After tubes were added to the electrophoresis unit, 20 mA of electrical current was applied (5 mA per tube for four tubes being run simultaneously, typically). Following electrophoresis, each tube was sliced into eight sections roughly equal in size. In an effort to elute any Cu present within the gel, each portion was submerged in 1 mL of ddH2O for a few hours. Each of the resulting solutions were subsequently measured for Cu.

Size Exclusion Chromatography (SEC)

Separation using SEC involved a 100-mL Biogel P4 column equilibrated with ddH2O. Sample volumes that were added to the column were typically 1 mL and contained around 1000 ng of Cu. Samples, at times, required centrifugation in order to remove any precipitates that were present after reconstitution of the dried, lyophilized sample. Fractions of about 2 mL (60 drops) were collected continuously using the BIO-RAD Model 2110 Fraction Collector up to a total eluate

volume of around 120 mL. Each fraction was then measured for Cu and UV absorption at 220 nm.

Hydrophilic Interaction Liquid Chromatography (HILIC)

A PolyLC Inc. polyhydroxyl A 5uM hydrophilic interaction liquid chromatography (HILIC) column was used in the BioEad BioLogic DuoFlow instrument and software. The buffers used were 65% acetonitrile in 20 mmol tetramethylammonium phosphate buffer or 60% acetonitrile in the 20 mM tetramethylammonium phosphate buffer and 10% acetonitrile in the tetramethylammonium phosphate buffer. The software created a gradient that started with the high acetonitrile content to the 10% acetonitrile. The sample injected into the column was a concentrated 3 kDa filtrate dissolved in 1 ml of a high percentage acetonitrile in a 20 mmol tetramethylammonium phosphate buffer. The fractions collected were about 1 mL, due to the 1 mL flow rate. To start 7 ml of the high acetonitrile concentration was used, and the gradient consisted of 18 ml. The end of the run used 5 ml of the 10% acetonitrile in the tetramethyl ammonium phosphate buffer. Each fraction was collected into Eppendorf tubes, and the fractions were evaporated before testing, since the acetonitrile used is highly volatile. Once the buffer was removed, each fraction was then reconstituted in about 1 ml of ddH2O to ensure that all fractions had the same volume.

Atomic Absorption Spectrometry

Cu concentrations were measured using the Perkin-Elmer AAnalyst 600. The apparatus was set up and operated according to the manufacturer's suggested protocol for Cu analysis.

UV Spectrophotometry

To measure the overall amount of material in a sample, UV light absorbance at 220 nm was measured. This was done using the Beckman Coulter DU 800 Spectrophotometer which was set up and operated according to the manufacturer's suggested protocol for sample analysis.

Calculating Sample Purity and Cu Recovery

Cu recovery and sample purity were the primary metrics for gauging the success of a procedure and the overall purification. Recovery is given as a percentage by dividing total Cu in a sample after the procedure(s) by total Cu before the procedure(s). Purity is given by the ratio of Cu concentration to UV light absorbance at 220 nm (Cu/A220).

Results

Determining Dialysis Tubing Cutoff Size

In order to determine which dialysis tubing would generate the best SCC-Cu recovery and most effective removal of contaminants from 3 kDa ultrafiltered samples of pig plasma, several trials of dialysis were performed. For the dialysis solution, either ddH2O or 5 mM potassium phosphate buffer at pH 7 was used. For the dialysis tubing, molecular weight cutoffs ranging from 0.1 - 0.5 kDa to 1 kDa were used. The Cu concentration, purity and recovery of each sample were measured for each condition as shown in Table 1. Dialysis solutions were refreshed at increasingly long intervals, specifically at 30min, 30min, 1hr, and overnight.

Dialysis Solution	Tubing Cutoff Size (kDa)	[Cu] (ppb)	Cu/A220	Cu Recovery (%)
ddH2O	1	-1	4	1
5 mM K PO4 (pH 7)	1	-1	-10	-3
ddH2O	0.5 - 1	22	24	28
5 mM K PO4 (pH 7)	0.1 - 0.5	24	25	74
ddH2O	0.1 - 0.5	29	27	86

Table 1 - Dialysis Tubing Trials: Several trials were performed to determine which dialysis cutoff filter and which dialysis solution generated the best Cu recovery and purity. Doubly-deionized water and 5 mM K Phosphate were investigated as potential dialysis solutions. Tubing cutoff sizes ranged from 0.1 - 0.5 kDa up to 1 kDa. Dialysis solutions were refreshed at increasingly long intervals over a couple of days.

The data showed that tubing with the smallest pores, 0.1 - 0.5 kDa, gave the best recovery compared to those with larger sized pores. Using water as the solvent also seemed to be better for Cu recovery than using buffer. The data also show that the smallest pore dialysis resulted in better purity of the SCC, which is indicated by the ratio of Cu/A220, A220 measuring the presence of molecules as a whole in the sample. Based on these results it was decided that 0.1 - 0.5 kDa cutoff tubing should be used for future purifications.

The volume of 3 kDa filtrate before and after dialysis was too large to be used effectively in further purification techniques such as size exclusion chromatography (SEC) and hydrophilic interaction liquid chromatography (HILIC). As such, samples needed to be concentrated in order to test such techniques and thus, the effects of lyophilization to concentrate the dialyzed filtrate was examined. For this we allowed the sample to lyophilize over several days to complete

dryness. Afterward, we dissolved the dried sample in ddH20. For every 37.5 mL of dialyzed 3 kDa filtrate, we used 6 mL of ddH2O or an amount proportional to this ratio if the filtrate volume was different. The data in Table 2 show that concentration by lyophilization resulted in a more than 5-fold increase in SCC purity and that Cu recovery remained high. Thus, we concluded that lyophilization to complete dryness should follow dialysis as a preparatory step for further purification.

	Cu/A220	Cu Recovery (%)
3K Filtrate	5	N/A
3K Dialyzed	11	74
3K Concentrated	60	64

Table 2 – Sample purity (Cu/A220) and Cu recovery was measured at each major stage of the most current purification process, namely after ultrafiltration, after dialysis, and after lyophilization-reconstitution.

Although concentrated samples possessed fairly high SCC purity, further purification was needed in order to obtain meaningful results when subsequently applying the sample to various structural analysis methods such as solution NMR and mass spectrometry. One of the first methods that was trialed was native polyacrylamide gel electrophoresis (native PAGE) using 10% acrylamide tube gels. Each gel received 500 uL of concentrated 3K filtrate containing about 200 ng of Cu and subjected to about 5 mA of electrical current. Following electrophoresis, the gel was sliced into 8 sections and measured for Cu. Figure 1 shows that much of the Cu remained within the first section of the gel indicating that the SCC-Cu migrated slowly. The data also show, however, that the concentration of Cu in the first section is less than 10 ppb (ng/mL). As each section was less than 1 mL, this indicated that less than 10 ng of Cu was retained after the procedure. Indeed, all native PAGE trials resulted in less than 5% Cu recovery. As such, native PAGE was eliminated as a potential purification step following ultrafiltration, dialysis, and lyophilization.
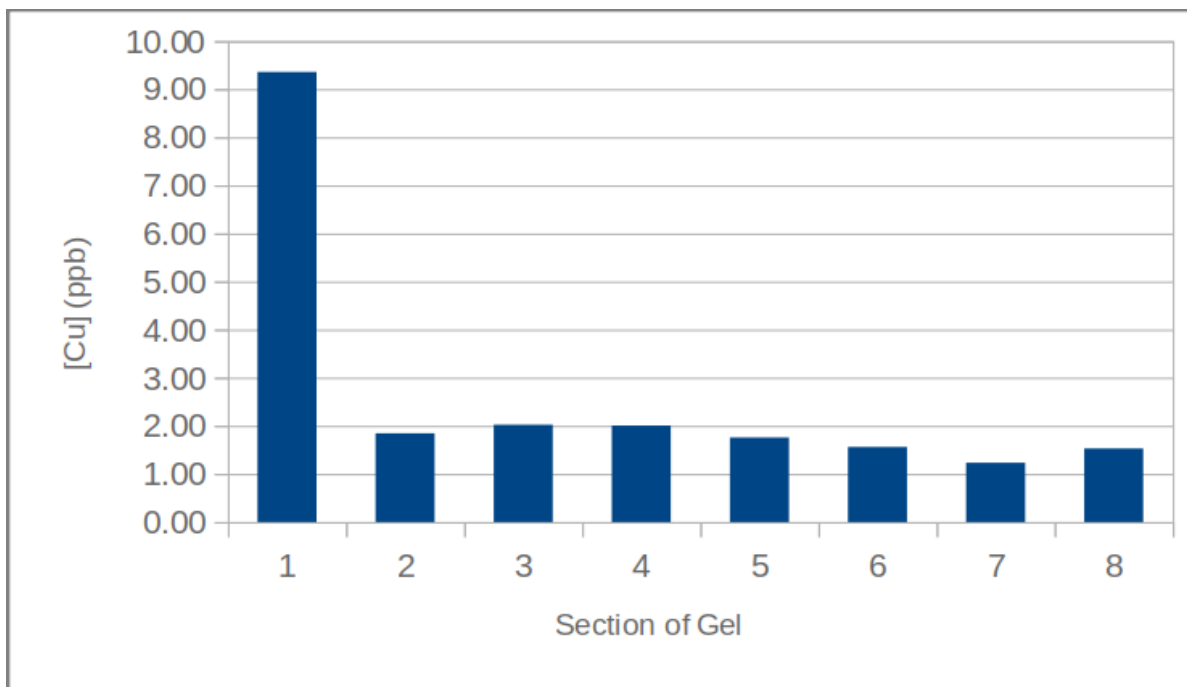
Figure 1 – Further separation using native PAGE: 500 uL of concentrated 3K filtrate containing ~200 ng of Cu was added to each 10% acrylamide gel. Cu peaks were mostly found in early sections of the gels. Cu recovery in peak sections was consistently less than 5%.

With native PAGE eliminated for the time being, SEC was another option to investigate as a potential further purification method. For this, a 100 mL Biogel P4 SEC column equilibrated with ddH2O was used. Any precipitates that formed during reconstitution of the dried sample to be added were removed using centrifugation. Concentrated samples added to the column were typically 1 mL in volume and contained about 1000 ng of Cu. Fractions of about 2 mL were collected up to a total eulate volume of about 120 mL and each fraction was measured for Cu and absorption at 220 nm as shown in Figure 2. The data show that in peak-Cu fractions, purity was very high given by a Cu/A220 ratio of around 330. Total Cu recovery across multiple Cu peaks was nearly 100%. From these results, it was determined that SEC was a better separation method than native PAGE and should be refined further.
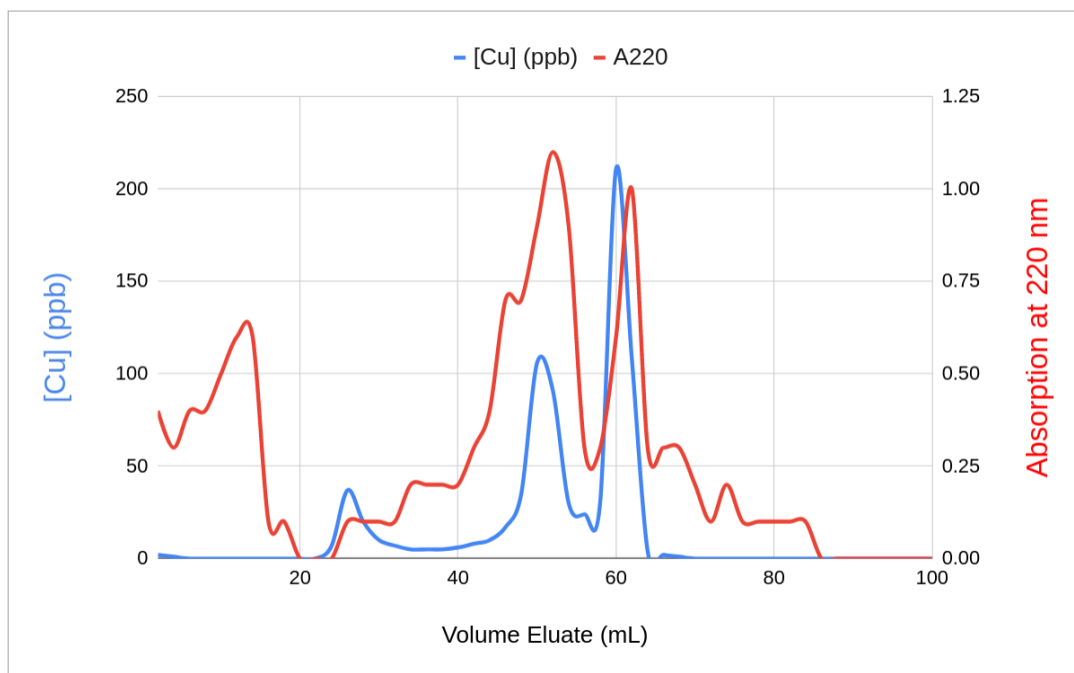
Figure 2 – Further purification using SEC: 1 mL containing ~1000 ng of Cu was added to a 100-mL Biogel P4 SEC column equilibrated with ddH2O. The highest and most pure peak was found at 60 mL of eluate containing more than 400 ng of Cu. The Cu/A220 ratio here was 330. Total Cu recovery across the three peaks was nearly 100%.

Hydrophilic interaction liquid chromatography (HILIC) was also used as a subsequent purification technique. The preliminary data showed poor Cu recovery, so the column was cleaned with 40 mM EDTA to remove any contaminants retained by the column. The first sample injected into the column was dissolved with the 60% acetonitrile in 20 mM tetramethylammonium phosphate buffer. The sample was 1 ml in volume, had a Cu content of 588 ug/L. The column had a flow rate of 1 ml/min for the entire run. For the first 7 ml the 60% acetonitrile in the tetramethylammonium buffer was used. For the following 18 ml a gradient from 60 to 10 % acetonitrile in the tetramethylammonium buffer was used. For the last 5 ml of the run, only the 10% acetonitrile in the tetramethylammonium buffer was used. The column had an 80% Cu recovery, which indicated that the EDTA helped to cleanse the column, and produced better results.

HILIC 60% ACN (2)
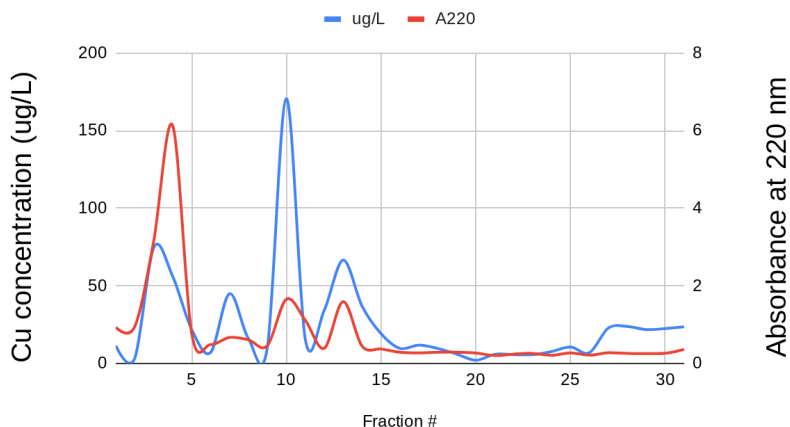
Figure 3 – Effect of running dialyzed and concentrated 3k filtrate on a HILIC column; 1 ml containing 588 ng of Cu was added to the HILIC column and run with a 60% ACN to 10% ACN gradient in 20 mM tetramethylammonium phosphate buffer pH 6 after being cleaned with 40 mM EDTA. This column had an 80% Cu recovery.

The next column was run under similar conditions, except that the dissolving solution, and high acetonitrile buffer was now 65% acetonitrile in a 20 mmol tetramethylammonium phosphate buffer pH 6 instead of 60%. The Cu recovery for this column was only 20%, indicating that the 60% acetonitrile might produce better recovery.

65% ACN
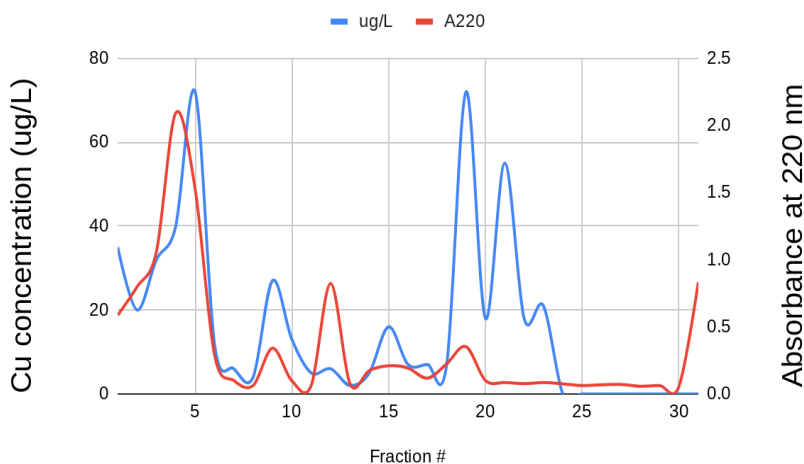
Figure 4 – Effect of using 65% acetonitrile in the column starting buffer; 1 ml containing 1950 ng of Cu was added to the HILIC column and run with a 65% ACN to 10% ACN gradient in 20 mM tetramethylammonium phosphate buffer pH 6 after being cleaned with 40 mM EDTA. This column had a 20% Cu recovery.

After collaborating with the producer of the column, it was suggested to cleanse the column again but this time using hexafluoro-2-propanol, to remove anything that the EDTA could not. The subsequent data shows the results after the column was cleaned. The starting buffer contained 60% acetonitrile, since it had produced a better Cu recovery as shown in Figure 1. On this column, 400 ng of Cu were injected, and had a 38% Cu recovery.

60% after hexafluoro-2-propanol clean
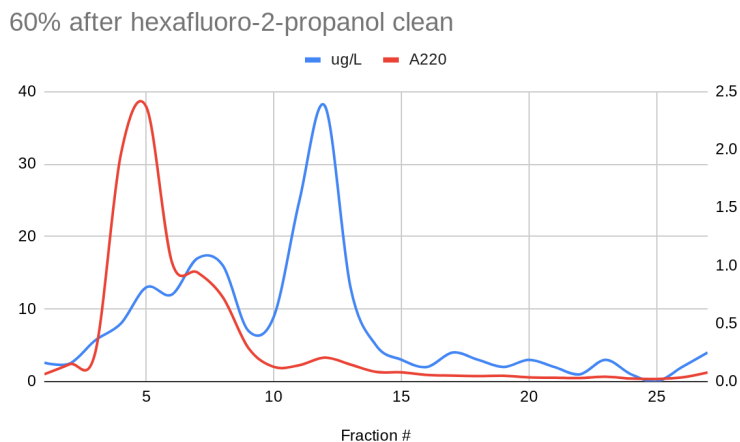
Figure 5 – Effects of cleaning the HILIC column with hexafluoro-2-propanol; a 1 ml sample containing 400 ng of Cu was injected and was run with a 60% ACN to 10 % ACN gradient in 20 mM tetramethylammonium phosphate buffer pH 6. This column had a 38% Cu recovery.

The collaborator also suggested using an overall 20 mmol tetraethylammonium phosphate buffer pH 6 for the high acetonitrile buffer. The next column was run using this new buffer. The sample used had 600 ng of Cu, and the column had a 43% Cu recovery.

60% ACN after hexafluoro-2-propanol clean Trial 2
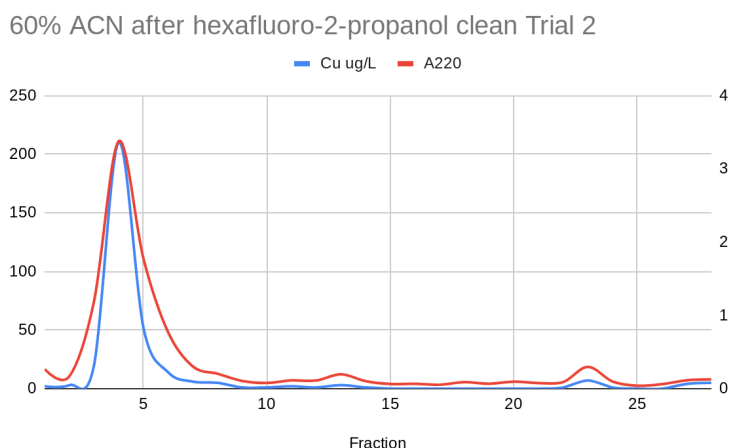
Figure 6 – Effects of using an overall concentration of 20 mmol tetraethylammonium phosphate pH 6 for the buffers; This HILIC column also used a 60% to 10% ACN gradient, but the 60% ACN had a final concentration of 20 mM tetramethylammonium phosphate buffer at pH 6 and the 10% ACN used a 20 mM tetramethylammonium phosphate buffer pH 6. 1 ml containing 600 ng of Cu was injected into the column, and this column had a 43% Cu recovery.

Discussion

One of the major obstacles in this effort to establish an effective purification method for SCC in pig plasma was maintaining high Cu recovery at each step. In previous purification trials, much Cu was lost during the lyophilization-reconstitution stage of the process. It was hypothesized that presence of salts in the 3 kDa ultrafiltrate was causing SCC-Cu to precipitate out of the solution. Thus, dialysis was included after 3 kDa ultrafiltration to remove salts. Doing so led to Cu recovery values consistently above 50% for overall purification.

To determine the most optimal parameters for dialysis, several trials were performed using varying dialysis solutions and filter tubing cutoff sizes. It was determined that using dialysis tubing with a cutoff size of 0.1 - 0.5 kDa yielded the highest Cu recovery and most effective removal of salts compared to other cutoff sizes. Additionally, ddH2O as the dialysis solution gave slightly better recovery and separation compared to 5mM potassium phosphate buffer pH 7 as seen in Table 1.

Presently, two methods of purification are being investigated, namely size exclusion chromatography (SEC) and hydrophilic interaction liquid chromatography (HILIC). Native polyacrylamide gel electrophoresis (native PAGE) using 10% acrylamide gels was being studied previously as well, but as shown in Figure 1, consistently poor Cu recovery (less than 5%) eliminated this as a potential separation procedure. It should be noted that during these studies, Cu was typically found in the early portion of the gel. This indicates that SCC-Cu migrated slowly and appears to have a low charge/mass ratio.

Comparatively, SEC has given more promising results with average Cu/A220 ratios of about 400 in peak-Cu fractions and Cu recovery of nearly 100% across all peak-Cu fractions. In many cases, 2 peaks were observed, one at around 40 mL of eluate and another at around 90 mL of eluate. In other cases, 3 peaks were observed, one at around 25 mL eluate, another at 50 mL eluate and a third at around 60 mL eluate. Inconsistency in recent results is being investigated and may be due to Cu accumulation within the column. Preliminary efforts to mitigate this issue include using 0.5 M EDTA to flush out lingering Cu from previous trials. Nevertheless, multiplicity in peaks suggests that SCC may be an aggregate molecule, with the most disaggregated forms eluting in later fractions. This is being further investigated by rerunning samples from a single peak-Cu fraction through the SEC column and seeing if multiple peaks form in the same manner.

Additionally, presence of salts appears to promote aggregation and thus, future studies will also include observing how the SEC elution profile changes when using phosphate buffered saline as the column's solvent.

Previous HILIC columns were giving poor Cu recovery, and it was suspected to be because the column needed to be passivated to clean any residues that could still be in the column. The column was passivated with 40 mM EDTA to remove any metals or minerals that could still be in the column. At first there was a significant increase in Cu recovery (Figure 3), but these results did not stay consistent (Figure 4). The acetonitrile concentration was increased to test if it would increase Cu recovery, but only 20 % was recovered in that trial (Figure 4). The column was then cleansed with hexafluoro-2-propanol in hopes that any contaminants left would be removed. The first column run after this clean was run with a 60% acetonitrile, and it did show an increase from the previous run with a Cu recovery to 38% (Figure 5). A second column was run under the same conditions (Figure 6), to see if the data was reproducible. This column did show an increase of Cu recovery, 43%, but the peaks shown in the graphs were not identical. Similar to data obtained from SEC trials, multiple distinct Cu peaks were observed in HILIC data, further suggesting that SCC may be an aggregate. More studies are being conducted to investigate why the data for every HILIC column tends to be different, however.

In any case, consistently high Cu recovery and good separation using 10K ultrafiltration followed by 3K ultrafiltration, then dialysis and lyophilization-reconstitution indicates that this is an effective purification method for isolating SCC from pig plasma. SEC and HILIC, though still being investigated, both seem to give better results than separation by native PAGE. The purest samples obtained from purification efforts will be analyzed for full structure determination using mass spectrometry (including electrospray), EPR, and solution NMR. Elucidating SCC structure can reveal more about the specific function of SCC which has numerous clinical applications including better diagnostic tools and therapies for Cu-overload diseases.

References

Linder, Maria C. "Copper Homeostasis in Mammals, with Emphasis on Secretion and Excretion. A Review." International Journal of Molecular Sciences 21.14 (2020): 4932. Web.

GEOLOGICAL SCIENCE

Examining The Concentric Pattern of The Box Springs pluton, Riverside, California

Author: Alex Valenzuela
Faculty Advisor: Dr. Vali Memeti
California State University, Fullerton
Department of Geological Sciences

Abstract

Plutons are an important part of the magma plumbing system that feed dangerous volcanic eruptions, yet the mechanisms of construction and evolution through time are poorly understood. This is also the case for the Cretaceous Box Springs Pluton in Riverside, CA. The geologic map pattern of tonalite, granite, and granodiorite shows a concentric unit arrangement with the youngest units in the core and the oldest units on the edges. The goal was to investigate the structure and composition of this pluton and test the hypotheses that the concentric structure represents a vertical transcrustal volcanic feeder and that the magmas are petrologically related.

An area of 1.3 km^2 was mapped at 1:10,000 scale in the oldest marginal units on the west side of the Box Springs pluton northeast of UC Riverside. Three of the units were mapped: Kbgg, a porphyritic granodiorite and granite; Kbhg, a heterogenous porphyritic granodiorite and granite; and Kbhg$_1$, a diked and layered version of Kbhg. Samples of the three units and the fine-grained leucogranite dikes within Kbhg$_1$ were collected for petrography and XRF whole rock geochemistry.

Petrographic observations show that each unit has a similar relative abundance of quartz, K-feldspar, and plagioclase. Mafic mineral abundance varies considerably, mainly within Kbhg$_1$ where hornblende and biotite make up nearly 40% combined in localized schlieren layers, and only appear at up to 5% in the other units. Accessory minerals include muscovite, epidote, hematite, sphene, ilmenite, and allanite. The dikes within Kbhg$_1$ are very felsic with only 5% mafics, and contain muscovite and garnet as accessory minerals. Since mineralogy is similar and modal abundances vary, the units might be petrologically related through fractional crystallization. This will be further tested with XRF geochemistry. All units have a consistent structure with magmatic foliations striking NW-SE and dipping toward the core of the pluton to the NE. The foliations dip steeper (50-65 degrees) near the center of the mapped area and are shallower (25-40 degrees) near the edges of the mapped area, resembling the structure of a lopolith. The dikes within Kbhg$_1$ strike in the same orientation to the units but dip in a SW direction. Contacts between Kbhg and Kbhg$_1$ are gradational while the contact between Kbgg and Kbhg is sharp. These relationships suggest that a transcrustal volcanic feeder cannot be excluded, but that the geometry of the magma plumbing in this part of the pluton is not a simple vertical cylindrical shape.

An Examination of The Structural and Petrologic Evolution of The Box Springs Plutonic Complex, Riverside, CA

Author: Andrew Culleton
Faculty Advisor: Dr. Vali Memeti
California State University, Fullerton
Department of Geological Sciences

Abstract

Although ancient magma plumbing systems (plutons) are commonly exposed at the surface, such as the Cretaceous Box Springs pluton of the Peninsular Ranges batholith, Riverside, CA, their emplacement and evolution through time are not well understood. Previous mapping of the Box Springs pluton shows concentric rock units of biotite granodiorites and biotite tonalites decreasing in age towards the core of the complex. The elliptical shaped kbt unit at Blue Mountain in the northeast of the pluton is hypothesized to represent a vertical magma feeder zone. This was tested by mapping the orientation and dip of the contact with adjacent units. A steep contact fits this interpretation, while a subhorizontal contact indicates a sill. Petrography and XRF analyses test whether the Box Springs pluton units are the product of one large fractionating magma body that traveled through the feeder zone.

Ten days were spent in the field mapping and collecting samples for petrography and whole-rock geochemistry of the 2 km2 NE section of the Box Springs pluton at Blue Mountain. Mapping revealed three compositionally diverse units decreasing in foliation intensity and sub-solidus deformation moving inwards towards the summit of Blue Mountain. The oldest unit located to the northeast of the summit, kbft, is a biotite tonalite with a strong magmatic foliation and parallel enclave alignments dipping southwest 71-82°. The kbgr is the second youngest unit located to the southwest of the summit. Here magmatic foliations strike northwest to southeast parallel to the contact with the kbft and dip 80-85° to the southwest. The youngest unit, kbt, is a biotite tonalite showing a weak subvertical magmatic foliation that strikes northwest to southeast and dips 60-89° to the southwest. Petrography shows higher biotite abundances and sub-solidus deformation in the kbft and kbgr units compared to the kbt. Biotite-rich schlieren layers line the elliptical kbt contacts on the north and east sides of the unit, which dip subvertically and young inwards towards the summit as indicated by cross-cutting relationships of individual layers. Schlieren along the contact between kbgr and kbft young towards kbgr.

Steeply dipping contacts with schlieren around the core of the pluton confirm that kbt at Blue Mountain likely represents a vertical feeder for the complex. Schlieren along the contacts indicate flow of magma against a fully or partially solidified wall of older units allowing mafic minerals to accumulate. XRF data is being analyzed to determine the petrologic relationships between the three units.

Lidar analysis of the high-slip section of the Hector Mine Earthquake Surface Rupture

Author: Brandon Cugini
Advisor: Dr. Sinan Akciz
California State University, Fullerton
Department of Geological Sciences

Abstract:

The Hector Mine earthquake (Magnitude 7.1) that occurred on 16 October 1999 within the eastern California shear zone, parallel to the San Andreas Fault, produced 48 km of surface rupture. This was the first large earthquake to have occurred along the fault strand in the Holocene Epoch. This is also the only fault with a surface rupture that has repeat (2000 and 2012) Light Detection and Ranging (lidar) that not only provide an opportunity to see if offset measurements using lidar data are reproducible, but also help document how much modification offset geomorphological features experience at a decadal timescale. In this study, I made 43 horizontal offset measurements using DEMs generated from lidar data collected along the maximum slip zone section Hector Mine earthquake surface rupture in 2012. The offset measurement values in the maximum slip zone area ranged from .5 +/- 0.1 m (CT-181) to 7.5 +/- 0.4 m (CT-162) with evidence of rapid changes in short distances. Comparison of measurements made using the 2000-lidar data and the 2012-lidar data shows that measuremend discrepancies (11 out of 43) are exclusively due to the subtlety of the geomorphological feature that was identified as displaced and the subjective determination of matching features between the geologists. My analyses also suggest that that no significant erosion or deposition has occurred in the maximum slip zone of the Hector Mine earthquake surface rupture in the 12 years the two lidar datasets were collected.

Introduction:

Earthquakes pose a severe hazard to people living in areas that are prone to have high risk of seismic ruptures and activity. Therefore, it is crucial that the kinematics and dynamics of earthquakes are better understood. Seismic and satellite instruments are commonly used to detect and characterize current seismic activity and provide abundant data for analyses, but

they have not been around long enough to record data that span multiple earthquake cycles. Paleoseismology aims to fill this gap by providing information about the age and magnitudes of past surface rupturing earthquakes. Measurements of surface displacements are one of the most widely used pieces of data collected by paleoseismologists to determine the magnitude of past earthquakes (e.g., Wallace, 1968; Sieh 1978; Hubert-Ferrari et al., 2002; Zielke et al., 2020).

Offset ephemeral or persistent stream channels along with abandoned terrace risers are the most commonly used geomorphic markers that are surveyed manually for making displacement measurements (e.g., Gold et al., 2009; Lienkaemper, 2001; Salisbury et al., 2012; Sieh 1978; Wallace, 1969; Woeard et al., 2002). Geologists typically identify the offset feature in the field, trace the offset marker on either side of the fault, and measure the distance between the projections of the offset markers on to the fault trace. Multiple measurements are made using different markers that are associated with the same offset feature (e.g., channel thalweg and two other margins). In the last two decades, georeferenced high-resolution digital topographic data, derived from airborne light detection and ranging (lidar) instruments, has been utilized in documenting and measuring offsets of past earthquakes (e.g., Arrowsmith & Zielke, 2009; Chen et al., 2015; Frankel et al., 2007; Hudnut et al., 2002; Zielke et al., 2010). LadiCaoz (Zielke & Arrowsmith, 2012) and 3D_Fault_Offsets (Stewart et al., 2017) are the two main software packages that are utilized to make these offset measurements using lidar data. These offset measurements are then used to estimate rupture lengths and magnitudes of the past earthquakes. However, in nearly all the cases, these geomorphological offsets are at least a century, sometimes several centuries old (e.g., >100 years for the northern San Andreas Fault and >150 years for the south-central San Andreas Fault). Do lidar-derived measurement datasets have adequate spatial density to represent the slip distribution along active faults? How well do geomorphological offset features get preserved to represent old displacements? While Chen et al. (2015) showed that measurement density offset geomorphic features identified on lidar datasets can, if conditions are right, can at least match and even be higher than field collected data. However, to this date, there has not been a study investigating possible uncertainties associated with geomorphic modifications these offset markers might have experienced since their formation and displacement. Some outstanding that remain to be answered are the following: Is there any evidence for erosional or depositional modification of these geomorphic features that are commonly used in offset measurements? How well do these offset measurements made

on geomorphic features that have been potentially modified for at least a century represent displacements that occurred during previous earthquakes? Do the uncertainties reported with offset measurements large enough that they include uncertainties associated with such post-earthquake modification, or are they consistently less and require an alternative method of measurement? Since most of the active faults within the San Andreas Fault system and elsewhere go through Quaternary-aged sediments, the effect of erosion and deposition to offset measurements made along strike-slip faults is of utmost importance to evaluate the relevance of the offset measurements that are now made routinely using lidar data.

The Hector Mine rupture (Figure 1) provides a unique opportunity to evaluate effects of erosion and deposition on surface displacement measurements along strike slip faults in arid settings because it is the only historic fault rupture where slip data has been collected on the same features using different data sets and over a period of twelve years. Field measurements were made immediately after the earthquake, and lidar was initially collected six months later (Hudnut et al., 2002). Lidar-based offset measurements were then made using this dataset and later published (Chen et al., 2015). In May 2012 a repeat lidar dataset was collected (NCALM, 2012). New field measurements were then made in 2012-14 (Akciz personal communication). In the context of these different datasets, several questions motivated this thesis: How well do different types of measurements made by different researchers at different times using different datasets agree? What data or technique issues, or natural processes, are be responsible for any differences? Is there any evidence of significant modification of geomorphological offset markers in the arid setting of the Hector Mine earthquake surface rupture in 12 years (the time in between the collection of two lidar datasets)? I examine these issues by comparing the multiple datasets of field and lidar-based offset measurements.

Geologic Background

In 1999, the 7.1 magnitude Hector Mine earthquake caused a north trending surface rupture of 48 km along the eastern California shear zone (Figure 1; Treiman, 2002). The Lavic Lake fault at the northern portion of the fault and the Bullion Fault to the south make up the two strands of the fault associated with the Hector Mine earthquake (Treiman et al., 2002). These two strands were mapped previously but showed no evidence of activity in the Holocene Epoch (Treiman et al., 2002). Therefore, the displacement observed along the fault strand are thought to have

occurred during the 1999 seismic event (Treiman et al., 2002). Based on field measurements, average horizontal displacement along the entire rupture was determined to be 1.5 meters, with the maximum being 5.2 meters (Treiman et al., 2002).
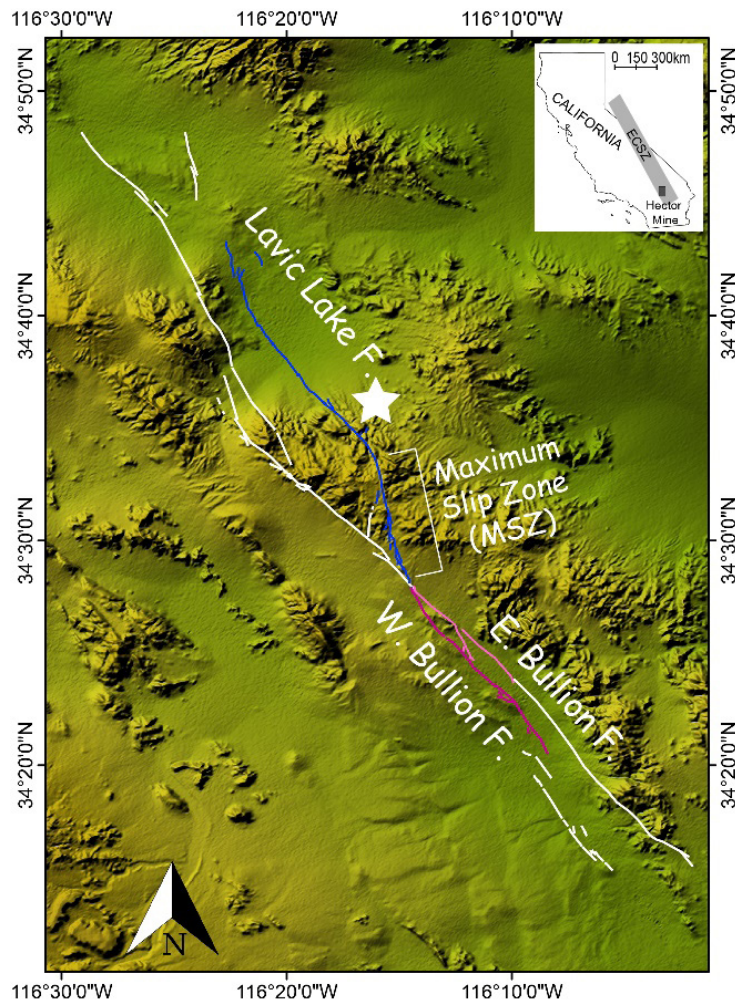
Figure 1: Overview of the surface rupture (colored lines showing the Lavic Lake, East and West Bullion faults) caused by the Hector Mine Earthquake shown on a shaded relief map derived from Shuttle Radar Topography Mission data (modified after Chen et al., 2015). White star shows the epicenter location. White lines represent the adjacent fault zone system that did not rupture during this earthquake. The extent of the study area, the Maximum Slip Zone (MSZ), is also shown. . The inset map indicates the context of the Hector Mine rupture within the northwest-striking dextral faults of the eastern California shear zone (ECSZ, gray band on inset map of California)

In April 2000, an airborne laser scanning survey was performed, creating the first high resolution lidar coverage of topography along the entire extent of a major earthquake surface rupture (Hudnut et al., 2002). Chen et al. (2015) used this ALS dataset along with LaDiCaoz, a Matlab-based cross-correlation tool for measuring offset topographic features from digital elevation models (Zielke and Arrowsmith, 2012), to measure over 250 individual offsets along the surface rupture. The analyses of lidar data from 2000 indicated that the maximum slip to be 6.6 meters, with a displacement of 1.72 meters being the average (Chen et al., 2015). This lidar analyses study also showed that while the lidar data can potentially increase the spatial density of offset measurements, there was more uncertainty in the measurements in the lidar sets than in the field work (Chen et al., 2015). Both of the previous studies are generally in good agreement regarding the amount of maximum horizontal offset (5-6 m dextral horizontal) and the location of a maximum slip zone (MSZ, Figure 1) within the Bullion Mountains section of the surface rupture where dextral horizontal offsets exceed four meters (e.g. Fialko et al. 2001; Peltzer et al., 2001; Treiman et al., 2002; Ayoub et al., 2009; Chen et al., 2015).

Methods

The lidar data analyzed in this study was collected by an airborne laser scanner system. As the aircraft flies over the fault, the scanners emit laser impulses that travel to the surface, are reflected back to the scanner, and are recorded. The onboard equipment measures how long the light was traveling to the ground and back and interprets the topography of the ground using data such as the laser travel times, the speed and location of the aircraft and the precise location and orientation information of the scanner unit with respect to the ground. This process is similar to SONAR and RADAR systems in concept. These point data is then converted to a digital elevation model (DEM) through additional processing after the completion of the data collection phase. All the DEMs used in the analyses described below were downloaded from opentography.org (last access January 2021).
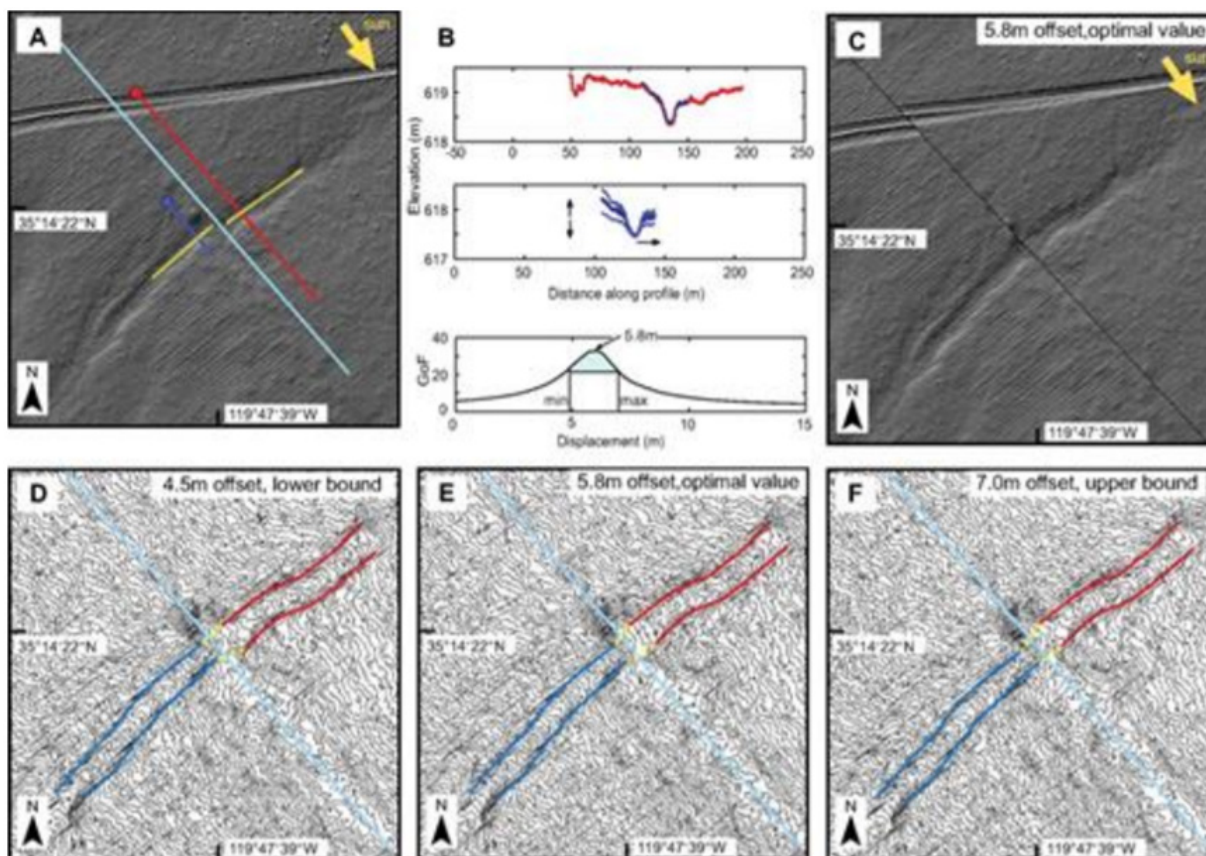
Figure 2. Examples of different stages of the LadiCaoz analyses shown on a well-developed channel in the Carrizo Plain (from Zielke and Arrowsmith, 2012). (A) Hill shade plot of the study area with NE illumination. Fault trace (cyan line) and topographic profile lines (red and blue), and channel trends (in yellow) are also shown. (B) A graphic representation of how well the upstream and downstream topographic profiles match and the slip magnitude that will maintain that goodness-of-fit is also shown as a probability density function. (C) Hill shade plot of the study area with NW illumination. Identification of subtle geomorphic features such as the fault trace and the small stream channel is sometimes difficult and the ability to look at multiple base maps help gain sufficient understanding of each site prior to identifying the location of faults and offset geomorphological features. Quick generation of those different illumination angles is a feature of the provided MATLAB tools. (D), (E), and (F) Back-slipped topography of the offset features using different slip amounts to determine the minimum-, optimal- and maximum-horizontal slip amounts.

The DEMs are assessed and interpreted using LaDiCaoz (lateral displacement calculator) software, which is a MATLAB cross-correlation code with a graphic user interface (GUI) (Zielke and Arrowsmith, 2012). Offsets are measured using the MATLAB GUI after certain basic parameters, including the fault location and trends of the thalwegs both downstream and upstream of the fault are entered manually. Additional steps were taken in order to complete the displacement measurements. Once the channel cross-sections were drawn, LadiCaoz automatically reconstructed the displaced channel profiles to determine the optimal displacement value and provide a goodness of fit value for how well the profiles matched as they were reconstructed. In the following steps, the DEM data was back-slipped to visually determine the minimum and maximum possible offset magnitudes for each offset. (Figure 2)

Results:

For this study, a total of 43 horizontal displacement measurements were made along approximately 4.5 km section of the fault (Figures 1 and 3). Locations of the offset features were obtained either from the Chen et al. (2015) investigation or from the field investigation notes collected in 2012-2014 (Akciz, personal communication). No new offset markers were identified using the 2012 lidar data. In comparison to the 15 measurements made by the field teams immediately after the earthquake (Treiman et al., 2002), this dataset contains the greatest number of offset measurements made to date along this section of the Hector Mine earthquake surface rupture. My study was able to reproduce the results of Chen et al. (2015) and agree with one of their conclusions which stated that measurements made on lidar-data alone can be as spatially dense or better than field data alone if geomorphological offset markers offset by the fault rupture are also abundant. In other words, the controlling factor is not the quality of the lidar data, but the existence of a linear geomorphological feature, such as a stream channel, across the fault prior to the rupture.

The offset measurement values in the maximum slip zone area ranged from .5 +/- 0.1 m (CT-181) to 7.5 +/- 0.4 m (CT-162) with evidence of rapid changes in short distances as evidenced along the surface ruptures of other earthquakes (e.g. McGill and Rubin, 1999; Rockwell et al., 2002; Treiman et al., 2002). The spatial distribution of the horizontal offset measurements is denser in the northern portion (the first 2 km) of the fault then the southern 2.5 km of the studied section

of the fault (Figure 3). The northern section has a total of 26 measurements, while the southern section has a total of 17. Offset measurements made using only the 2012 lidar data have no major (>300 m) spatial gaps in the fault segment studied.

The distribution of the horizontal slip within the MSZ is asymmetric with a cluster of >6 m slip within northwesternmost 750 meter of the study area (Figure 3). The northern portion (0 – 2 km) has measurements from as low as 2m, and as high at 7.5 +/- 0.4 m (CT-162). The average of these horizontal slip values being 4.1 +/- 0.61 m. While in the south (2 – 4.5 km), the measurements cluster around 2 m, with the highest being 6 +/- 1m, and the average of these values is 2.9 +/- 0.54 m of displacement. The average horizontal displacement along this 4.5 km high slip segment is calculated to be 3.65 +/- 0.58 m.
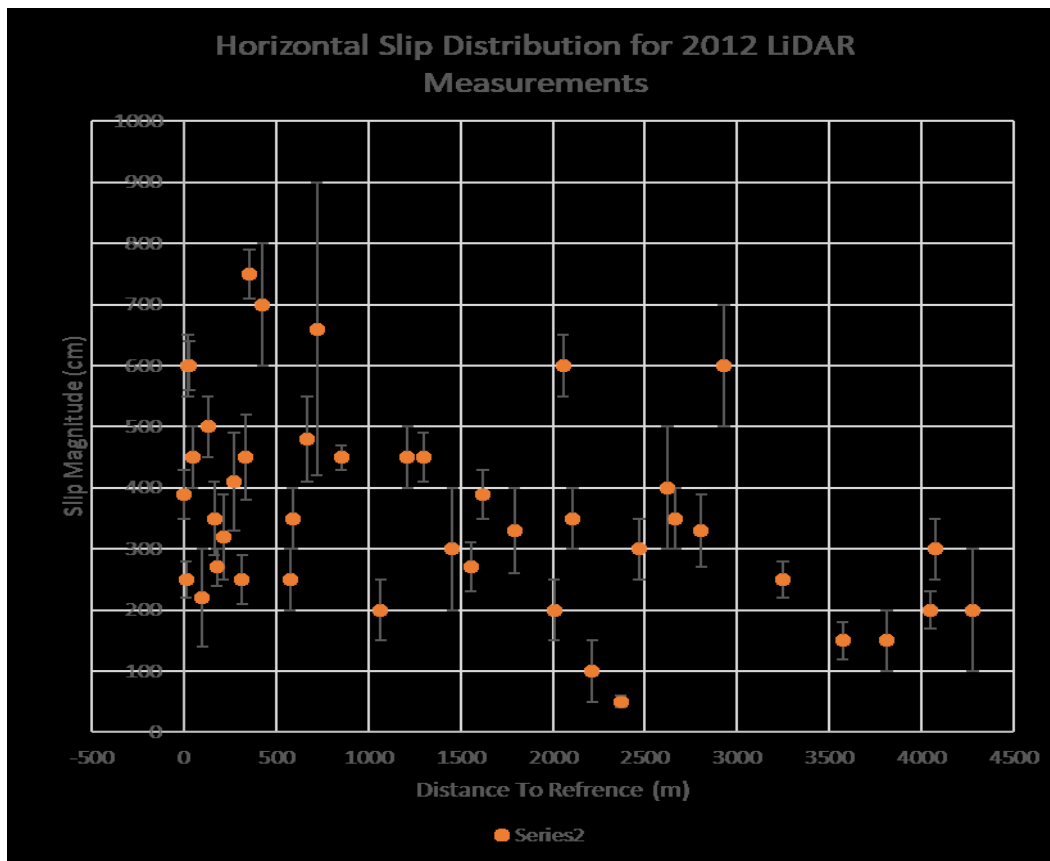
Figure 3. Distribution of horizontal offsets and measurement uncertainties plotted against distance along the surface rupture within the maximum slip zone.

Discussion

Comparison of measurements made using the 2000-lidar data (Chen et al., 2015) and the 2012-lidar data (this study) shows that 32 of the 43 measurements are similar to each other within uncertainties (Fig. 4). The similarity of the magnitudes of the measurements made by two different geologists using two different topographic datasets suggest that majority of the geomorphological features within the study area were well-preserved, easily identified, and had relatively simple geometries that resulted in offset measurements of similar magnitude.

The discrepancies in all other eleven locations are due to the subtlety of the geomorphological feature that was identified as displaced which resulted in different back-slipping solutions and the alternative interpretations of the projection lines by different geologists. In both datasets, these were all labeled as low-quality measurements because of these limitations. Two of the 40 measurement locations (sites CT169 and CT170) were investigated further as they are the only channels that are cut into unconsolidated alluvium deposits that were susceptible to erosion or deposition. In comparing the locations of the channel margins used for the offset measurements I conclude that no quantifiable morphological modifications occurred at these locations between 2000 and 2012. In such an arid-climate setting, this conclusion is not unexpected, however, to this date, no such repeat measurements of horizontal slip have been made to support this hypothesis. 12 years are unquestionably too short of a time period to determine how much do offset measurements made using lidar data on faults that last ruptured >100 years ago get modified.
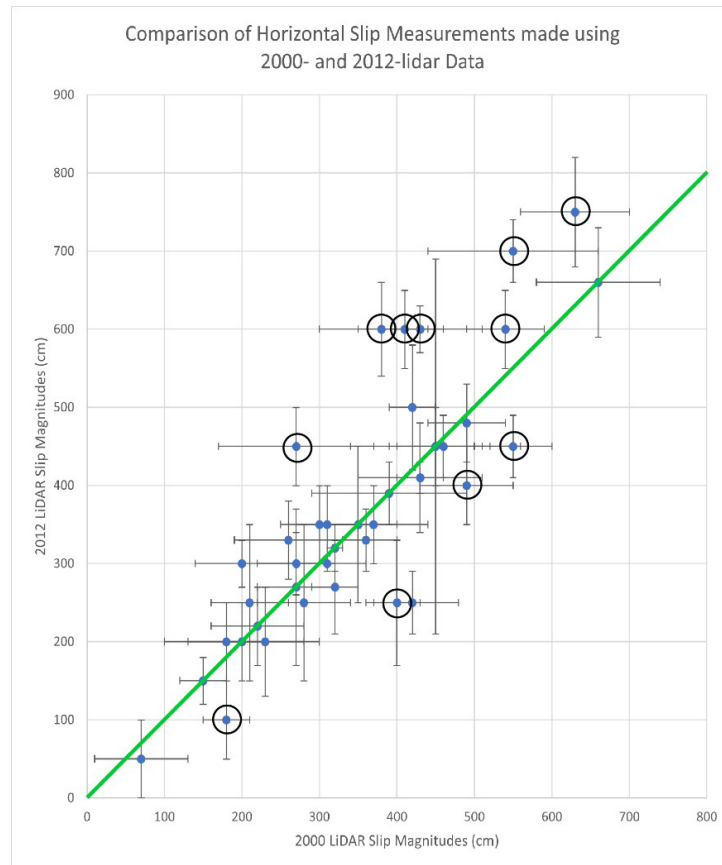
Figure 4. Comparison plot of the horizontal offset measurements made using the 2000- and 2012-lidar data. The solid green line is the 1:1 ratio line that was used to identify the 11 locations (black circles) where magnitudes of offset measurements did not match within their uncertainties.

Conclusions

I made 43 horizontal offset measurements using DEMs generated from lidar data collected along the maximum slip zone section Hector Mine earthquake surface rupture in 2012. The offset measurement values in the maximum slip zone area ranged from .5 +/- 0.1 m (CT-181) to 7.5 +/- 0.4 m (CT-162) with evidence of rapid changes in short distances. Comparison of measurements made using the 2000-lidar data (Chen et al., 2015) and the 2012-lidar data shows that 32 of the 43 measurements are similar to each other within uncertainties. The discrepancies in all other eleven locations are due to the subtlety of the geomorphological feature that was identified as displaced. My analyses also suggest that that no significant erosion or deposition has occurred in the maximum slip zone of the Hector Mine earthquake surface rupture in the 12 years the two lidar datasets were collected.

References:

Arrowsmith, J. R., & Zielke, O. (2009). Tectonic Geomorphology of the San Andreas Fault Zone from High Resolution Topography: An example from the Cholame Segment. Geomorphology, 113(1-2), 70– 81.

Chen, T., Akciz, S.O., Hudnut, K.W., Zhang, D.Z., Stock, J.M. (2015) . Fault-Slip Distribution of the 1999 M 7.1 Hector Mine Earthquake, California, Estimated from Post Earthquake Airborne LiDAR Data. Bulletin of the Seismological Society of America. 105, no. 2A, p. 776-790.

Frankel, K. L., Dolan, J. F., Finkel, R. C., Owen, L. A., & Hoeft, J. S. (2007). Spatial Variations in Slip Rate Along the Death Valley-Fish Lake Valley Fault System Determined from LiDAR Topographic Data and Cosmogenic Be-10 Geochronology. Geophysical Research Letters, 34, L18303.

Gold, R. D., Cowgill, E., Arrowsmith, J. R., Gosse, J., Chen, X., & Wang, X.-F. (2009). Riser Diachroneity, Lateral Erosion, and Uncertainty in Rates of Strike-Slip Faulting: A Case Study from Tuzidun Along the Altyn Tagh Fault, NW China. Journal of Geophysical Research, 114, B04401.

Hubert-Ferrari, A., Armijo, R., King, G., Meyer, B., & Barka, A. (2002). Morphology, Displacement, and Slip Rates Along the North Anatolian Fault, Turkey. Journal of Geophysical Research, 107(B10), 2235. https://doi.org/10.1029/2001JB000393

Hudnut, K. W., Borsa, A., Glennie, C., and Minster, J.-B.,(2002). High-Resolution Topography Along Surface Rupture of the 16 October 1999 Hector Mine, California, Earthquake (Mw 7.1) From Airborne Laser Swath Mapping: Bulletin of the Seismological Society of America, v. 92, no. 4, p. 1570-1576.

Lienkaemper, J. J. (2001). 1857 Slip on the San Andreas Fault southeast of Cholame, California. Bulletin of the Seismological Society of America, 91(6), 1659– 1672.

Rockwell, T. K., S. Lindvall, T. Dawson, R. Langridge, W. Lettis, and Y. Klinger (2002). Lateral offsets on surveyed cultural features resulting from the 1999 Izmit and Duzce earthquakes, Turkey, Bulletin of the Seismological Society of America, 92, no. 1, 79–94.

Salisbury, J. B., Rockwell, T. K., Middleton, T. J., & Hudnut, K. W. (2012). LiDAR and Field Observations of Slip Distribution for the Most Recent Surface Ruptures Along the Central San Jacinto Fault. Bulletin of the Seismological Society of America, 102(2), 598– 619.

Sieh, K.E. (1978). Slip Along the San Andreas Fault Associated with the Great 1857 Earthquake. Bulletin of the Seismological Society of America. 68, no. 5. 1421-1448.

Sieh, K.E.and Jahns, R.H. (1984). Holocene Activity of the San Andreas Fault at Wallace Creek, California, Geological Society of America Bulletin, 95, 883-896.

Treiman, J.A., Kendrick, K.J., Bryant, W.A., Rockwell, T.K., McGill, S.F. (2002). Primary Surface Rupture Associated with the M 7.1 16 October 1999 Hector Mine Earthquake, San Bernardino County, California. Bulletin of the Seismological Society of America. 92, no. 4. 1171- 1191.

Van der Woerd, J., Tapponnier, P., Ryerson, F. J., Meriaux, A. S., Meyer, B., Gaudemer, Y., Zhiqin, X. (2002). Uniform postglacial slip-rate along the central 600 km of the Kunlun Fault (Tibet), from 26Al, 10Be, and 14C dating of riser offsets, and climatic origin of the regional morphology. Geophysical Journal International, 148(3), 356– 388.California, Geological Society of America Bulletin. 95,. 883-896.

Wallace, R.E. (1968) Notes on Stream Channels Offset by the San Andreas Fault, Southern Coast Ranges, California, United States Geological Survey.

Zielke, O., Arrowsmith, J.R. (2012). LaDiCaoz and LiDARimager -MATLAB GUIx for LiDAR Data Handling and Lateral Displacement Measurement, Geosphere. 8, 206-221.

Testing The Efficacy Of 3D-Printed Geologic Block Models As Tools For Fostering Spatial Visualization Abilities

Authors: Joseph Gutierrez[1], Sinan Akçiz[1], Natalie Bursztyn[2], and Kevin Nichols[3]
[1]Department of Geological Sciences - California State University, Fullerton
[2]Physical Sciences Division - Quest University, Canada
[3]Department of Mathematics - California State University, Fullerton

Abstract

Spatial visualization is crucial to success in the geosciences (Alles and Riggs 2011; Gobert 2005; Titus and Horsman 2009). Students with the spatial abilities necessary to succeed in early geoscience coursework are more likely to pursue the geosciences further. Those that are lacking in their spatial abilities are more likely to struggle, resulting in a significant proportion of students who disregard or abandon the geosciences. Previous studies have demonstrated that spatial abilities are malleable, and can be trained efficiently and inexpensively, yet widespread spatial training is uncommon (National Research Council 2006; Gold et al., 2018; Hespanha et al., 2009).

This research involved testing the use of 3D-printed geologic block models as tools for fostering spatial abilities in conjunction with an introductory-level geologic structures exercise in physical geology lab courses at California State University, Fullerton. In particular, we sought to answer the following two research questions - (i) Does teaching geologic structures with 3D-printed block models impact students' spatial abilities? (ii) Do student demographics such as gender, ethnicity, and academic background affect the impact of block models? It was hypothesized that the block models used to teach geologic structures would have some positive impact on students' spatial visualization abilities. It was also hypothesized that historically underrepresented and underserved students would receive the most significant benefits from the models, putting them on a more level playing field with their peers who have had more opportunity to develop their spatial abilities. This hypothesis was based on previously conducted research (Levine et al., 2005), which found that underrepresented students tend to have poorer spatial visualization abilities.

Three spatial visualization abilities were targeted during this study – spatial rotation, spatial manipulation, and visual penetrative ability. Though there are many spatial abilities, these three were deemed to be especially important to the geosciences (Titus and Horsman 2009). Spatial rotation – the ability to mentally rotate or reorient objects – is required to understand how

rocks are reoriented over time, as well as higher-level concepts such as pebble imbrications and the preferential reorientation of minerals in magma chambers (Shepard and Metzler 1970). Spatial manipulation – the ability to mentally reconfigure or distort an object – is necessary to understanding deformational processes such as folding and shear (Ekstrom et al., 1970). Lastly, visual penetration – the ability to mentally pierce through an object and envision its internal form – is frequently used when trying to interpret three-dimensional subsurface information from two-dimensional data presented on the Earth's surface (Kali and Orion 1996).

Pre- and post-lab spatial visualization surveys and a demographic background questionnaire were used to collect data from 593 introductory geology laboratory students who completed a geologic structures lab as part of their coursework. Data was processed in R statistical software, and missing data were imputed using the MissForest package. One-way MANOVA were used to determine the main effect of block model access on students' spatial visualization survey score differentials. Two-way MANOVA were used to determine the interaction effect of student demographics on block model impact.

The data collected indicate that access to the 3D-printed block models – as they were incorporated during this study – hindered the development of spatial rotation, while having insignificant effects on the development of spatial manipulation or visual penetrative abilities. This lack of efficacy is attributed to the means by which the blocks were incorporated, as well as the duration of the intervention. Additional manipulative and penetrative operations that incorporated the blocks were not required, and rotational operations were discouraged as students realized they could use the blocks to pattern-match when solving rotation-based problems. Furthermore, the students who participated in this study had approximately three hours to practice their spatial visualization abilities – previous studies have shown that students need longer-term, repeated practice if they are to develop their abilities (Duesbury and O'Neill, 1996; Ormand et al., 2014; Titus and Horsman, 2009).

These data also indicate that the block models had equitable results across student demographics. 3D-printed block models do not depend on any sort of cultural context or academic background, so who our students were did not change their experiences with the models. Though further research is required to determine the appropriate application of the block models as tools for fostering spatial visualization abilities, we take pride in knowing that our resources are equitable to students of all cultural and academic backgrounds.

References Cited

Alles, M., and Riggs, E. M., 2011, Developing a process model for visual penetrative ability, *in* Feig. A.D., and Stokes, A., eds., Qualitative Inquiry in Geoscience Education Research, Geological Society of America Special Paper 474, p. 63–80.

Duesbury, R.T., and O'Neil, H.F. Jr., 1996, Effect of type of practice in a computer-aided design environment in visualizing three-dimensional objects from two-dimensional orthographic projections: Journal of Applied Psychology, v. 81, no. 3, p. 249–260.

Ekstrom, R.B., French, J.W., Harman, H.H., and Derman, D., 1976, Manual from Kit of Factor-Referenced Cognitive Tests, Princeton, N.J., Educational Testing Service.

Gobert, J.D., 2000, A typology of causal models for plate tectonics: Inferential power and barriers to understanding: International Journal of Science Education, v. 53, no. 4, p. 444–455.

Hespanha, S.R., Goodchild, F., and Janelle, D.G., 2009, Spatial Thinking and Technologies in the Undergraduate Social Science Classroom: Journal of Geography in Higher Education, v. 33, no. 1, p. 17–27.

Kali, Y., and Orion, N., 1996, Spatial Abilities of High-School Students in the Perception of Geologic Stuctures: Journal of Research in Science Teaching, v. 33, no. 4, p. 369–391.

Levine, S.C., Vasilyeva, M., Lourenco, S.F., Newcombe, N.S., and Huttenlocher, J., 2005, Socioeconomic Status Modifies the Sex Difference in Spatial Skill: Psychological Science, v. 16, no. 11, p. 841–845.

National Research Council, 2006, Learning to think spatially—GIS as a support system in the K–12 curriculum. Washington DC: National Academies Press, 313 p.

Ormand, C.J., Manduca, C., Shipley, T.F., Tikoff, B., Harwood, C.L., Atit, K., and Boone, A.P., 2014, Evaluating Geoscience Students' Spatial Thinking Skills in a Multi-Institutional Classroom Study: Journal of Geoscience Education, v. 62, p. 146–154.

Shepard, R.N., and Metzler, J., 1970, Mental Rotation of three-dimensional objects: Science, v. 171, p. 701–703.

Titus, S., and Horsman, E., 2009, Characterizing and Improving Spatial Visualization Skills: Journal of Geoscience Education, v. 57, no. 4, p. 242–254.

Geochemistry of basalts and trachyandesites in northern Owens Valley, Inyo County, California

Authors: Matthew Pilker[1], Jeffrey Knott[1], Fred Phillips[2], Jade Star Lackey[3] and Angela Jayko[4]
[1]California State University, Fullerton
[2]New Mexico Tech
[3] Pomona College
[4]U.S. Geological Survey

Abstract

As the Mendocino Triple Junction migrated north past the latitude of northern Owens Valley (OV) at approximately 5 Ma, the plate boundary progressed from a subduction zone to a transform fault. One hypothesis is that the composition of rising basalt magmas (e.g. Ce, Y, Zr, and Ba) altered as the plate boundary changed. A second hypothesis is that delamination of the lithospheric root of the Sierra Nevada batholith around 4 Ma also altered the composition of basalt magmas (i.e., higher K_2O).

Basalts in northern OV have $^{40}Ar/^{39}Ar$ ages of 3 Ma and 11 Ma, which precede and follow these tectonic events. Basalt flows in the White Mountains (WM) to the east are dated to 11.5 Ma. Relations between the OV and WM flows and magma are unclear. We make three hypotheses. If the Sierra delamination raised K_2O in magma, then the 3 Ma basalts will have higher K_2O content. If the Ce, Y, Zr, and Ba concentrations change, then the 3 Ma basalts reflect plate boundary and magma pathway changes. Also, if the OV and WM basalts are geochemically similar, then they may represent the same flow or magmatic source. To test these hypotheses, one 3 Ma and three 11 Ma basalts in northern OV were sampled and field observations were made about the geologic nature of the outcrops. Samples were analyzed by x-ray fluorescence.   Our findings do not support the hypothesis that the OV and WM basalts represent the same flow or magmatic source. The samples are geochemically dissimilar: 11 Ma MZP-OV-2 is a trachyandesite (59.50 wt% SiO_2; 3.51 wt% Na_2O; 3.35 wt% K_2O) and 11 Ma MZP-OV-3 is a mantle-sourced basalt (0.32 Zr/Ba; 3.44 Ce/Y; 1.64 wt% K_2O). MZP-OV-4 is an 11 Ma lithosphere-sourced basalt (0.19 Zr/Ba; 4.36 Ce/Y; 2.72 wt% K_2O) geochemically similar to the WM basalts; however, field observations suggest MZP-OV-4 is a small vent and not part of a larger flow. Our findings also do not support the hypothesis that the 3 Ma basalts in northern OV are mantle-sourced. The 3 Ma "basalt", MZP-OV-1, is a basaltic trachyandesite (51.42 wt% SiO_2; 2.43 wt% Na_2O; 3.68 wt% K_2O) showing lithosphere interaction (0.15 Zr/Ba; 2.35 Ce/Y). The 3 Ma basaltic trachyandesite and 11 Ma trachyandesite had similar K_2O concentrations, thus we were unable to discern the influence of Sierra delamination on magma composition in this area. These relations may instead be linked to crustal assimilation.

Mathematics

MANCALA: DIVIDE AND TIED

Authors: Salma Albakri and Chloe Rhee
Advisor: Alison S. Marzocchi, Ph.D.
California State University, Fullerton
Department of Mathematics

Abstract.

Traditional mancala has been played for millennia and requires a lot of problem-solving skills and strategies. We wanted to see how the game mancala would change if the twelve pits were reduced to six pits with only one gemstone instead of four gemstones in each pit, along with mirrored moves and no capturing allowed. First, we show that there are a total of sixteen different possible games. Then, we show that every game results in a tie of three gemstones each for both Players A and B. We concluded that the order of who plays first does not make a difference in the end. Our version of mancala did not require many strategies and critical thinking since the second player's move would mirror the first player's move.

1. Introduction

1.1. **History of Mancala.** Mancala is a board game that has been played for thousands of years, dating back to the sixth century in Africa [2]. The term mancala comes "from the Arabic word naqala", which translates to "to move" [3, pg. 1]. There are many different ways of playing this board game, some versions involve two, three, or even four players. Some mancala games have a different amount of rows, pits, and/or gemstones. Overall, the basics and goals of mancala are the same. In our paper, we will cover what the traditional mancala consists of along with the altered mancala game we created.

1.2. **Traditional Mancala.** The most popular version of mancala consists of twoplayers, one mancala board, forty-eight gemstones, twelve total pits, and two man-calas. The wooden game board is long and is placed between the two players vertically. It is made up of two rows with six pits each, and mancalas on the oppo-site ends of the board. Players A and B have the six pits straight in front of them on their right-hand side. In each of the twelve total pits, four gemstones are placed. Refer to Figure 1 to see what the traditional mancala board setup looks like.

Now that we have explained the equipment and setup of the traditional mancala, let us explain the objective and gameplay of it. The objective of mancala is to find out which of the two players can collect the most gemstones into their mancala. Then, we have the gameplay, which starts with Player A picking up the four gemstones from any one of the pits on their side. Player A will then place onegemstone at a time into their pits in a counterclockwise manner until the gemstones run out [3]. If Player A encounters their mancala, a gemstone is placed there as well. If Player A's last gemstone is placed in their mancala, then Player A receivesan extra turn.

On the other hand, if there are extra gemstones even after placing a gemstone in Player A's mancala, Player A can continue to place gemstones onto Player B's pits [3]. Yet, if Player A's extra gemstones encounter Player B's mancala, then that mancala is skipped over. If Player A's last gemstone is placed in an empty pit on

its own side, then Player A will take this gemstone along with Player B's gemstones across from the empty pit that it was placed in [3]. These gemstones get placed in Player A's mancala. This process is repeated for Player B. The players continue taking turns until the game is completed.

There are specific terms that can be used to describe a few of the processes that were explained above. First, the game mancala revolves around the idea of picking up the gemstones and placing them into different pits. Another term that can be used for placing the gemstones into the pits is called "sowing the seeds" [4]. The seeds are the gemstones and the action of placing them into the pits is called sowing. This is similar to when farmers sow seeds into the ground. Secondly, above we have mentioned that if the player's last gemstone is placed into their mancala, that player receives an extra turn. Another term to describe this process of gaining an extra turn is "multiple laps" or "relay sowing" [5]. Multiple laps and relay sowing are not universal features in the game of mancala, but they are frequently used throughout the world [5]. Lastly, we stated earlier that if a player's last gemstone is placed in an empty pit on its own side, then that player will take this gemstone along with the opposing player's gemstones across from the empty pit that it was placed in. This process is called "capturing gemstones." Not all cultures use this technique, but it is also very commonly used.

The game is completed when all six pits on one side are emptied. To decide who the winner of the game is, each player will then count the number of gemstones that were placed in their mancalas. Player A will count the gemstones in mancala A and Player B will count the gemstones in mancala B. The player with the most gemstones in their mancala wins the game.

1.3. **Modified Mancala.** We wanted to create our own version of the game man- cala, where there are a total of six pits instead of twelve pits. So, there would be two rows of three pits instead of six pits. In each of the six pits, there will only be one gemstone instead of four. In addition, the player's moves are mirrored and no capturing is allowed. The overall basics of mancala remain the same though. How would the game mancala change if the twelve pits were reduced to six pits with only one gemstone instead of four gemstones in each pit, along with mirrored moves and no capturing allowed? Refer to Figure 2 to see what the modified mancala board setup looks like.

Our altered version of mancala is simpler than the traditional mancala game. The reason why this version of mancala is worth researching is because it is a kid- friendly version where the game would require fewer materials, time, and confusion. Although it is a kid-friendly game, it is still not meant for people under the age of eight [1]. This mancala version would be great for a kid who would want to play this game at school or a park. They can easily gather six rocks in place of the gemstones. Gathering six rocks is much easier than gathering forty-eight rocks. Then, using chalk, the kids can draw the six pits and two mancalas on the floor. That way, this game can be played anytime even if one does not have the game board or gemstones.

Our altered version of mancala has mirrored moves, meaning that whatever Player A chooses to do, Player B will do the same but on their side as if there is a mirror. From mirroring, the other player will not be confused as to what the next move they should make since they have no choice but to follow Player A's move.

This makes the game fair to play because the moves are the same. Even if we were to take away the mirroring in the modified version, the rules are still simple to follow along.

To conduct our research, we met up and played over a hundred games of mancala together. We used the physical traditional mancala board but implemented our own rules. Not much research has been done on the alterations of mancala and what patterns may arise. After collecting numerous amounts of data, we came up with some patterns and conclusions. We expand on these findings in the Results.
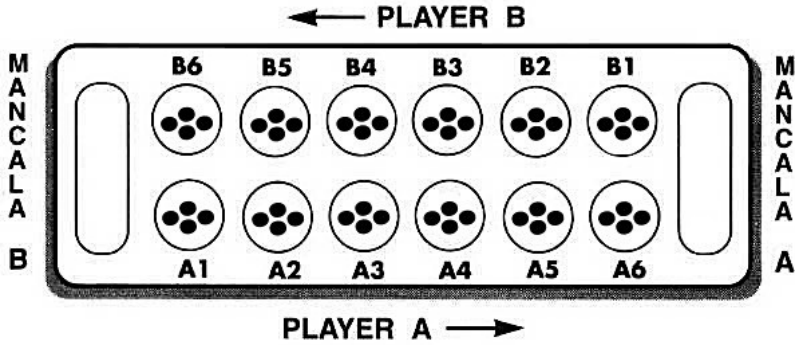
Figure 1. Traditional Mancala Board
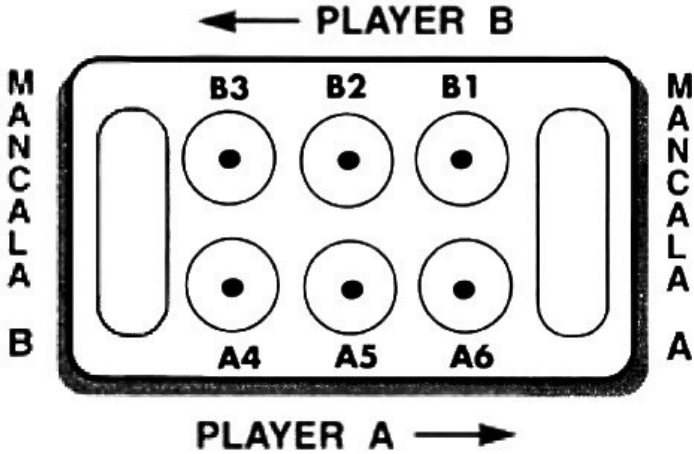
Figure 2. Modified Mancala Board

2. Results

We reduced the number of pits to six and gemstones to one while keeping mancala A and mancala B on each side (see Figure 1). Our rules are the same except we excluded capturing, and we are mirroring Player A's turns for Player B's turn. Wewanted to create a new version, so that it would be simple and quick to play. We inputted our results from playing multiple games of our version into three different

tree diagrams grouping them with three rounds consisting of pits A4, A5, and A6. There are three different pits to choose from and then from there, we had multiple games when choosing from within the three pits. This includes having different routes from a given turn.
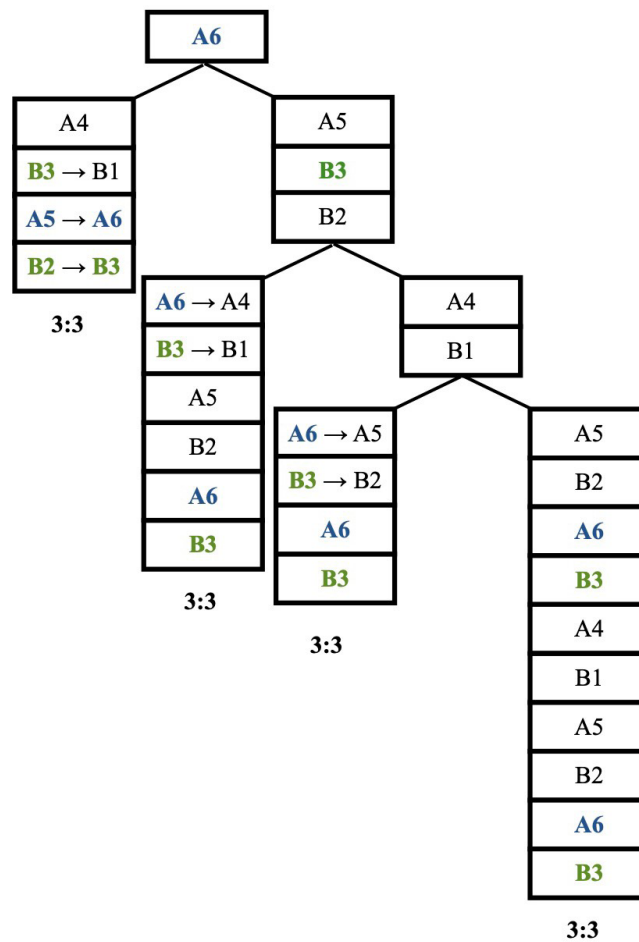
Figure 3. Round 1 Tree Diagram

Game 1 Example. An example of how to read our tree diagrams from Figures 3-6, is by first looking at Figure 3. Each branch of the tree diagram is considered one game. The arrows represent an extra turn after getting a gemstone into the mancala. Next, we color-coded whenever Players A and B got a gemstone intothe mancala. Player A is blue while Player B is green.

According to the rules, to start the game we had Player A go counterclockwise first for each game, then Player B mirrored Player A's moves. The first game consisted of starting with pit A6. The reason why we started with pit A6 is because it is closer to the mancala and starting with that pit would end up inside the mancala on our first turn. From the first turn, since pit A6's gemstone went into the mancala they get an extra turn. Rather than mirroring Player B right after it goes into the mancala, we continue just as the rules say with Player A's turn until they end their turn. Each box is considered one turn.
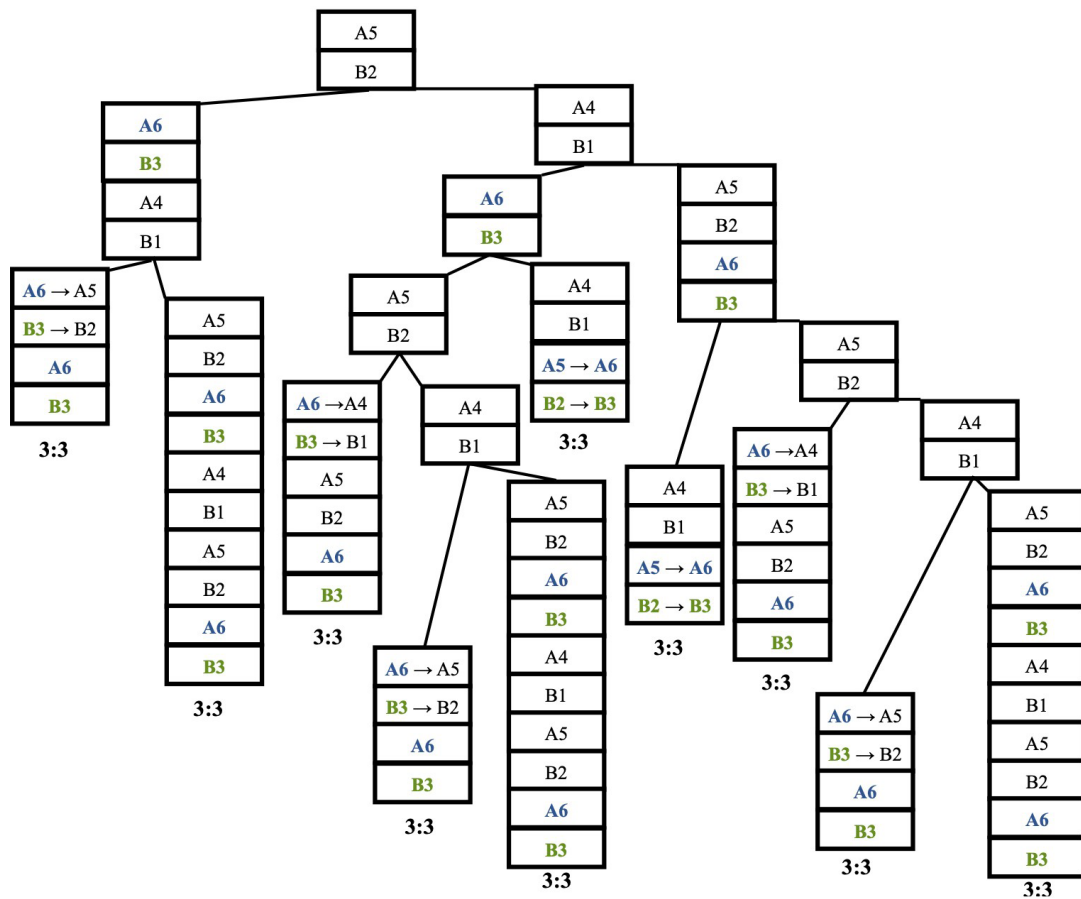
Figure 4. Round 2 Tree Diagram

So, an example of one gamewould be (see Figure 3) when we pick up the gemstone from pit A6, this would automatically end the turn resulting in it being inside a box. Since the gemstone ends up in the mancala, we colored A6 blue at the top indicating that the first turnresulted in mancala. That was Player A's first turn and now we get an extra turn. From here we see that we have two different routes to choose from, pit A4 and pitA5. So, we chose pit A4, which ended the turn, and that is why it is in its own box. Now, it is finally Player B's turn to mirror Player A's move. Player B picks up the gemstones from pit B3 which goes into the mancala, so we colored it greenin our figure, ending up in an extra turn so we added the arrow. The next turn forPlayer B was to pick up the gemstones from pit B1, which resulted at the end of Player B's turn so we need a new box for Player A's next turn. Player A picks up the gemstones from pit A5 resulting in it being in the mancala where we colored itblue getting an extra turn. Since there is only one gemstone in pit A6, Player A picks that up and it goes into the mancala. Now we automatically mirror Player Bbecause there are no more gemstones on Player A's side. So, Player B picks up thegemstone in pit B2 resulting in an extra turn. Leading to picking up the gemstone in pit B3 ending inside the mancala. This ends the first game and so we are tied3:3 gemstones in each mancala.
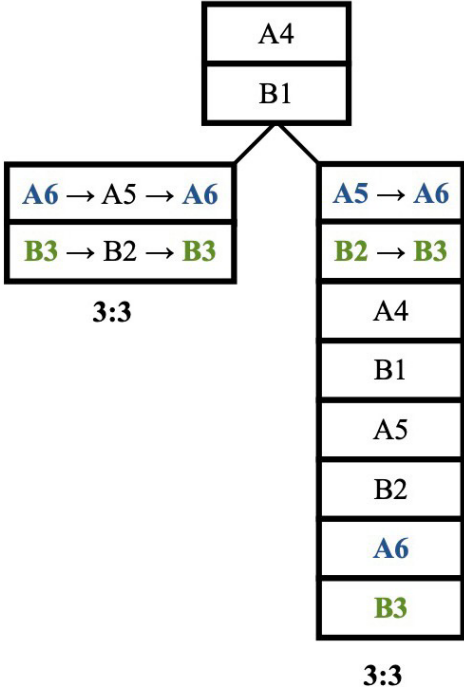
Figure 5. Round 3 Tree Diagram

1.1. **Game 2 Example.** Another example of a game (see Figure 5) is when PlayerA first picks up the gemstone in pit A4. This ends their turn since it does not go into the mancala. Now we mirror Player A's move for Player B's turn. PlayerB picks up the gemstone in pit B1 ending their turn. Now we have two different routes either Player A can choose to start with pit A6 or pit A5. Let us choose pit A6 first. The gemstone goes into the mancala resulting in an extra turn. PlayerA picks up the two gemstones in pit A5 which goes into the mancala, resulting in an extra turn. So Player A's only option is to pick up the gemstone in pit A6, ending the turn. Player B then mirrors everything from Player A by picking up the gemstone in pit B3 into the mancala. Then, picking up the gemstone in pit B2into the mancala again and lastly picking up the gemstone in pit B3 resulting at the end of Player B's turn. It is a tie once again.

As we were playing each round, finding all the possible different routes within the games, we had a total of sixteen games. From these games, both Players A and B were tied with 3:3 gemstones in each mancala. Even if we were to switch turns of which player went first, the results would be the same. Through this, we see that no computer can have power over a human being when playing this game, since the outcome always resulted in a tie. Our altered mancala did not require many strategies and critical thinking since we are mirroring the player's moves.

3. Conclusion

The traditional mancala game has twelve pits with four gemstones in each pit. The way we approached the traditional mancala game was first asking ourselves how can we change the game to make it a kid-friendly version where the game would require fewer materials, time, and confusion. We wanted to see how the game mancala would change if the twelve pits were reduced to six pits with only one gemstone instead of four gemstones in each pit, along with mirrored moves and no capturing allowed.

The way we gathered our data was first separating the games into three different rounds that start with the pits, A4, A5, and A6. Although you may think that Player A has an advantage since they go first in each game, the results would be the same since our game is mirrored. We would still end up with a 3:3 tie. By playing all the different possible routes in each round, we had a total of sixteen games, all of which ended in a 3:3 tie. However, Player A likely has more fun because they are solely controlling the gameplay.

Not much research has been done on mancala especially with the idea of mirroring and ending with a tie. For further research, we would like to know whether there are other versions in which the final result of the game ends in a tie. For ex-ample, would we result in a tie if we used the full traditional board with our same alterations? We also would like to further discover whether the final score would still be a tie if we used a four-player mancala board with our alterations. If the traditional mancala game has mirrored turns with no capturing, would the end also result in a tie? Overall, we are interested in what caused the ending tie result. Un-der what circumstances did the tie occur? It could be the change in gemstones, the change in the number of pits, no capturing mode, and/or the idea of mirroring. As you can see, we are curious about many things and about what would happen if we kept altering the traditional mancala game. Not many people have explored these topics and made research, and we believe it is worth looking into. Data can easily be gathered by playing many rounds over and over again. Computer simulations could greatly increase sample size and provide more accurate results, as well as a method for studying alternations of the game more quickly. Greater correlations, connections, and conclusions could be made if more research was done.

References

[1] Arneson, E. (2020, August 26). Basics, History and Origin of the Mancala Board Game. https://www.thesprucecrafts.com/mancala-411837
[2] Dewar, G. (2009). Mancala games for kids. https://www.parentingscience.com/mancala-games.html
[3] How to Play Mancala. (n.d.). https://www.ymimports.com/pages/how-to-play-mancala
[4] MacQuaid, M. (2020, October 1). How to Play Mancala: Rules for Popular Versions. https://bargames101.com/how-to-play-mancala-rules/
[5] Mancala. (1999). https://rocketsnail.com/mancala

Solvability of Peg Solitaire on Different Graphs

Author: Gustavo Alonso
Advisor: Roberto Soto
California State University, Fullerton
Department of Mathematics

Abstract

The traditional version of peg solitaire challenges the player to remove all but one of the pegs from a board. We consider how this task may change in difficulty if the game board is of different graphs, such as the cyclical or windmill variants. The vertices act as containers for the pegs that allows peg movement provided that the vertices connect via edges. Additionally, the set of natural numbers differentiate peg colors to create and play various versions of these games. By developing a computer program, we determine that the cyclical and windmill graphs are solvable, i.e., there is a strategy to win the game based on peg placement on the graph. Finally, we consider the possibility that peg placement is irrelevant to solve these graphs when specific sizes and peg colors are selected to play the game.

Statistical Modeling for Discovery:
How Movements Within Counties Affect the Spread of the COVID-19 Pandemic

Authors: Gwendolyn Lind, Caleb Peña, and Seth Arreola
Advisor: Dr. Sam Behseta
California State University, Fullerton
Department of Mathematics

Introduction

In March of 2020, the World Health Organization declared the novel coronavirus or COVID-19, a global pandemic. Currently, over 100 million cases and over 2 million deaths worldwide have been reported and these figures continue to grow (see World Health Organization's COVID-19 Dashboard.) While there is a significant research associated with building predictive models for the number of infected cases, mortality, and hospitalization, the literature on models excavating the effects of mobility among populations, and their relationship with economic status of their communities is surprisingly scarce. It is with the intention of filling that void that we have attempted to take into account those predictors in our statistical modeling approach.

Purpose

In this work, we aim to construct a series of straightforward yet robust statistical and machine learning methods to understand the effect of economic disparities in conjunction with the mobility among selected counties in Southern California, on the nuances and patterns of variation of the reported COVID-19 infection rates. Broadly, our working hypothesis is that communities that struggle financially, witnessed a significantly higher rates of infection in certain time periods, and one way to tease that effect out is to look at the intensity of the mobility in the residential and business areas in those communities. We mine and collate varied sources and volumes of data for building a data repository for our statistical modeling approaches. This enables us to subsequently examine, and eventually select, a series of robust models for predicting COVID-19 cases through the economic status and mobility of the communities of interest.

Literature Review

According to the journal Nature, about 4% of the world's research output was devoted to corona virus in 2020 which translates to thousands of published materials and preprints, peer-reviewed or otherwise (Else, 2020). Consequently, a cloud of articles on the analytics of the COVID-19

data appeared in a variety of biological, mathematical, statistical, and computational journals, among others. The most frequent theme in all of those is models for prediction of the diagnosis, infection, mortality, and hospitalization rates (Zaobi, 2021; Shinde et al., 2020; Sujath, 2020; Liu et al., 2020; Tsallis and Tirnakli, 2020; He et al., 2020; Zeroual 2020). Nevertheless, there is a significantly limited published work, so far, on the statistical or mathematical models for studying the dynamics of COVID-19 fluctuations when viewed through the prism of the economic status of their communities. Since economic disparity in the U.S. often correlates with racial disparity, there is an urgent need to tackle the issue, and thereby fill that void in the literature (McLaren, 2020; Tomaskovic-Devey, 2020; Yechezkel 2020). In this work, we draw from the rich and extensive literature on the statistical modeling for the time-series regression, when the response variable is the number of incidents or random occurrences, and the predictors of the model are time dependent as well (for a comprehensive review of the literature, see the references cited in Kedem and Fokianos, 2002).

Data Sources and Data Wrangling

Data Sources: Google, Apple, and USA Facts

Prior to building models, there needs to be a data repository that weaves together the multiple sources of data, namely information regarding mobility in the residential and business areas, number of infected cases and mortality rate associated with each community, and its socio-economic data. These data sources are scattered across the web and are of varying formats and modalities.

We began our work by extracting publicly available mobility data from two sources (Google and Apple mobility data dashboards), as well as the infection and mortality data from USA FACTS, a public service data dashboard, pulling reliable information across multi engines such as the Centers for Disease Control and Kaiser Permanente.

Information from Google were extracted from one of their side-repositories, titled Community Mobility Reports (CMR) which has been available to the public, since early January 2020. The site facilitates importing data by location, at the county level, and is intended to help remediate the impact of COVID-19 in concert with the changes in Federal and State policies, public awareness, and other evolving dimensions affecting the spread of the disease at the macro level. Therefore, the numerical reports acquired from this resource, tabulate the movement trends over time, and by geographic regions. There are six categories within the CMR of places, namely Retail and Recreation (restaurants and shopping centers), Grocery and pharmacy (farmers markets

and drug stores), Parks (national and public), Transit Stations (subways and buses), Workplaces, and Residential data. Google collects this information in an aggregated fashion, followed by anonymizing it, to uphold the privacy of its mobile users. A multi-dimensional data frame is then structured through the location history of the users' cellphone and other mobile apps. The information is refined through tracing the spatiotemporal position of the users, and according to their interaction with at least one of the six above listed categories. Each category has a different metric for a certain day of the week, and in contrast to a baseline which represents a normal value for that day of the week. More precisely, the baseline value is taken as the median of the data from the 5-week period during January 3 to February 6, 2020, to reflect a sense of "normalcy" prior the widespread of the virus. Table 1 provides an example of the type of data that can be accessed from Google's CMR repository. As shown in there, for each particular day since January 2020, one can acquire the percentage of changes, whether they are upward or downward, from the baseline value, and according to one of the six categories of interest. From Table 1, specifically, demonstrates an 18% downward shift from the baseline in the Retail and Recreation category.

Table 1. Google Data Mobility trends reporting of the Retail and Recreation category.

In contrast to the CMR data, Apple data mobility repository focuses is built through tracking transportation, rather than the time spent in a particular location. The data is generated by counting the number of requests made to Apple Maps for multiple applications such as driving directions, walking, and cumulative transit reports. Similar to the CMR data, Apple created a baseline value, but only using a single day information, associated with January 13th, 2020. To this end, we find these "baseline-oriented" data quite useful, as our models will track the nuances or uncertainty resulted from the changes in mobility across time. Therefore, the concept of baseline allows us to incorporate scale-free and comparable predictors to our models.

As shown in Table 2, Apple data, called Apple Metric, are reported in the units of hundreds. Nevertheless, due to the existence of the baseline, and as part of our data wrangling efforts, we transformed the data to changes, in percentage, to the baseline. More specifically, as shown in Table 2, in the County of Orange, on January 15, 2020 there was a 9% (referring to the value

109.61) increase from the baseline in the "walking" category, whereas on January 16, 2020, one can observe a 5% (95.27) decrease in the same category.

Table 2. Apple Data Mobility trends reporting of navigation requests.

	geo_type	region	transporta	alternative_	sub-region	country	1/13/2020	1/14/2020	1/15/2020	1/16/2020	129.25	107.79
3860												
3861	county	Orange County	walking		California	United Sta	100	104.74	115.1	122.04	144.22	168.84
3862	county	Orange County	walking		New York	United Sta	100	102.78	109.61	95.27	127.25	98.68
3863	county	Orange County	walking		Florida	United Sta	100	97.45	101.7	104.94	134.51	162.93

We accumulate the Google and Apple data, and the USA Facts data to build a comprehensive repository of the modes of mobility and the rates of infection and mortality. To include the COVID-19 cases the USA FACTS, we applied a technique known as "web-scraping" (Glez-Peña et al., 2014), allowing us to select, filter, and register the data on the dates for which we also have collected the mobility information. The data reflects the number of new cases, total number of cases daily, and total number of deaths daily across the U.S. counties.

Lastly, to complete our data repository, we employed the same web-scraping techniques to collect the information representing the average income of the counties of interest.

Data Wrangling: Building Moving Windows as a Mean for Standardizing Data

Given the magnitude of information from each of the three sources (Google, Apple, and USA Facts), it is imperative to condense the data into one seamless, accessible, and comprehensible format. Taking a step back before combining these datasets together, we needed to incorporate important key facts about the nature of the Coronavirus and its manifestation in human subjects. To give an example, according to the literature, the exposure to symptoms onset or the incubation period, takes somewhere between 2 to 14 days (Lauer et al., 2020). Moreover, research suggests that people are most contagious in the 48 hours before symptoms develop and they remain contagious for up to 10 days after onset of symptoms (Byrne et al. 2020). To account for the above, we adjusted the time epochs or time frames in our data repository. This is similar to creating a moving average or moving sum of data, so that it takes into account the incubation period associated with the disease. Assuming the infection data is of size n, we can refer to the time series of data as follows:

$$X_{cases,1} \dots X_{cases,n}$$

Since the cases reported in a given day are influenced by past interactions, we created a new variable Y, by considering a moving window of length or lag 4 over the data set of the infected cases. This is equivalent with the following:

$$Y_{cases,j} = \sum_{j=1}^{7} X_{cases,i-j-4}$$

Since there is a good deal of variation in how the disease is manifested in each individual, this approach may lead to overestimations of the true cases reported, but the benefit of this exercise is that it generates a holistic data during time bins of equal lengths. This strategy will greatly empower our model approaches. Utilizing the strategy of building moving windows is particularly beneficial for the mobility data, as we do not know which days people were infected, who was infected, and their travel destinations, prior to developing any symptoms.

To achieve the abovementioned tasks, we wrote elaborate R code, taking more than 4 months of time, for collecting data, merging and binding data features, and automatically updating data from their varied sources. Moreover, though data wrangling, we were able to extract the information related to seven Southern California counties (Imperial, Riverside, San Diego, San Bernardino, Los Angeles, Orange, and Ventura). These counties were chosen specifically with the idea that they represent a host of socioeconomic characteristics, mobility patterns, and COVID-19 cases (Azar et al., 2020).

Methods: Modeling the Spread of COVID-19 via Mobility and Economic Data

As described in the above, the goal of this project is to explore the effect of mobility variables, and the average-income of the counties from which the data were obtained on the nuances of COVID-19 infection rates. The data are in the form of time-series. That is, the values associated with the COVID-19 cases, and mobility features, are time dependent. Counties are assumed to be independent. This may not be entirely true, as they follow certain federal and statewide guidelines. Nevertheless, we do not think the latter assumption is an overt simplification or oversight in terms of the structure and thereby the outcomes of our models.

In this project, we would like to be able to build a regression model whose response variable is the rates of infection, and the predictors in the model are mobility and county values. This means we aim to construct a regression model for time series. This concept has been explored thoroughly in the literature (Ke-dem and Fokianos, 2004). In particular, we model the time series of the cases or the response variable with a so-called non-homogeneous Poisson process (Zegerand Qaqish, 1988). This approach considers a distinct Poisson distribution for the value of the response at time t, and allows for those distributions to be related to each other via a time lag whose value can be determined through statistical hypothesis tests. More importantly, a regression model can be structured through a mechanism known as generalized linear models or GLM in short(Nelder and Wedderburn, 1972). When GLM is coupled with ARMA, the result is a powerful modeling engine known as GLARMA (Davis et al., 2005). We employ a class of GLARMA models for modeling the rate of COVID-19 among counties of interest.

There is a vast literature on the foundations of modeling time-series data (for a review of the literature see Box, Jenkins, Reisnel, 1994; Brockwell and Davis,2007). A general paradigm for modeling time-series data was introduced by Box and Jenkins (1970) in which one models data based on two sets of components: an auto-regressive component that takes into account the number of lags or the period of dependency among data, and a moving average component that incorporates the effect of errors resulted from fitting the model to the previous lags. When combined, the ARMA model (AR stands for Auto Regressive and MA for Moving Average) is a powerful tool for modeling time-series data. Specifically, we let , i.e. the number of cases or infection rate at time , to be model with a Poisson distribution whose intensity function represent a time-dependent regression model or

$$\lambda_t = exp(\alpha + \beta x_t^T + z_t),$$

in which

$$z_t = \sum_{i=1}^{p} \phi_i(z_{t-i} + e_{t-i}) + \sum_{j=1}^{q} \theta_i e_{t-i,}$$

Where α indicates the baseline, is the vector of the mobility predictors is comprised of auto-regressive terms of lag p and moving-average terms of lag q, and represents the error terms. This will allow us to estimate all the parameters in the model by an iterative maximum likelihood estimation procedure.

Results

In addition to GLARMA, we examined a series of other machine learning models. A comparative analysis confirmed the optimality of the GLARMA in the context of our data analysis. To implement the model, we utilized the statistical package *GLARMA* in the programming language R (Dunsmuir and Scott, 2015) for the purpose of modeling the pandemic data. We calibrated and selected optimal values for the parameters of the model, through a technique known as cross-validation (Mosteller and Tukey, 1968; Stone, 1974). We also applied Bootstrap (Efron et al., 1984; Bühlmann, 2002) for constructing confidence bands for the predicted values.

Figure 1 provides a nice summary of our analysis. In Figure 1, the actual cases are depicted with the red curve, our predicted values are in blue, and the dotted curves are confidence bands around the predictive model.
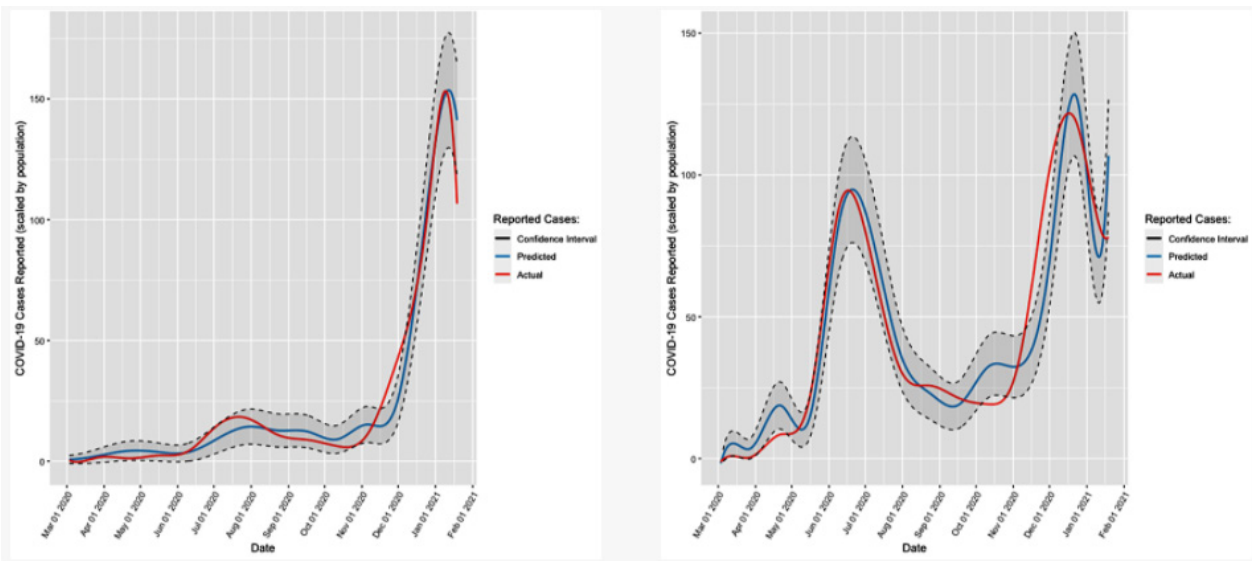
Figure 1. The left graph features the number of COVID-19 cases in Ventura County and the right graph features Imperial County. Here, we can see the trends differ drastically.

It is assuring that all actual values fall within the moving confidence interval of the model. Two counties in Southern California are showcased in that figure: Ventura County (on the left) and Imperial County (on the right). Notably, Imperial County suffers from a significantly lower median income in contrast to Ventura County (according to the U.S. Census Bureau's *American Community Survey* in 2019, the median income in Ventura county was reported at $33,814, whereas the same metric at Imperial county was reported at $18,245.) Strikingly, during the months of June to September 2020, when Imperial County experienced a significant surge in the reported cases of COVID-19, Ventura's pattern of the spread of the disease remained nearly flat. This outcome was not only consistent across the Southern California counties, but also was clearly observable when we fitted the model nation-wide, as shown in Figure 2. In Figure 2, the bimodal pattern of the spread of the pandemic, referring to two local surges, are observed.
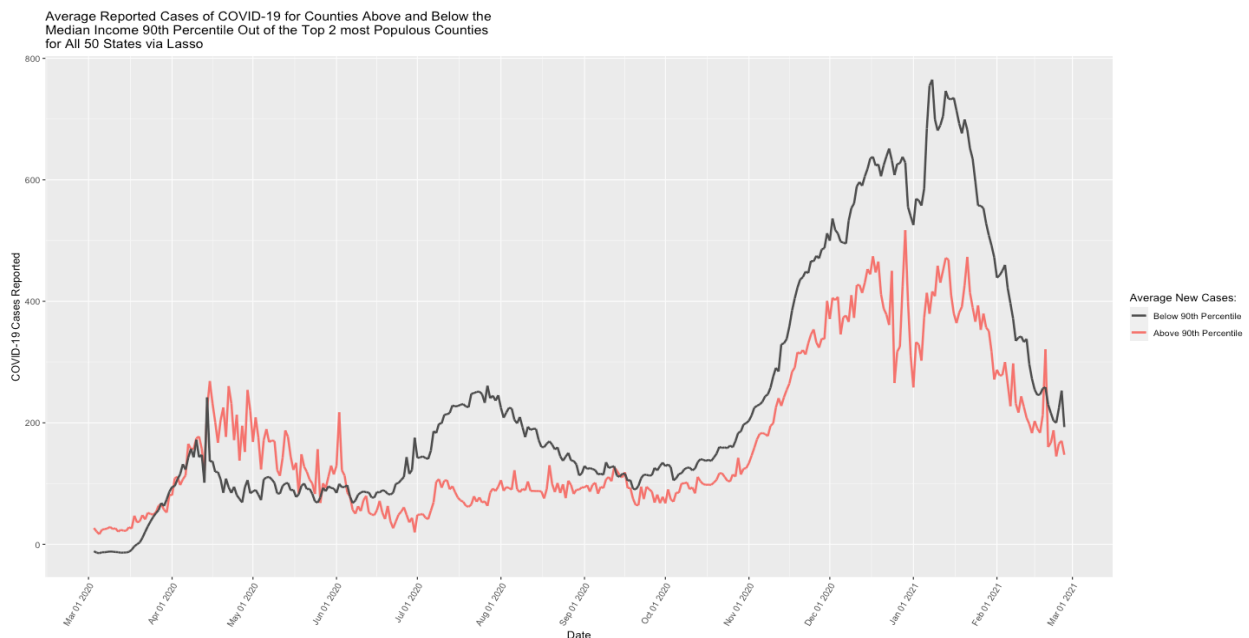
Figure 2. The nation-wide counties (two most populated counties per state in the U.S.) trend mimics the pattern of the spread of the disease to the one observed in the more economically challenged counties in Southern California.

Interestingly, there were counties such as Orange County, California, in which the spread of the disease signifies a combination of the patterns of both Ventura and Imperial counties, as shown in Figure 3.
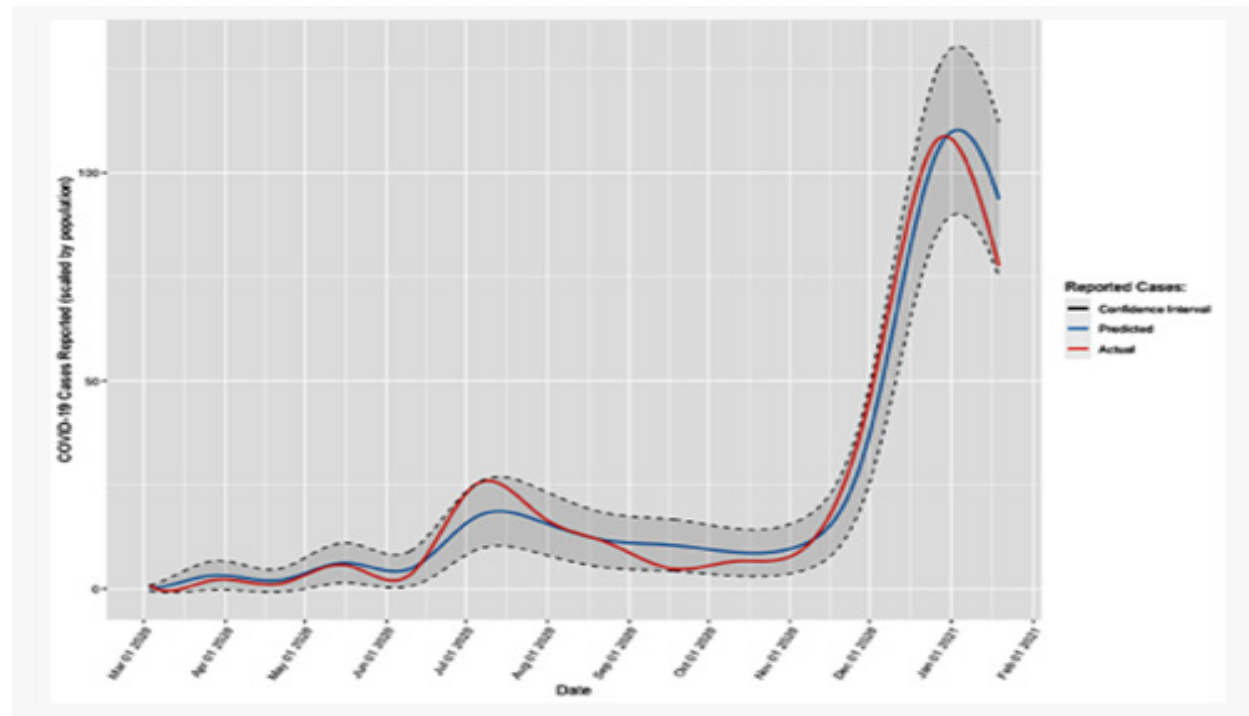
Figure 3. Trend of Orange County in the number of COVID-19 cases is a combination of the ones for Ventura and Imperial counties.

This may be explained due to the significant variation of the median income, and considerable economic disparities, among the cities of Orange County (Scott, 1986; Morello-Frosch 2002, Bohn and Danielson, 2016).

Discussion/Conclusion

In this work, we demonstrate how statistical models can be utilized to capture the patterns of the spread of pandemics among the communities with distinct economic dynamics. Specifically, these patterns are more vividly at display, when the modeling approach includes different forms

of mobility among the residents in seven Southern California counties (Imperial, San Diego, San Bernardino, Los Angeles, Orange, Riverside, Ventura). In those communities, we found that counties that were less prosperous showed to have a higher increase in the number of coronavirus cases, particularly in the early window of the pandemic, compared to counties that are more affluent economically. Remarkably, this pattern is also observable when more prosperous counties across the country are contrasted with those that are economically more challenged. Moreover, by using a class of time-dependent models known as GLARMA, we were able to accurately predict the number of COVID-19 cases from January 2020 to date. Our model proved to be a robust method for connecting mobility to predicting the number of cases. To give an example, during the months of June to September in 2020, Imperial County experienced a significant increase in its reported number of cases, whereas the pattern of cases in Ventura, a considerably more prosperous county, remained steady and flat.

In general, the approach discussed in our work, allows for adopting fresh outlook for tracking economic disparities under the pandemic. To offer a pragmatic solution, by identifying different trends of the growth, our models can provide critical information to determine which areas are in the most need of additional resources. This is important for decision makers when equally allocating sufficient resources to different counites. These resources, in many forms, may result in more informed policymaking, and may contribute to providing better healthcare assistance, and job security, among others. Additionally, our research is showing the power of data science by promoting statistical methodologies that can address real major local and global challenges such as this pandemic.

Future Work
Currently leading a new team of undergraduate and graduate students to further investigate this multilevel research project. We wish to add new relevant information to our dataset to understand these disparities based within the seven Southern California counties. The next logical step is to examine the economic disparities within each of the southern California counties by obtaining data at the city level. We plan to incorporate new COVID-19 data associated with the number of vaccines, as well as preventative measures such as masks mandates policies, and the overall growth of the pandemic. Lastly, the aim for our research is to provide insights on how the pandemic behaves in addition to becoming valuable information that can create a positive impact.

Acknowledgements

References

Apple. "Mobility Trend Reports." 2020, covid19.apple.com/mobility.

Azar, Kristen M. J., et al. "Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California: Health Affairs Journal." Health Affairs, 21 May 2020, www.healthaffairs.org/doi/full/10.1377/hlthaff.2020.00598.

Bhon, Sarah, and Caroline Danielson. "Income Inequality and The Safety New in California." *Public Policy Institute of California*, 2016, www.ppic.org/content/pubs/report/R_516SBR.pdf.

Box, George Edward Pelham, and Gwilym Jenkins. *Time Series Analysis, Forecasting and Control* . 5th ed., Holden-Day, Inc, 1990.

Box, George E.P., et al. *Time Series Analysis; Forecasting and Control*. 3rd ed., Prentice Hall, 1994.

Brockwell, Peter J, and Richard A. Davis. Introduction to Time Series and Forecasting. Springer New York, 2002. Print

Byrne AW, McEvoy D, Collins AB, Hunt K, Casey M, Barber A, Butler F, Griffin J, Lane EA, McAloon C, O'Brien K, Wall P, Walsh KA, More SJ. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. BMJ Open. 2020 Aug 5;10(8):e039856. doi: 10.1136/bmjopen-2020-039856. PMID: 32759252; PMCID: PMC7409948.

Bühlmann, Peter. "Bootstraps for time series." *Statistical science* (2002): 52-72.

Davis, Richard A, et al. "Maximum Likelihood Estimation for an Observation Driven Model for Poisson Counts." *Methodology and Computing in Applied Probability*, Kluwer Academic Publishers, 1 Jan. 1992, link.springer.com/article/10.1007%2Fs11009-005-1480-4.

Efron, Bradley, and R.J. Tibshirani. *An Introduction to the Bootstrap*. 1st ed., Chapman and Hall/CRC, 1994.

Else, Holley. *COVID in Papers: A Torrent of Science*. page 553 ed., vol. 588, 2020.

Glez-Peña, Daniel, et al. "Web Scraping Technologies in an API World, *Briefings in Bioinformatics*" Volume 15, Issue 5. *OUP Academic*, Oxford University Press, 30 Apr. 2013, Pages 788-797 academic.oup.com/bib/article/15/5/788/2422275.

Google. "COVID-19 Community Reports." 2020, www.google.com/covid19/mobility/.

He, Shaobo, and Yuexi Peng. "SEIR Modeling of the COVID-19 and Its Dynamics." *Nonlinear Dynamics*, Springer Netherlands, 18 June 2020, link.springer.com/article/10.1007/s11071-020-05743-y.

Kedem, Benjamin, and Konstantinos Fokianos. *Regression Models for Time Series Analysis*. John Wiley & Sons, 2002.

Lauer, Stephen A, et al. "The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application." *Annals of Internal Medicine*, American College of Physicians, 5 May 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7081172/.

Liu, Zhi Hua, et al. "Predicting the Cumulative Number of Cases for the COVID-19 Epidemic in China from Early Data." *Mathematical Biosciences and Engineering : MBE*, U.S. National Library of Medicine, 7 Apr. 2020, pubmed.ncbi.nlm.nih.gov/32987515/.

McLaren, John. "Racial Disparity in COVID-19 Deaths: Seeking Economic Roots with Census Data." *NBER*, National Bureau of Economic Research, 22 June 2020, www.nber.org/papers/w27407.

Mosteller, F, and J Turkey. *Data Analysis, Including Statistics. In Handbook of Social Psychology*. 2nd ed., Wiley, 1968.

Morello-Frosch, Rachel, et al. "Environmental Justice and Regional Inequality in Southern California: Implications for Future Research." *Environmental Health Perspectives*, U.S. National Library of Medicine, Apr. 2002, www.ncbi.nlm.nih.gov/pmc/articles/PMC1241158/.

Nelder, J. A., and R. W. M. Wedderburn. *Generalized Linear Models. Journal of the Royal Statistical Society*. Vol. 135, Wiley, 1972.

Scott, Allen J. "HIGH TECHNOLOGY INDUSTRY AND TERRITORIAL DEVELOPMENT: THE RISE OF THE ORANGE COUNTY COMPLEX, 1955–1984." *Taylor & Francis*, Urban Geography, 15 May 2013, www.tandfonline.com/doi/abs/10.2747/0272-3638.7.1.3.

Shinde, Gitanjali R., et al. "Forecasting Models for Coronavirus Disease (COVID-19): A Survey of the State-of-the-Art." *SN Computer Science*, U.S. National Library of Medicine, 11 June 2020, pubmed.ncbi.nlm.nih.gov/33063048/.

Stone, M. *Cross-Validatory Choice and Assessment of Statistical Predictions*. 2nd ed., vol. 36 Pages 111-147, Wiley, 1974.

Sujath, R., et al. "A Machine Learning Forecasting Model for COVID-19 Pandemic in India." *Springer Link*, Springer Berlin Heidelberg, 30 May 2020, link.springer.com/article/10.1007/s00477-020-01827-8.

Tomaskovic-Devey, Donald. "The COVID-19 Pandemic: Normal Accidents and Cascading System Failures (Inequality, Poverty, and Mobility)." *American Sociological Association*, Center for Employment Equity at the University of Massachusetts, 6 Aug. 2020, Page 25 Vol 48

Tsallis, Constantino, and Ugur Tirnakli. "Predicting COVID-19 Peaks Around the World."

Frontiers, Frontiers, 20 May 2020, www.frontiersin.org/articles/10.3389/fphy.2020.00217/full.

USA Facts. "US Coronavirus Cases & Deaths by State Track COVID-19 Data Daily by State and County." 2020, usafacts.org/visualizations/coronavirus-covid-19-spread-map/.

World Health Organization ,"WHO Coronavirus (COVID-19) Dashboard." *World Health Organization*, , 2020, covid19.who.int/.*https://covid19.who.int/*

Yechezkel, Matan, Amit Weiss, Idan Rejwan, Edan Shahmoon, Shahaf Ben Gal, Dan Yamin et al. "Human Mobility and Poverty as Key Drivers of COVID-19 Transmission and Control." *X-MOL*, MedRxiv- Epidemiology, 19 July 2020, www.x-mol.com/paper/1285073892858839040.

Zeroual, Abdelhafid, Harrou F., Dairi A., Sun Y et al. "Deep Learning Methods for Forecasting COVID-19 Time-Series Data: A Comparative Study." *Chaos, Solitons & Fractals*, Pergamon, 15 July 2020, www.sciencedirect.com/science/article/pii/S096007792030518X.

Zoabi, Yazeed, Deri-Rozov, S. & Shomron, N et al. "Machine Learning-Based Prediction of COVID-19 Diagnosis Based on Symptoms." *Nature News*, Nature Publishing Group, 4 Jan. 2021, www.nature.com/articles/s41746-020-00372-6.

Zeger, Scott L., and Bahjat Qaqish. *JSTOR*. 4th ed., vol. 44 Pages 1019-1031, International Biometric Society, 1988.

Yechezkel, Matan, Amit Weiss, Idan Rejwan, Edan Shahmoon, Shahaf Ben Gal, Dan Yamin et al. "Human Mobility and Poverty as Key Drivers of COVID-19 Transmission and Control." *X-MOL*, MedRxiv- Epidemiology, 19 July 2020, www.x-mol.com/paper/1285073892858839040.

Zeroual, Abdelhafid, Harrou F., Dairi A., Sun Y et al. "Deep Learning Methods for Forecasting COVID-19 Time-Series Data: A Comparative Study." *Chaos, Solitons & Fractals*, Pergamon, 15 July 2020, www.sciencedirect.com/science/article/pii/S096007792030518X.

Zoabi, Yazeed, Deri-Rozov, S. & Shomron, N et al. "Machine Learning-Based Prediction of COVID-19 Diagnosis Based on Symptoms." *Nature News*, Nature Publishing Group, 4 Jan. 2021, www.nature.com/articles/s41746-020-00372-6.

Zeger, Scott L., and Bahjat Qaqish. *JSTOR*. 4th ed., vol. 44 Pages 1019-1031, International Biometric Society, 1988.

Riemann Sums: Approximating Integrals With Semi-Ellipses

Authors: Holly C. Anderson, Christine N. Gamez, and Summer K. Andrews
Advisor: Alison S. Marzocchi, Ph.D.
California State University, Fullerton
Department of Mathematics

Abstract

This paper explores whether Riemann sum approximations using semi-ellipses are more accurate than rectangular Riemann sum approximations and whether they increase in accuracy as n increases. We derived a formula for our Semi-Ellipse Approximation Method (SEAM), tested it with four equations using different n-values, and compared with the results we obtained using Right and Left Rectangular Approximations, as well as with the true value obtained using integration. Our results using SEAM were the most accurate with the function that was concave up on the given interval and increased in accuracy as n increased. The functions that were concave down were less accurate and became imprecise as n increased.

1. Introduction

Riemann sums are approximations of definite integrals, the area under the curve of a function, over a closed interval. When function f is positive, the Riemann sum is the sum of areas of the approximating shapes. When function f has both positive and negative values, the Riemann sum is the net area, or the sum of areas above the x-axis subtracted by the sum of areas below the x-axis (Stewart, 2016). Rie- mann sums are useful when a function does not have a clear antiderivative, which is often the case in real-world functions (Herman & Strang, 2016). Seeing that real-world functions may not have an exact antiderivative, determining the most accurate approximations of these antiderivatives is crucial. This paper deals with the question of whether integrals could be approximated more accurately with a different shape; the semi-ellipse. We aim to explore the accuracy of approximations using semi-ellipses and to determine which functions the semi-ellipse approxima- tions work well with.

There are some functions for which rectangular Riemann sum approximations have a large error. The accuracy of the rectangular Riemann sum typically depends on the graph and how many rectangles are used to approximate the sum. Generally, the more rectangles used, or the larger n value, the more accurate the approximation (Leclair, 2010). In this paper, we will explore whether these observations are more accurate using our formula for Riemann sum approximations using semi-ellipses.

From (Alexendre et al., n.d.), we know the formula for a Riemann sum using rectangles is the summation of the area of the rectangles in the given interval. To create our equation, we looked at the formula for the area of an ellipse,

$$A = \pi r_1 r_2$$

where r1 and r2 are the vertical and horizontal radii of the ellipse. We divided this equation in half in order to determine a formula for the area of a semi-ellipse. We will use this to produce (1).

To calculate the integral approximations using semi-ellipses, we created a formula for SEAM (Semi-Ellipse Approximation Method). For some function $f(x)$, to approximate $\int_a^b f(x)\, dx$ where $a, b \in \mathbb{R}$, we have

(1)
$$SEAM_{n=N} = \sum_{k=0}^{n-1} \frac{\pi}{2}\left(f\left(k\Delta x + \frac{\Delta x}{2}\right)\right)\frac{\Delta x}{2}$$

where n is the number of semi-ellipses we are using to partition the integral, $\Delta x = \frac{b-a}{n}$,and k is the counter we are using to sum each semi-ellipse.

Using this formula, we will explore estimates using SEAM and compare them to the estimates using Left Rectangular Approximation Methods (LRAM) and Right Rectangular Approximation Methods (RRAM) and to the true value using integra- tion. With these comparisons, we will determine how accurate our estimations are, learn which functions SEAM works well with, and speculate why this approximation method works better or worse for certain functions.

We hypothesize that as n increases, the accuracy of our approximations using SEAM will also increase, similar to LRAM and RRAM. We also suspect that our approximation method will be able to determine the integrals of graphs such as f (x) = sin x since the graph appears elliptical in shape between each x-intercept.

2. Results

We will explore the accuracy of our estimates using SEAM by comparing them to the estimates using LRAM, RRAM and to the true value using integration. We compare these values in the tables given. Note that most of the values in the tables are approximations rounded to the neared thousandth. To visualize how we are approximating these integrals, we have included images of the graphs on their respective intervals of integration with the semi-ellipses drawn onto them. We chose these functions because they vary in shape, allowing us to seek commonalities in an assortment of graphs.

2.1. **Equation 1.** We will be approximating the integral of $f(x) = x$ from $[0,5]$, splitting the interval into $n = 2, 5, 10$ partitions. The results of these calculations are shown in Table 1. The numbers in bold are closest in their respective row to the true value of $\int_0^5 x\,dx$.

n-value	LRAM	RRAM	SEAM	True Value
2	6.25	18.75	**9.81748***	
5	**10**	**15**	9.81748*	
10	**11.25**	**13.75**	9.326	
True Value				12.5

TABLE 1. Approximations of $\int_0^5 x\,dx$
*Extra digits are shown to indicate the values for $n = 2$ and $n = 5$ are equivalent

We will show how we used (1) to calculate $SEAM_{n=5}$.

$$\Delta x = \frac{b-a}{n} = \frac{5-0}{5} = 1$$

$$SEAM_{n=5} = \sum_{k=0}^{4} \frac{\pi}{2}\left(f\left(k\Delta x + \frac{\Delta x}{2}\right)\right)\frac{\Delta x}{2}$$

$$= \sum_{k=0}^{4} \frac{\pi}{2}\left(f\left(k + \frac{1}{2}\right)\right)\frac{1}{2}$$

$$= \frac{\pi}{4}\left[f\left(0 + \frac{1}{2}\right) + f\left(1 + \frac{1}{2}\right) + f\left(2 + \frac{1}{2}\right) + f\left(3 + \frac{1}{2}\right) + f\left(4 + \frac{1}{2}\right)\right]$$

$$= \frac{\pi}{4}\left[f\left(\frac{1}{2}\right) + f\left(\frac{3}{2}\right) + f\left(\frac{5}{2}\right) + f\left(\frac{7}{2}\right) + f\left(\frac{9}{2}\right)\right]$$

$$= \frac{\pi}{4}[.5 + 1.5 + 2.5 + 3.5 + 4.5]$$

$$\approx 9.817$$

Using SEAM to approximate $\int_0^5 x\,dx$ is generally an *underestimate* (Table 1) (Figure 1)
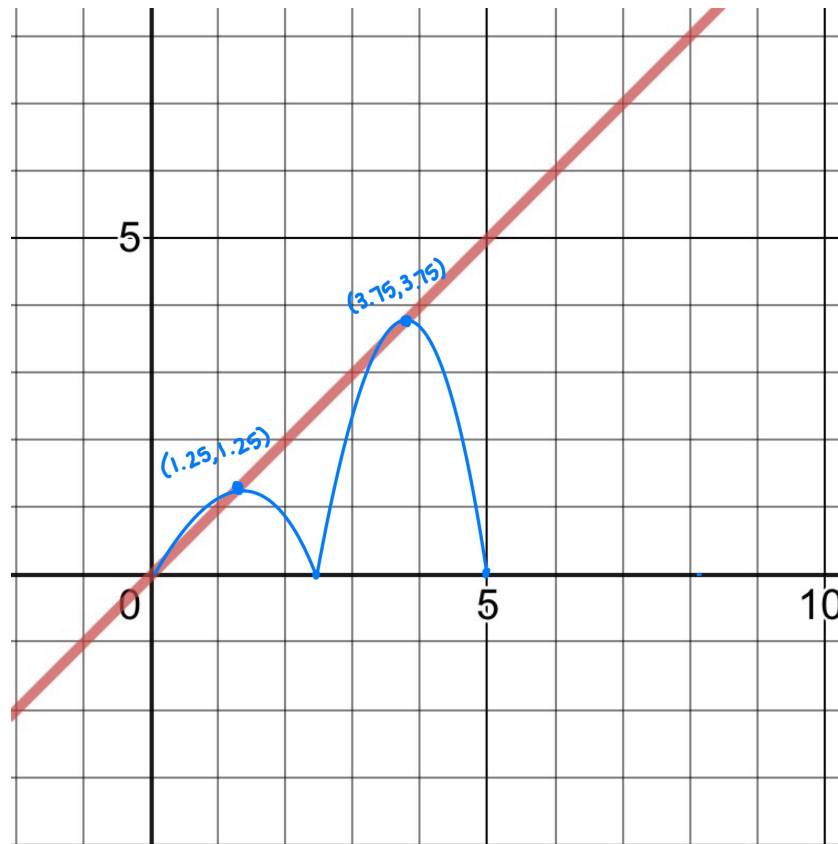
The values for $SEAM_{n=2}$ and $SEAM_{n=5}$ were equivalent. We can see from Table 1 that SEAM was the most accurate for $n = 2$; however, as n increased, SEAM became less accurate than RRAM or LRAM. This contradicts the hypothesis we made in the introduction. We conjecture that this is likely because $f(x)$ isa linear function.

2.2. **Equation 2.** We will be approximating the integral of $f(x) = x^2$ from [0,2], splitting the interval into $n = 1, 2, 4$ partitions. The results of these calculations
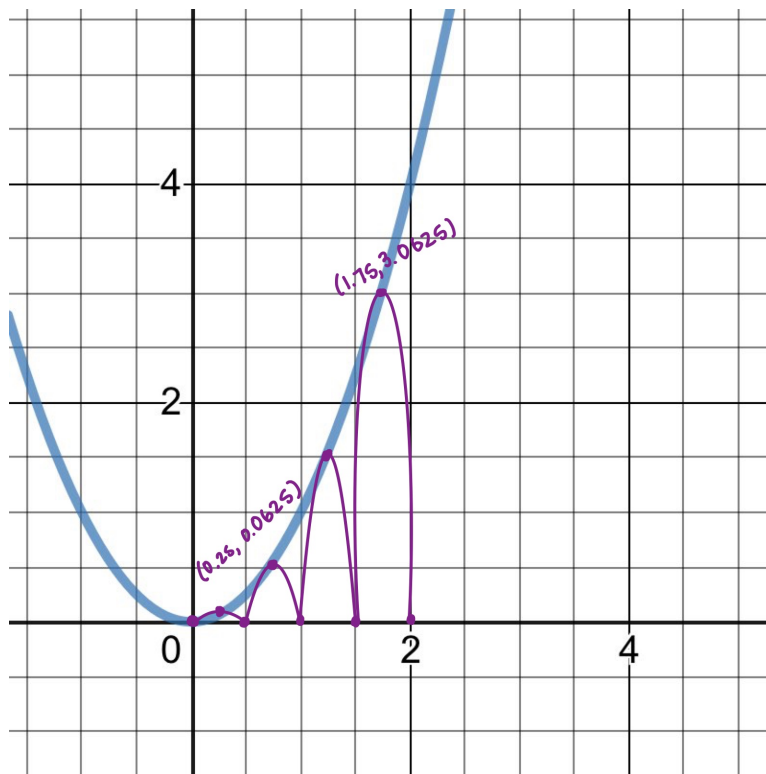
FIGURE 2. Caption

are shown in Table 2. The numbers in bold are closest in their respective row to the true value of $\int_0^2 x^2 \, dx$.

n-value	LRAM	RRAM	SEAM	True Value
1	0	8	$\frac{\pi}{2} \approx \mathbf{1.571}$	
2	1	5	**1.964**	
4	1.75	3.75	**2.061**	
True Value				$\frac{8}{3} \approx 2.667$

TABLE 2. Approximations of $\int_0^2 x^2 \, dx$

For this equation, and future equations, we will not be showing how we derived the values for the SEAM column of the tables, as it is the same process as we demonstrated in Equation 1.

In approximating $\int_0^2 x^2 \, dx$, our SEAM approximation outperformed the LRAM and RRAM for all explored values of n. The closest value to the true value was $SEAM_{n=4}$, our largest n value explored. Adhering to our hypothesis, the approximations became more accurate as n increased. We suspect SEAM was a better approximation for $f(x) = x^2$ because unlike $f(x) = x$, $f(x) = x^2$ is a nonlinear function.

2.3. **Equation 3.** We will be approximating the integral of $f(x) = \sin x$ from $[0, \pi]$, splitting the interval into $n = 1, 2, 4$ partitions. The results of these calculations are shown in Table 3. The numbers in bold are closest in their respective row to the true value of $\int_0^\pi \sin x \, dx$.

n-value	LRAM	RRAM	SEAM	True Value ‖
1	0	0	$\frac{\pi}{2}^2 \approx \mathbf{2.467}$	
2	$\frac{\pi}{2} \approx 1.5708$	$\frac{\pi}{2} \approx 1.571$	**1.744**	
4	**1.896**	**1.896**	1.611	
True Value				2 ‖

TABLE 3. Approximations of $\int_0^\pi \sin x \, dx$
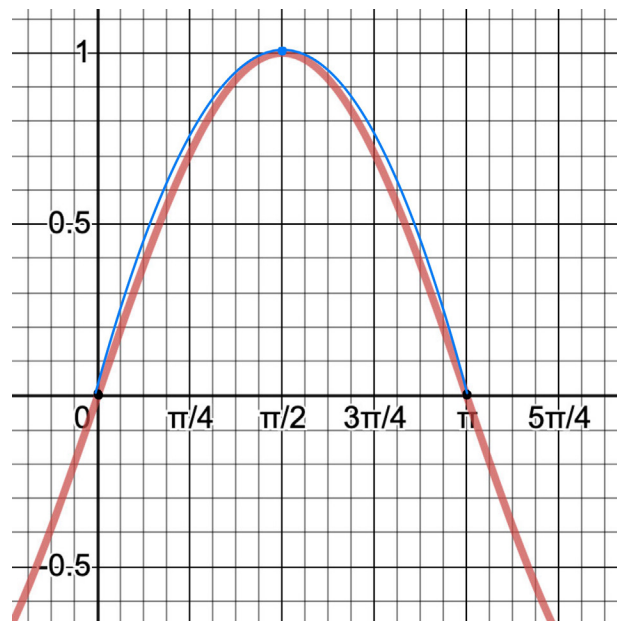
FIGURE 3. Semi-Ellipse Approximation of $\int_0^\pi \sin x \, dx$ with $n = 1$

We presumed our SEAM approximation of $\int_0^\pi \sin x\, dx$ would be exactly the true value since periodic graphs such as $\sin x$ or $\cos x$ appear elliptical in shape. As seen in Table 3, $SEAM_{n=1}$ and the true value of $\int_0^\pi \sin x\, dx$ are not equal. After drawing Figure 3, we could visually see why our hypothesis was incorrect. Although our semi-ellipse and the graph of $\sin x$ from $[0,\pi]$ share the same endpoints and maximum, the curve of the semi-ellipse is slightly broader than the curve of $\sin x$. A future point of study would be to investigate why this is and see if we can manipulate graphs such as $\sin x$ and $\cos x$ to be more elliptical.

Also noteworthy, as with Equation 1, as n increased, the SEAM approximation became less accurate. Utilizing the error bounds for LRAM and RRAM provided from Leclair (2010), we see that with RRAM and LRAM, as n increases, the approximations become more accurate. After comparing the results of Equation 1 to Equation 2, we had hypothesized that our results became less accurate for equation 1 because $f(x) = x$ is linear, whereas $f(x) = x^2$ is nonlinear. However, as we see with our nonlinear graph of $f(x) = \sin x$, the SEAM approximation still became less accurate as n increased, thus proving our hypothesis incorrect. We question why this is since $f(x) = x$ and $f(x) = \sin x$ are very different graphs.

2.4. **Equation 4.** We will be approximating the integral of $f(x) = \ln x$ from $[1, 4]$, splitting the interval into $n = 2, 3, 6$ partitions. The results of these calculations are shown in Table 4. The numbers in bold are closest in their respective row to the true value of $\int_1^4 \ln x \, dx$.

n-value	LRAM	RRAM	SEAM	True Value
2	1.347	3.454	**2.047**	
3	1.792	3.178	**2.022**	
6	2.183	**2.876**	2.005	
True Value				2.545

TABLE 4. Approximations of $\int_1^4 \ln x \, dx$
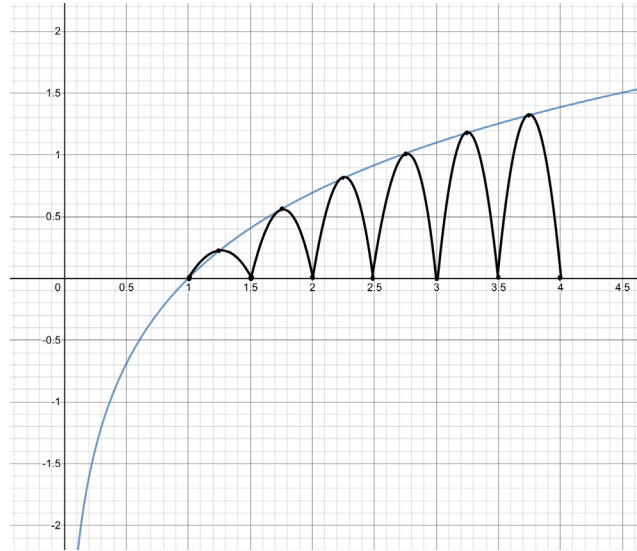
FIGURE 4. Semi-Ellipse Approximation of $\int_1^4 \ln x \, dx$ with $n = 6$

Using SEAM to approximate the integral of $\ln x$ gave us results similar to those in Equations 1 and 3. As n increased, the SEAM approximation became less accurate. We then tried to determine similar characteristics between the graphs of $f(x) = x$, $f(x) = \sin x$, and $f(x) = \ln x$. The most plausible explanation was the concavity of the graphs. SEAM proved to be an excellent approximation for the integral of $f(x) = x^2$, a graph that is concave upward on $[0,2]$. However, SEAM became less accurate in graphs that concave downward, specifically $f(x) = \sin x$, and $f(x) = \ln x$. As for $f(x) = x$, the function can be interpreted as concave upward or concave downward (Pierce 2017). This is a special case because $f(x) = x$ is a straight line. We would like to continue exploring this hypothesis in future research.

3. Conclusion

The Riemann Sum Approximation Methods are useful tools in calculus for approximating the net area under a curve. Within the area under the curve of a graph, important information often relating to probability is found. This is why it is important to make our approximations as accurate as possible. Our mathematical question explored whether we could create a method for a Riemann sum approximation using semi-ellipses that was more precise than Left Rectangular Approximation Method and Right Rectangular Approximation Method. We predicted that similarly to LRAM and RRAM, as the number of shapes n increased, our Semi-Ellipse Approximation Method would become more accurate. We also predicted that for the function $f(x) = \sin x$, the SEAM would be accurate due to the shape of our graph.

In our research in Sections 2.1, 2.3, and 2.4, we found that as n increased, the approximations using SEAM became less accurate. However, in Section 2.2, our SEAM was more accurate than both LRAM and RRAM. Contrary to the other equations, as n increased, SEAM became more accurate. Initially, we believed this was because $f(x) = x^2$ is nonlinear. However, we now believe this was due to the concavity of the functions. Since $f(x) = x^2$ is concave up, SEAM was the most accurate approximation method.

For further research, we would like to compare the accuracy of SEAM with a set of functions concave up on a given interval to a set of functions concave down on a given interval. This will help us determine whether concavity has an influence on the accuracy of the SEAM. In addition to comparing the results to LRAM and RRAM, we would compare them to the Midpoint Rectangular Approximation Method and the Trapezoidal Approximation Method.

References

[1] Alexandre, C., Chebrolu, S., Georgieva, T., Kau A., Khim, J., Kumar, A., Pao, H., Plaza, J., Sharma, G., Sreckovic, M., & Turtle, N. (n.d.). Riemann Sums. Retrieved from https://brilliant.org/wiki/riemann-sums/.

[2] Herman, E. & Strang, G. (2016). *Calculus Volume* I. OpenStax.

[3] Leclair, A. (2010). Error Bound Theorems. Personal Collection of Aaron Leclair, Carnegie Mellon University, Pittsburgh, PA.

[4] Pierce, Rod. (2017). "Concave Upward and Downward". *Math Is Fun*. Retrieved from http://www.mathsisfun.com/calculus/concave-up-down-convex.html.

[5] Stewart, James. (2016). *Calculus: Early Transcendentals (8th ed.)*. Cengage: Boston, MA.

Applying Machine Learning Methods in Solid State Chemistry: Classification and Prediction of Perovskites

Authors: Maria D. Diaz, Brandon Tomich, and Juan Cabrera
Advisors: Dr. Allyson Fry-Petit and Dr. Sam Behseta
Dept. of Chemistry, Dept. of Mathematics, Dept. of Computer Science

Abstract

The mineral family perovskites are compounds with applications in a wide variety of technologies. The rational design of perovskites has the potential to improve the performance of technologies such as MRI machines and solar energy, however, their discovery to date has often been limited to a slow synthesis and testing process mostly based on trial and error. This work employs a series of machine learning techniques aimed at advancing our understanding of the structural forces that drive perovskite formation in hopes that predictive rational design of perovskites can be a reality. The machine learning method used incorporates an unsupervised learning technique from a set of 883 perovskite compounds with 72 predictors, to categorize known materials by structure type. The classification of data is a primary step towards the identification of features that can potentially contribute to building a predictive model for perovskite design, thus, Principal Component Analysis (PCA) was applied to the final space of 67 variables to reveal natural clustering of certain perovskites in the direction of the most significant loadings or eigenvectors of the sample covariance matrix. Specifically, results show that predictors that separate structure type are very sensitive to anion type. This result has the potential to elucidate the currently unknown structural driving forces that can favor perovskite formation with differing anions.

Background

In solid state chemistry, the class of minerals called perovskites express high structural versatility that allow these compounds substantial industry applications. Perovskites belong to a class of minerals that are defined by corner sharing of octahedra (Fig. 1). Double perovskites, the structural type described in the dataset, have the general formula ofwhereandare cations andare anions. Unique to the perovskite structure is the ability to support nearly every element in the periodic table. Additionally, various site occupations, tilts, and lattice directions intensify structural permutability. The importance of perovskite's heightened structural variability arises

from the dependent relationship between physical structure and chemical behavior, that is, in solids it is well established that the chemical properties a compound displays are intimately tied to its underlying crystal structure. Indeed, perovskite behavior varies greatly, with certain distinct permutations owning a considerable amount of utility. Some perovskites possess properties that allow them to find application in memory devices (Kang, et al., 2019), ion conductors (Vicente, et al., 2017), negative thermal expansion materials (Yamada, et al., 2011), and superconductors (Maeno, et al., 1994). The synthesis of commercially appealing perovskites is highly desirable not only for their immediate application, but for the modification of existing materials to suit new applications as well.

While software enabled prediction of some perovskite structures have been observed before synthesis (Ramakrishnan, et al., 2015), others are still elusive and costly in both time and resources. Machine learning offers an approach that mitigates both of these challenges. Recent advancements in machine learning have lent themselves to accurate predictive analysis in quantum chemistry approximations (Hansen, et al., 2015), molecular properties (Coley, et al., 2018), and computer-aided synthesis planning (Lufaso, Woodward, 2001). Similarly, methods employed in this study harnesses machine learning techniques in the accurate prediction of new perovskite materials with structural characteristics that lend themselves to commercial utility. Vast amounts of quality data are required in the use of machine learning algorithms, therefore, a dataset has been compiled from various sources which consists of 2,971 rows of chemical compounds and 109 columns. The ultimate challenge will be identifying which features are most influential in separation based on crystal structure. Therefore, the purpose of this article is to describe the process of feature reduction using Principal Component Analysis, a mathematical and computational approach, in pairing with domain expertise of solid state chemistry.
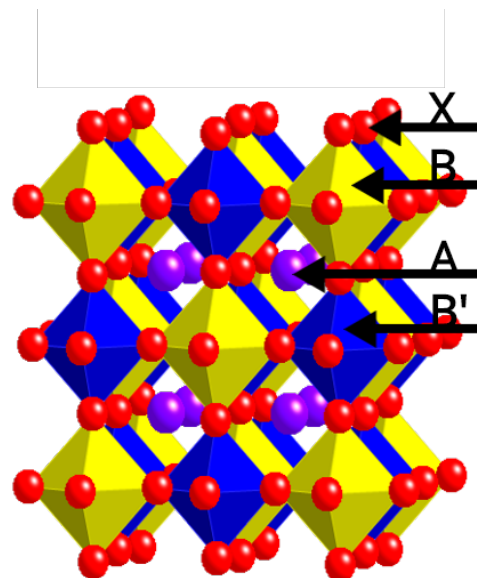
Figure 1: The distribution of A, B, and X atoms within a single cubic perovskite structure where A and B are both cation groups and X is the anion group. The model displays the general size differences among each atom. Size differences among cations and anions within the cubic contribute to the perovskite's chemical properties, physical structure, and behavior.

Data

Preparation of the dataset for use within a machine learning environment entailed comprehensive data cleaning. This feature-oriented dataset is comprised of information retrieved from literature data, experimental results, and existing third party perovskite datasets. Prior to data preprocessing, the dataset was joined together to form more than 2,000 rows of perovskite compounds and 109 columns of feature variable characteristics for each compound respectively. The overall structure of the dataset resembles a simple table. Each row of the dataset constitutes a perovskite compound while each feature is a chemical or physical property of each atom within that row compound entry. The dataset was further organized by way of exploration and categorization of known perovskites possessing structural and chemical properties of interest. N/A values were withdrawn from each column and headers were adapted to match the format requirement for MatLab software. An initial statistical learning algorithm was run on the data and yielded convoluted, insensible results which meant that further data preprocessing needed to be performed. Categorical data was extricated from the set which reduced the row count to roughly 883 compound rows and 67 feature columns (Table 1) and was then exported to Rstudio for principal component analysis.

Table 1: Overview of the 67 features used to represent the dataset

Features of the Perovskite Dataset		
Vc	cation group	Atomic mass of A
Atomic mass of A'	Atomic mass of B	Atomic mass of B'
AREN of A	AREN of A'	AREN of B
AREN of B'	R_XIIofA	R_XIIofA'
r_VIofB'	r_VIofB	r_VIIofB
Crystal Radius of A	Crystal Radius of A'	Crystal Radius of B'
Crystal Radius of B	Ionic Radii of B	Ionic Radii of B'
Ionic Radii of A	Ionic Radii of A'	Ionization Potential of B
Ionization Potential of B'	Ionization Potential of A	Ionization Potential of A'
Pauling Energy Number of B	Pauling Energy Number of B'	Pauling Energy Number of A
Pauling Energy Number of A'	Zunger Radii of B	Zunger Radii of B'
Zunger Radii of A	Zunger Radii of A'	Boiling Point of A
Boiling Point of A'	Boiling Point of B	Boiling Point of B'
Coefficient of Linear Thermal Expansion at Specific Heat of A	Coefficient of Linear Thermal Expansion at Specific Heat of A'	Coefficient of Linear Thermal Expansion at Specific Heat of B
Coefficient of Linear Thermal Expansion at Spec of B'	Density at Specific Heat of A	Density at Specific Heat of A'
Density at Specific Heat of B	Density at Specific Heat of B'	Enthalpy of Fusion of A
Enthalpy of Fusion of A	Enthalpy of Fusion of A prime	Enthalpy of Fusion of B
Enthalpy of Fusion of B'	Melting Point of A	Melting Point of A'
Melting Point of B	Melting Point of B'	Specific Heat Capacity of A
Specific Heat Capacity of A'	Specific Heat Capacity of B	Specific Heat Capacity of B'
Thermal Heat Conductivity of A	Thermal Heat Conductivity of A'	Thermal Heat Conductivity of B
Thermal Heat Conductivity of B'	Radius of A	Radius of B
Tolerance BVP		

Chemistry Methods

Solid state chemists used the intrinsic relationship between structural and chemical properties to identify perovskite features that would be the most influential in the prediction of commercially viable perovskites, meaning feature selection served as the first data preprocessing step prior to the machine learning assisted dimensionality reduction method described later on. It was determined that additional information related to physical features should be appended to the dataset to increase the predictors. To this end, a lookup script was employed to add information to each row given that a specified atom substring matched within the compound data entry.

Machine Learning Methods

Principal component analysis was used in conjunction with unsupervised learning to project for viable perovskites. Principal component analysis allowed us to reduce the entire feature space of the data set onto its most valuable features.

PCA is a linear transformation, facilitating the creation of new features, as a weighted sum of original features, that have two, important properties: 1. They are independent, and 2. Their variance shrinks accordingly. This process then makes it possible to identify and select the most prominent features of data, contributing to the response.

To elaborate, suppose the covariance matrix associated with the feature space of the Perovskites' population is represented by Σ. Let $\Sigma = Q\Lambda Q^{-1}$ represent the eigen-decomposition of Σ, so that Λ is the diagonal matrix of its eigenvalues, and Q is a matrix whose columns are eigenvectors of Λ. Let q_i represent the ith column of Q. The ith principal component of the Perovskites covariance matrix, can be then represented as $q_i^T = (q_{1i}, \ldots, q_{di})^T$, where d is the dimensionality of the feature space. Since the variance of the ith principal component is given by the ith diagonal component of Λ, we can then reduce the dimensionality of the feature space by first, selecting a level, say $(1 - \alpha) \leq 1$, followed by, reducing the space to the integer $k \ll d$, so that the ratio $\frac{\lambda_k}{\sum_{i=1}^{d} \lambda_j}$ is less than or equal to $1 - \alpha$.

Furthermore, suppose $X_{\sim}^T = (x_{1,\ldots,} x_p)$ be a column vector of our data set. We will then calculate the covariance matrix via the dot product $X \cdot X^T$. If the covariance is close to zero, then the dimensions are highly correlated. Next, we calculate the eigenvectors and eigenvalues of the covariance matrix. The eigenvalues of the covariance matrix are the so-called loadings of the PCA.

$$Z_1 = X_1 P_1 + X_2 P_2 + \ldots + X_p P_p$$

We can also observe this from a different angle if each row of the data set matrix is written as a linear combination. Suppose that we represent the first PC via the following $Z_1 = X_1 P_1 + X_2 P_2 + \ldots + X_p P_p$. The coefficients associated with $x_1, x_2, \ldots, x_p$ are the loadings of the first principal component. We can normalize the feature space so that each feature has a mean of zero and moreover we can control the sum squared of the terms in the first PC in order to impose an upper bound on the variance of each PC.

The loadings of the vector $X_{\sim}$ define a direction in the PC space through which the most of the total variance is expressed. The second principal component is calculated with the caveat that it is the linear combination of our features $x_1, x_2, \ldots, x_p$ uncorrelated with the first principal component.

To implement the principal component analysis on our high dimensional data, we utilized the programming language R and selected packages.

Preliminary Results

The following plots were generated using the "stats" library in Rstudio:

Figure 2: A biplot for the perovskite population. In this plot, the features in the dataset are represented as dots and the loadings are the vectors. The loadings indicate not only the direction, but also the weight of each linear transformation. Therefore, we can detect natural clusters by focusing locally on each direction, followed by identifying perovskites that are in the same direction.

The loadings of the biplot (Fig. 2) distinguish features that affect the total variation in the perovskites population. Notable among these are X-site features, AREN of A, and coefficient of linear thermal expansion of A. Loading vectors that overlap with each other indicate strong similarity within the PC space. Features exhibiting strong similarity to each other can be extricated from the dataset on the principle that most of the information that can be gathered from one variable can be represented by the overlapping similar variable. The size of the vector, or its magnitude, is interpreted as the level of importance the variable has to the PC space. A vector with a high magnitude captures much of the information in the PC space and thus defines itself as a feature which comprehensively represents the maximum variation in the data. Those features will be used as the test set for the machine learning algorithm. Likewise, features with low

magnitude can be liberated from the dataset, as they do not influence the variation in the dataset. The dots behind the loadings on the biplot are the score plot which describes clusters and shape of the data.

Loading plot (Fig. 2) shows how definitively each feature influences a principal component. Vectors that are close together mean positive correlation while those that meet at 90° have features that are likely not correlated. Vectors that diverge at 180° have features that are negatively correlated. Loadings that have high magnitude are said to be prominent features of the PC space.
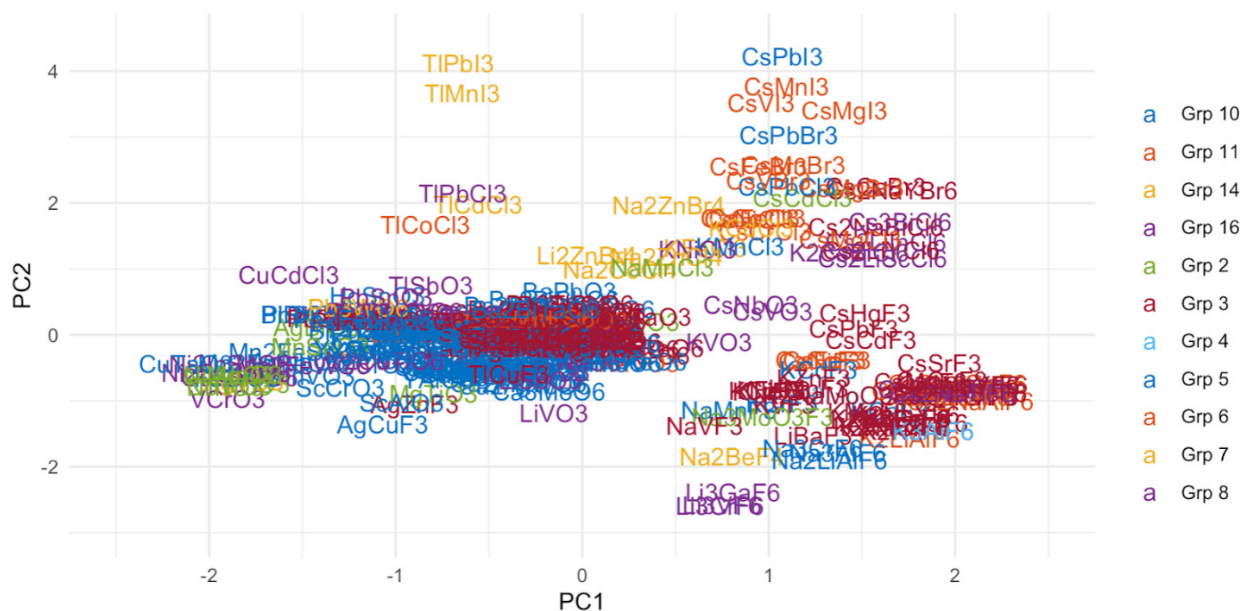
Figure 3: All compounds plotted with respect to first and second principal components. Color differences added to increase visibility of various compounds.

Figure 4: Similar to Figure 3, compounds are plotted with respect to first and second principal components. Two colors are used to highlight clustering of anion compounds.

From the two plots above (Fig. 3, 4), the prominence of X-site features in conjunction with the natural anion clustering exhibited in the PC score plots suggest false positivity of X-site feature influence. This is illustrated by the various compounds represented in the PC space clustering together with other like anion groups. This demonstrates a sensitivity to anion type which will need to be addressed via an amendment to the design rules.

Summary

After the initial PCA trial, X and X' site features were shown to be a falsely diverting influence in the PC space, additionally, properties that showed the most similarity were atomic mass, atomic radius, Pauling energy number, ionization potential, and AREN site and site prime properties, however, to accurately determine if these properties can be condensed, X-site false positivity will need to be addressed. Viewing the plots above (Fig. 3, 4), it is shown that oxygen anion compounds are clustered together over the PC space, but an empirical understanding of solid-state chemistry will need to be used to interpret results and inform future direction.

Future Work

The set of analyses included in this paper may be taken as initial steps towards identification and categorization of perovskites. However, further modeling is needed to tease out the features that directly contribute to a potential response variable. For example, as an immediate future step, the authors are planning to employ a group of generalized linear models to not only gauge the effect of select features but their direction on a response variable which is a natural grouping of the perovskites from the feature space. Moreover, compounds will be divided into their own anion groups and PCA will be run yet again in order to create more precise natural clusters.

References
1. Kang, Keehoon, et al. "Perovskite Memory Devices: High-Performance Solution-Processed Organo-Metal Halide Perovskite Unipolar Resistive Memory Devices in a Cross-Bar Array Structure (Adv. Mater. 21/2019)." *Wiley Online Library*, John Wiley & Sons, Ltd, 21 May 2019, onlinelibrary.wiley.com/doi/abs/10.1002/adma.201970149.

2. Vicente, Nuria, and Germà Garcia-Belmonte. "Organohalide Perovskites Are Fast Ionic Conductors." *Wiley Online Library*, John Wiley & Sons, Ltd, 3 July 2017, onlinelibrary.wiley.com/doi/abs/10.1002/aenm.201700710.

3. Yamada, Ikuya, et al. "Giant Negative Thermal Expansion in the Iron Perovskite $SrCu_3Fe_4O_{12}$." *Wiley Online Library*, John Wiley & Sons, Ltd, 6 June 2011, onlinelibrary.wiley.com/doi/full/10.1002/anie.201102228.

4. Maeno, Y., et al. "Superconductivity in a Layered Perovskite without Copper." *Nature News*, Nature Publishing Group, 1994, www.nature.com/articles/372532a0.

5. Ramakrishnan, Raghunathan, et al. "Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach." *Journal of Chemical Theory and Computation*, vol. 11, no. 5, 2015, pp. 2087–2096., doi:10.1021/acs.jctc.5b00099.

7. Coley, Connor W., et al. "Machine Learning in Computer-Aided Synthesis Planning." *Accounts of Chemical Research*, vol. 51, no. 5, 2018, pp. 1281–1289., doi:10.1021/acs.accounts.8b00087.

8. Lufaso, Michael W., and Patrick M. Woodward. "Prediction of the Crystal Structures of Perovskites Using the Software ProgramSPuDS." *Acta Crystallographica Section B Structural Science*, vol. 57, no. 6, 2001, pp. 725–738., doi:10.1107/s0108768101015282.

Acknowledgements

Fibonacci Gaps

Author: Madeline Ceccia
Advisor: Dr. Martin Bonsangue
California State University, Fullerton
Department of Mathematics

Abstract. The Fibonacci number sequence has been a source of fascination for both mathematicians and lay persons for hundreds of years. In this paper we look not at the Fibonacci numbers, but the numerical "negative space," or gaps, between them. These Fibonacci Gaps are based on the number of natural numbers between two successive Fibonacci numbers. We prove the relationship between the Fibonacci sequence and their related Fibonacci Gaps, and explore other mathematical relationships found in these gaps.

Introduction. The Fibonacci number sequence has been a source of fascination for both mathematicians and lay persons alike. Works have ranged from applications of Fibonacci numbers in stock market trading (Boroden, 2008) to exploratory activities for young children (Garland 1990). Mathematical interests include the relationship between Fibonacci numbers and the golden ratio, as well as other more advanced explorations commonly taught at the undergraduate level (e.g., Posamentier & Lehmann, 2007; Vorob'ev, 2011). The proofs and observations presented in this paper are built on the Fibonacci sequence, which further provides students with reinforcement of the connecting nature of mathematics, especially those in proof courses. Perhaps one of best-known (American) references of Fibonacci numbers and the golden ratio is from Walt Disney's classic 1959 cartoon, *Donald in Mathmagic Land*. All of these applications and explorations are based on the Fibonacci number sequence 1, 1, 2, 3, 5, 8, 13, ...

Number Gaps. The difference of two numbers is typically thought of as the continuous "distance" between them on a number line. For example, any eight-year old boy knows that his five-year old sister is three years younger than he is (especially if they were born in the same month). However, there are only two "ages" that are separating the children, namely, age 6 and age 7. Thus, the "age gap" between the children is actually 3 – 1, or 2 (Figure 1).
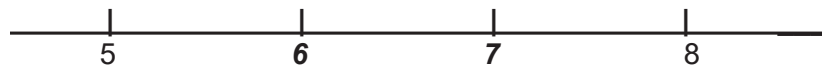
Figure 1: The "Age Gap" Number Line

If we think of the number line as a continuum that includes non-integer values, then, indeed, we would say that the boy is (about) three years older than his little sister. However, in keeping with the discrete nature of the Fibonacci numbers (think bunnies!), we will choose to focus here on the age gap viewpoint. It is in this spirit that we define the *number gap* between two natural numbers, a and b, where $a < b$ and $b - a > 1$, to be the number of natural numbers that are greater than a and less than b. Thus, the number gap between a and b is $b - a - 1$.

Fibonacci Gaps. The Fibonacci sequence is

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

Using the standard notation for the Fibonacci numbers, $f_1 = 1$, $f_2 = 1$, $f_3 = 2$, $f_4 = 3$, $f_5 = 5$, $f_6 = 8$, ..., where

$$f_{k+2} = f_{k+1} + f_k \tag{1}$$

for $k = 1, 2, 3, \ldots$

The first non-zero number gap occurs between the successive Fibonacci numbers f_4 and f_5. From Figure 2 we can see that the first five Fibonacci Gap numbers (in pink) are $g_1 = 1$, $g_2 = 2$, $g_3 = 4$, $g_4 = 7$, and $g_5 = 12$.

Figure 2: Counting Fibonacci Gap Numbers Between Fibonacci Numbers

If we denote the Fibonacci Gap numbers by g_1, g_2, g_3, etc., we can compare the Fibonacci numbers with the two Fibonacci sequence numbers which reside to the left and right of each Fibonacci Gap number. We will refer to these residing Fibonacci numbers as Fibonacci "border" numbers. For example, the Fibonacci Gap number 1 is associated with the Fibonacci numbers 3 and 5. Similarly, the Fibonacci Gap number 2 is associated with the Fibonacci numbers 5 and 8. (Table 1). Please note that, again, one could interpret the gap between two Fibonacci numbers to be the difference of those numbers rather than the number of integers between them. In this case, the new sequence is, by definition, the original Fibonacci sequence.

Table 1: Fibonacci Numbers and Their Associated Fibonacci Gap Numbers																	
Fib. Nos.	f_1	f_2	f_3	f_4		f_5		f_6		f_7		f_8		f_9		f_{10}		f_{11}
	1	1	2	3		5		8		13		21		34		55		89
Fib. Gaps					1		2		4		7		12		20		33	
					g_1		g_2		g_3		g_4		g_5		g_6		g_7	

Two observations can be quickly gleaned from Table 1. As noted earlier, each Fibonacci Gap number is one less that the difference of its two bordering Fibonacci numbers. For example, $20 = 55 - 34 - 1$, that is, $g_6 = f_{10} - f_9 - 1$. In general, the kth Fibonacci Gap number, g_k, is given by

$$g_k = f_{k+4} - f_{k+3} - 1 \tag{2}$$

for k = 1, 2, 3, ... In addition, since $f_{k+4} = f_{k+2} + f_{k+3}$, we have that

$$g_k = f_{k+4} - f_{k+3} - 1 = (f_{k+2} + f_{k+3}) - f_{k+3} - 1 = f_{k+2} - 1.$$

Thus, for k = 1, 2, 3, ... ,

$$g_k = f_{k+2} - 1. \qquad\qquad (3a)$$

For example, $g_4 = 7 = 8 - 1 = f_6 - 1$.

Three Fibonacci Gap Number Relationships. If we consider the set of Fibonacci Gap numbers to be a sequence it is own right, we can readily observe three basic relationships between Fibonacci numbers f_k and Fibonacci Gap numbers g_k (Table 2).

Table 2: The Sequence of Fibonacci Numbers and the Sequence of Fibonacci Gap Numbers														
k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
f_k	1	1	2	3	5	8	13	21	34	55	89	144	233	377	610
g_k	1	2	4	7	12	20	33	54	88	143	232	376	609	986	1596

First, as noted above,

$$g_k = f_{k+2} - 1 \text{ or} \qquad\qquad (3a)$$
$$f_{k+2} = g_k + 1. \qquad\qquad (3b)$$

Second, since $f_{k+4} = f_{k+3} + f_{k+2}$, we adjust subscripts in equation (3b) to obtain
$g_{k+2} + 1 = (g_{k+1} + 1) + (g_k + 1)$, or

$$g_{k+2} = g_{k+1} + g_k + 1. \qquad\qquad (4)$$

That is, each Fibonacci Gap number is one more than the sum of the two previous Fibonacci Gap numbers. And third, since $g_k = g_{k-1} + g_{k-2} + 1$ and $g_{k-2} + 1 = f_k$, we have

$$g_k = g_{k-1} + f_k. \qquad\qquad (5)$$

Equation (5) gives a quick way of computing Fibonacci Gap numbers using a "diagonal sum" approach in Table 2. For example, starting with $g_3 = 4$, we have $4 + 3 = 7$; $7 + 5 = 12$; $12 + 8 = 20$; and etc.

Other relationships between the Fibonacci numbers and the Fibonacci gap numbers can be seen as well. For example, a well-known identity for the sum of the first k Fibonacci numbers is

$$f_1 + f_2 + \ldots + f_k = f_{k+2} - 1.$$

From equation 3a, $g_k = f_{k+2} - 1$. Thus, the kth Fibonacci gap number is the sum of the first k Fibonacci numbers:

$$f_1 + f_2 + \ldots + f_k = g_k.$$

A Fibonacci Gap Surprise. One well-known result of Fibonacci numbers is that the difference of the square of a Fibonacci number and the product of its two bordering Fibonacci numbers is –1 or 1, depending on whether the index of the number is odd or even (e.g., Posametier & Lehmann, 2008). For example, for $k = 5$ we have

$$f_6^2 - f_5 f_7 = 8^2 - 5 \times 13 = 64 - 65 = -1,$$

while, for $k = 6$,

$$f_7^2 - f_6 f_8 = 13^2 - 8 \times 21 = 169 - 168 = 1.$$

In general,

$$f_{k+1}^2 - f_k f_{k+2} = -1 \text{ if } k \text{ is odd, and} \qquad\qquad (6a)$$
$$f_{k+1}^2 - f_k f_{k+2} = 1 \text{ if } k \text{ is even.} \qquad\qquad (6b)$$

While applying this relationship directly to the sequence of Fibonacci Gap numbers did not seem to yield a similar result, we observed that by rewriting equations (6a) and (6b) we could translate the result to the sequence g_k. By adding in and subtracting out the product $f_k f_{k+1}$ into the left hand side of equations (6a) or (6b), we have

$$f_{k+1}^2 - f_k f_{k+2} \quad = f_{k+1}^2 + f_k f_{k+1} - f_k f_{k+1} - f_k f_{k+2}$$

$$= f_{k+1} (f_{k+1} + f_k) - f_k (f_{k+1} + f_{k+2})$$

$$= f_{k+1} f_{k+2} - f_k f_{k+3} .$$

Thus, equations (6a) and (6b) can be rewritten as

$$f_{k+1} f_{k+2} - f_k f_{k+3} = -1 \text{ if } k \text{ is odd, and} \qquad\qquad (7a)$$
$$f_{k+1} f_{k+2} - f_k f_{k+3} = 1 \text{ if } k \text{ is even.} \qquad\qquad (7b)$$

In the example above for $k = 6$, equation (7b) gives

$$f_7 f_8 - f_6 f_9 = 13 \times 21 - 8 \times 34 = 273 - 272 = 1.$$

Equations (7a) and (7b) produce an alternating sequence $\{s_k\}$ of –1's and 1's:

$$\{s_k\} = \{-1, 1, -1, 1, -1, 1, -1, 1, \ldots\}. \qquad\qquad (8)$$

What would happen if we applied the left-hand side of equations (7a or 7b) directly to the Fibonacci Gap numbers, that is, if we computed

$$g_{k+1}\, g_{k+2} - g_k\, g_{k+3}$$

for different values of k? For example, if $k = 1$, we have

$$g_2\, g_3 - g_1\, g_4 = 2 \times 4 - 1 \times 7 = 8 - 7 = 1.$$

For $k = 2$,

$$g_3\, g_4 - g_2\, g_5 = 4 \times 7 - 2 \times 12 = 28 - 24 = 4.$$

Surprisingly, though, for $k = 3$ we get the same result:

$$g_4\, g_5 - g_3\, g_6 = 7 \times 12 - 4 \times 20 = 84 - 80 = 4.$$

These computations produce the sequence $\{t_k\}$:

$$\{t_k\} = \{1,\, 4,\, 4,\, 9,\, 12,\, 22,\, 33,\, 56,\, ...\}. \qquad\qquad (9)$$

Was there a predictable pattern for the sequence $\{t_k\}$ and, perhaps more importantly, were the numbers in the Fibonacci number-based sequence $\{s_k\}$ related to those in the Fibonacci Gap-based sequence $\{t_k\}$ and, if so, how?

The Mystery Revealed. Equations (7a) and (7b) state that

$$f_{k+1}\, f_{k+2} - f_k\, f_{k+3} = -1 \text{ if } k \text{ is odd, and} \qquad\qquad (7a)$$
$$f_{k+1}\, f_{k+2} - f_k\, f_{k+3} = 1 \text{ if } k \text{ is even.} \qquad (7b)$$

Using the relationship between Fibonacci numbers and Fibonacci Gap numbers, $f_{k+2} = g_k + 1$ from equation (3b), equations (7a) and (7b) can be rewritten as

$$(g_{k-1} + 1)(g_k + 1) - (g_{k-2} + 1)(g_{k+1} + 1) = -1 \text{ if } k \text{ is odd, and} \qquad (10a)$$
$$(g_{k-1} + 1)(g_k + 1) - (g_{k-2} + 1)(g_{k+1} + 1) = 1 \text{ if } k \text{ is even.} \qquad (10b)$$

We consider each case separately.

Case 1: k is odd.   Removing parentheses in equation (10a), we have

$$g_{k-1}\, g_k + g_{k-1} + g_k + 1 - g_{k-2}\, g_{k+1} - g_{k-2} - g_{k+1} - 1 = -1.$$

Since $g_{k-1} + g_k + 1 = g_{k+1}$, we have

$$g_{k-1}\, g_k + g_{k+1} - g_{k-2}\, g_{k+1} - g_{k-2} - g_{k+1} - 1 = -1,$$

or

$$g_{k-1}\, g_k - g_{k-2}\, g_{k+1} - g_{k-2} = 0.$$

Rewriting the indices in the above equation based on the kth Fibonacci Gap number, we have

$$g_{k+1}\, g_{k+2} - g_k\, g_{k+3} = g_k. \tag{11a}$$

Equation (11a) is the Fibonacci Gap counterpart of equation (7a) for Fibonacci numbers, in which $t_k = g_{k+1}\, g_{k+2} - g_k\, g_{k+3}$ is not -1, but g_k. For example, for $k = 5$ and $g_5 = 12$, we have

$$t_5 = g_6 g_7 - g_5\, g_8 = 20 \times 33 - 12 \times 54 = 660 - 648 = 12 = g_5.$$

Table 3a shows (in bold) this result for odd values of k.

Table 3a: The Sequence of Fibonacci Gap Numbers and the Sequence of $\{t_k\}$ Numbers for Odd Values of k														
k	**1**	2	**3**	4	**5**	6	**7**	8	**9**	10	**11**	12	**13**	14	**15**
g_k	**1**	2	**4**	7	**12**	20	**33**	54	**88**	143	**232**	376	**609**	986	**1596**
t_k	**1**	4	**4**	9	**12**	22	**33**	56	**88**	145	**232**	378	**609**	988	**1596**

Case 2: k is even.   As before, removing parentheses in equation (10b), we have

$$g_{k-1}\, g_k + g_{k-1} + g_k + 1 - g_{k-2}\, g_{k+1} - g_{k-2} - g_{k+1} - 1 = 1.$$

Since $g_{k-1} + g_k + 1 = g_{k+1}$, we have

$$g_{k-1}\, g_k + g_{k+1} - g_{k-2}\, g_{k+1} - g_{k-2} - g_{k+1} - 1 = 1,$$

or

$$g_{k-1}\, g_k - g_{k-2}\, g_{k+1} - g_{k-2} = 2.$$

Rewriting the indices in the above equation based on the kth Fibonacci Gap number, we have

$$g_{k+1}\, g_{k+2} - g_k\, g_{k+3} = g_k + 2. \tag{11b}$$

Equation (11b) is the Fibonacci Gap counterpart of equation (7b) for Fibonacci numbers, in which $t_k = g_{k+1}\, g_{k+2} - g_k\, g_{k+3}$ is not 1, but $g_k + 2$. For example, for $k = 6$ and $g_6 = 20$, we have

$$t_6 = g_7 g_8 - g_6\, g_9 = 33 \times 54 - 20 \times 88 = 1782 - 1760 = 22 = 20 + 2 = g_6 + 2.$$

Table 3b shows (in bold) this result for even values of k.

Table 3b: The Sequence of Fibonacci Gap Numbers and the Sequence of $\{t_k\}$ Numbers for Even Values of k														
k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
g_k	1	2	4	7	12	20	33	54	88	143	232	376	609	986	1596
t_k	1	4	4	9	12	22	33	56	88	145	232	378	609	988	1596

Combining both even and odd cases for k, we get the sequence

$$\{t_k\} = \{1, 4, 4, 9, 12, 22, 33, 56, 88, 145, 232, 378...\}, \tag{9}$$

where $t_k = g_k$ for odd values of k and $t_k = g_k + 2$ for even values of k. The paired equations are:

$$g_{k+1}\, g_{k+2} - g_k\, g_{k+3} = g_k \text{ if } k \text{ is odd, and} \tag{11a}$$
$$g_{k+1}\, g_{k+2} - g_k\, g_{k+3} = g_k + 2 \text{ if } k \text{ is even.} \tag{11b}$$

An Unexpected Connection. The terms in the sequence

$$\{s_k\} = \{-1, 1, -1, 1, -1, 1, -1, 1, ...\} \tag{8}$$

can be expressed recursively as

$$s_{k+2} = s_{k+1} + s_k - 1 \text{ if } k \text{ is odd, and} \tag{12a}$$
$$s_{k+2} = s_{k+1} + s_k + 1 \text{ if } k \text{ is even.} \tag{12b}$$

What would the terms in the sequence $\{t_k\}$ look like if written recursively? Using the basic relation

$$g_{k+2} = g_{k+1} + g_k + 1, \tag{4}$$

we can easily consider the cases for when k is odd and when k is even. If k is odd, then $k + 2$ is odd but $k + 1$ is even. Thus, $t_k = g_k$, $t_{k+2} = g_{k+2}$, while $t_{k+1} = g_{k+1} + 2$, or $g_{k+1} = t_{k+1} - 2$ from equations (11a) and (11b). Substituting into equation (4) gives

$$t_{k+2} = (t_{k+1} - 2) + t_k + 1,$$

or

$$t_{k+2} = t_{k+1} + t_k - 1. \tag{13a}$$

If k is even, then $k + 2$ is even but $k + 1$ is odd. Thus, $t_k = g_k + 2$, $t_{k+2} = g_{k+2} + 2$, while $t_{k+1} = g_{k+1}$.

As before, substituting into equation (4) gives

$$t_{k+2} - 2 = t_{k+1} + t_k - 2 + 1,$$

or

$$t_{k+2} = t_{k+1} + t_k + 1. \qquad\qquad (13a)$$

Equations (13a) and (13b) are identical analogs to equations (12a) and (12b):

$$t_{k+2} = t_k + t_{k+1} - 1 \text{ if } k \text{ is odd, and} \qquad (13a)$$
$$t_{k+2} = t_k + t_{k+1} + 1 \text{ if } k \text{ is even.} \qquad (13b)$$

Thus, the recursive property for the Fibonacci Gap numbers accompanying sequence$\{t_k\}$(shown in equations 13a and 13b) is exactly the same as that for the Fibonacci numbers accompanying sequence$\{s_k\}$ (shown in equations 12a and 12b)!

Fibonacci Gaps and the Golden Ratio. The Fibonacci sequence is perhaps best-known for its link to the "Golden Ratio." The ratios of consecutive Fibonacci numbers, $f_{k+1} : f_k$, produce an alternating sequence of numbers approaching

What about the Fibonacci Gaps? Do they approach the golden ratio as well? Taking the ratios of successive Fibonacci Gap numbers, $g_{k+1} : g_k$, we get the sequence (to three decimal places)

$$2, 2, 1.75, 1.714, 1.666, 1.65, 1.636, 1.629, 1.625, 1.622, \qquad (14)$$

We can quickly observe two things about the sequence in (14). First, the ratios seem to be approaching some limit (perhaps phi). And second, unlike the Fibonacci number ratios which fluctuate, the Fibonacci Gap number ratios seem to be strictly decreasing. Table 4 compares the Fibonacci Gap number ratios with the Fibonacci number ratios.

Table 4: Fibonacci Gap Numbers and the Golden Ratio					
	Fibonacci Numbers		Fibonacci Gap Numbers		Difference
k	f_k	$f_{k+1}: f_k$	g_k	$g_{k+1}: g_k$	$(g_{k+1}: g_k) -$ $(f_{k+1}: f_k)$
1	1	1	1	2	1
2	1	2	2	2	0
3	2	1.5	4	1.75	0.25
4	3	1.666666667	7	1.714285714	0.047619048
5	5	1.6	12	1.666666667	0.066666667
6	8	1.625	20	1.65	0.025
7	13	1.615384615	33	1.636363636	0.020979021
8	21	1.619047619	54	1.62962963	0.010582011
9	34	1.617647059	88	1.625	0.007352941
10	55	1.618181818	143	1.622377622	0.004195804
11	89	1.617977528	232	1.620689655	0.002712127
12	144	1.618055556	376	1.619680851	0.001625296
13	233	1.618025751	609	1.619047619	0.001021868
14	377	1.618037135	986	1.618661258	0.000624122
15	610	1.618032787	1596	1.618421053	0.000388266
16	987	1.618034448	2583	1.618273326	0.000238878
17	1597	1.618033813	4180	1.618181818	0.000148005
18	2584	1.618034056	6764	1.618125370	0.000091313
19	4181	1.618033963	10945	1.618090452	0.000056489
20	6765	1.618033999	17710	1.618068888	0.000034889

The Fibonacci ratios seem to trail behind the Fibonacci Gap ratios. However, the difference between the two ratios, as seen in Table 4, decreases as we get further in the respective sequences, suggesting that as k gets large, the ratio of g_{k+1} to g_k approaches phi. Since $g_k = f_{k+2}$ – 1 (equation 3a), we have

As k approaches infinity, approaches zero and approaches phi. Thus, the ratio of successive terms of Fibonacci Gap numbers is, indeed, the Golden Ratio.

Conclusion. In this brief journey we have explored the number gaps between Fibonacci numbers. Fibonacci numbers and Fibonacci Gap numbers are closely related (Fibonacci Gap numbers essentially trail Fibonacci numbers by one), so that the ratio of successive terms in each sequence is phi, the Golden Ratio. However, we observed some differences as well, especially in looking at the difference of the product of pairs of four successive Fibonacci and Fibonacci Gap numbers, $f_{k+1} f_{k+2} - f_k f_{k+3}$ and $g_{k+1} g_{k+2} - g_k g_{k+3}$. However, each sequence of these differences, s_k and t_k, can be written recursively in exactly the same way.

Table 5 gives seven identities that are known for Fibonacci numbers (Sundstrom, p. 207), with Identity 6 already extended to Fibonacci Gap numbers in this paper.

Table 5: Seven Fibonacci Number Identities		
Identity	Fibonacci Identity	Fibonacci Gap Identity
1	f_{4k} is a multiple of 3	?
2	f_{5k} is a multiple of 5	?
3	$f_1 + f_2 + \ldots + f_{k-1} = f_{k+1} - 1$	?
4	$f_1 + f_3 + \ldots + f_{2k-1} = f_{2k}$	?
5	$f_2 + f_4 + \ldots + f_{2k} = f_{2k+1} - 1$	?
6	$f_{k+1}\,f_{k+2} - f_k\,f_{k+3} = -1$ if k is odd; $f_{k+1}\,f_{k+2} - f_k\,f_{k+3} = 1$ if k is even	$g_{k+1}\,g_{k+2} - g_k\,g_{k+3} = g_k$ if k is odd; $g_{k+1}\,g_{k+2} - g_k\,g_{k+3} = g_k + 2$ if k is even
7	$f_1^2 + f_2^2 + \ldots + f_k^2 = f_k\,f_{k+1}$	?

The question marks in Table 5 are open questions in which the interested reader can explore further and perhaps draw her or his own connection between the Fibonacci Identity and a possible Fibonacci Gap Identity. We invite the interested reader to play around with these and see if there is a Fibonacci Gap extension to be discovered for any of the other identities as well.

References

Boroden, C. (2008). *Fibonacci trading: How to master the time and price advantage*. McGraw-Hill: New York.

Garland, T. H. (1990). *Fascinating Fibonaccis*. Dale Seymour Publications: Reston, VA.

Posamentier, A., and Lehmann, I. (2007). *The fabulous Fibonacci numbers.* Prometheus Books: Amherst, New York.

Sundstrom, T. (2013). *Mathematical reasoning: Writing and proof, version 2.1*. www.tedsundtrom.com: Allendale, MI.

Vorob'ev, N. (2011). *Fibonacci numbers*. Dover Publications: Mineola, New York.

Minimal Generating Sets of the Monoid of Partial Order-preserving Injections

Author: Saul Lopez
Advisor: Scott Annin, Ph.D.
California State University, Fullerton
Department of Mathematics

Abstract

Monoids arise in such fields as computer science, physics, and numerous branches of mathematics including abstract algebra, cryptography and operator theory. In this research project we seek to determine minimal generating sets for the monoid of partial order-preserving injections of an n-element set, P OI(n). A generating set for a monoid is a collection of elements S such that every element of the monoid can be expressed as a product of elements from S. Generating sets are of fundamental importance across math and science, and mathematicians have great interest in studying generating sets of a variety of algebraic structures. By a minimal generating set, we refer to a generating set for which no proper subset is a generating set. In this paper, we provide necessary and sufficient conditions for a set to be a minimal generating set for P OI(n), and we show that there are exactly (n − 1)! minimal generating sets for *POI*(n).

Keywords 1. Monoid; Partial order-preserving injection; rank; minimal generating set; partial identity; Green's relations; directed graph; walk; path; cycle

1   Introduction

Partial order-preserving injections comprise an important subset of the collection of all partial injections from a set to itself. This latter collection forms a monoid (defined below) into which any inverse monoid can be embedded [5]. Embeddings preserve algebraic structure, so that the collection of all partial injections on a set becomes a central object of study in the theory of inverse monoids. In this paper, we will be specifically studying the monoid consisting of partial order-preserving injections. These order-preserving maps often arise in the context of studying transformations of partially ordered sets. With this motivational background, let us turn to a brief review of basic concepts in the theory of monoids.

Definition 1. A **monoid** is a non-empty set M together with a binary operation $*$ that satisfies the following axioms:

1) Closure: If $a, b \in M$, then $a * b \in M$.

2) Associativity: $a * (b * c) = (a * b) * c$, for all $a, b,$ and $c \in M$.

3) Existence of an Identity: An element $e \in M$ is called the **identity element** if $a*e = a = e*a$ for every $a \in M$.

Example 1. The set of positive real numbers under multiplication, $(\mathbb{R}^+, \times)$, is a classic example of a monoid, with identity element 1.

Example 2. Note $(\mathbb{R}^+, +)$ is not a monoid since 0 does not belong to $\mathbb{R}^+$. However, $([0, \infty), +)$ is a monoid with identity element 0.

Note that the monoid in Example 2 forms a **submonoid** of $(\mathbb{R}, +)$, since it is a subset of $\mathbb{R}$ that forms a monoid under $+$.

Definition 2. A **semigroup** is a non-empty set S that satisfies the closure and associativity axioms in Definition 1. Thus, every monoid is a semigroup.

Any semigroup can be made into a monoid by adjoining an identity element. In Example 2, for instance, starting with the semigroup $(\mathbb{R}^+, +)$, we form the monoid $([0, \infty), +) = (\mathbb{R}^+ \cup \{0\}, +)$.

Example 3. Let X be a non-empty set. An expression of the form $x_1 x_2 x_3 \cdots x_n$ (with $x_i \in X$) is called a **word** on the alphabet X. The set of all words on X is the **free semigroup** on X, denoted X^+. We can form a monoid by adjoining an empty word, say Λ, to X^+. This monoid, denoted X^*, is called the **free monoid** on X.

The operation in X^+ and X^* is concatenation. If we let X be the English alphabet of lowercase letters, then some elements of X^+ are

$$road \; * \; runner = roadrunner,$$
$$sledge \; * \; hammer = sledgehammer,$$
$$aababba \; * \; babbba = aababbababbba.$$

Note that we often omit the operation $*$ from the notation when the operation is known. Thus, for instance, we may write $a * b$ simply as ab.

Example 4. The **symmetric inverse monoid**, ***SIM(S)***, consists of all partial injections from a set S to itself. The operation on $SIM(S)$ is given by composition of mappings, where we will adopt the convention that mappings are composed from left to right.

The monoid $SIM(S)$ is of great relevance to this paper. Therefore, let us take some time to elaborate on its structure. Consider the particular case where $S = \{1, 2, 3, \ldots, n\}$, denoted as *SIM(n)*. The identity element of *SIM(n)* is the element that fixes each element of the set $\{1, 2, 3, \ldots, n\}$. The elements of *SIM(n)* can be represented by a classic two-line notation. For example,

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 6 & - & 5 \end{pmatrix} \in SIM(6). \tag{1}$$

In this element, we observe that 1 maps to 2, 2 maps to 1, 3 maps to 4, 4 maps to 6, 5 is not in the domain, and 6 maps to 5. We denote the domain of τ by $Dom(\tau)$ and the range of τ by $Rng(\tau)$.

Definition 3. The **rank** of an element $\tau \in SIM(S)$ is

$$rank(\tau) = |Dom(\tau)| = |Rng(\tau)|.$$

In (1), $rank(\tau) = 5$. Also, note that the set of all elements in $SIM(n)$ of rank n forms the well-known group called the symmetric group, denoted S_n. Consult any introductory text on modern algebra for more information([6], [8]).

To illustrate how composition of partial permutations works, consider this example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 3 & 4 & - & 5 & 6 & - \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 1 & 4 & 6 & - & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 4 & 6 & - & - & 5 & - \end{pmatrix}.$$

Of course, the rank of the composition of two partial mappings cannot exceed the rank of either element comprising the product. Here, for example, we see that the elements on the left side have rank 4 and 5, respectively, while the resulting product on the right side only has rank 3.

There are some fundamental equivalence relations which help us understand the structure of semigroups and monoids. These are known as the **Green's relations** [5]. To discuss these, we need some notation: given a monoid M and $t \in M$, denote $Mt = \{mt : m \in M\}$ and $tM = \{tm : m \in M\}$. Now, let $\alpha, \beta \in M$. Then

1. α and β are $\mathcal{L}$**-related**, denoted $\alpha \mathcal{L} \beta$, if and only if there exist $x, y \in M$ such that $x\alpha = \beta$, $y\beta = \alpha$. That is, $\beta \in M\alpha$ and $\alpha \in M\beta$.

2. α and β are $\mathcal{R}$**-related**, denoted $\alpha \mathcal{R} \beta$, if and only if there exist $u, v \in M$ such that $\alpha u = \beta$, $\beta v = \alpha$. That is, $\beta \in \alpha M$ and $\alpha \in \beta M$.

3. The $\mathcal{D}$**-relation**, $\mathcal{D} = \mathcal{L} \circ \mathcal{R} = \mathcal{R} \circ \mathcal{L}$, is the smallest equivalence relation containing both $\mathcal{L}$ and $\mathcal{R}$.

4. The $\mathcal{H}$**-relation** is the intersection of $\mathcal{L}$ and $\mathcal{R}$. That is, α and β are $\mathcal{H}$-related, denoted $\alpha \mathcal{H} \beta$, if and only if $\alpha \mathcal{L} \beta$ and $\alpha \mathcal{R} \beta$.

The Green's relations are equivalence relations, and hence form equivalence classes, which we will refer to as $\mathcal{L}$-classes, $\mathcal{R}$-classes, $\mathcal{D}$-classes, and $\mathcal{H}$-classes.

We are now ready to begin studying the particular monoid of interest to us in our research, and in the next section, we will determine its Green's relations.

2 Partial Order-preserving Injections

The monoid of interest to our research is the monoid of **partial order-preserving injections**, denoted by ***POI(n)***, which is a submonoid of $SIM(n)$. An element σ of $SIM(n)$ is an element of $POI(n)$ if whenever $i < j$ in $\{1, 2, 3, \ldots, n\}$, then $\sigma(i) < \sigma(j)$. For example,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & - & 5 & 6 & - \end{pmatrix} \in POI(6),$$

since the values on the second row appear in increasing order. However, the element τ in (1) is not an element of $POI(6)$.

Definition 4. A **partial identity** in $POI(n)$ is an element σ such that for each $i \in \{1, 2, 3, \ldots, n\}$, either $\sigma(i) = i$ or $i \notin Dom(\sigma)$.

As we will see in our main results, partial identities play an important role in describing economical generating sets of $POI(n)$. An example of a partial identity is the element

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & - & 3 & - & 5 & - \end{pmatrix} \in POI(6),$$

whose rank is 3. Throughout this paper, we denote the subset of $POI(n)$ of elements of rank k as $POI_k(n)$. For example, $rank(\tau) = 3$, so $\tau \in POI_3(6)$. Note that $POI(n) = \bigcup_{k=0}^{n} POI_k(n)$, and from [1],

$$|POI(n)| = \sum_{k=0}^{n} \binom{n}{k}^2.$$

The kth term of this summation counts the number of elements in $POI_k(n)$. For example,

$$|POI(4)| = \sum_{k=0}^{4} \binom{4}{k}^2 = \binom{4}{0}^2 + \binom{4}{1}^2 + \binom{4}{2}^2 + \binom{4}{3}^2 + \binom{4}{4}^2 = \binom{8}{4} = 70.$$

We now look at the Green's relations for the monoid $POI(n)$. The next lemma summarizes the results. The proof is routine and is omitted.

Lemma 1. *Let $\alpha, \beta \in POI(n)$. Then:*

- $\alpha \mathcal{R} \beta$ *if and only if* $Dom(\alpha) = Dom(\beta)$.

- $\alpha \mathcal{L} \beta$ *if and only if* $Rng(\alpha) = Rng(\beta)$.

- $\alpha \mathcal{D} \beta$ *if and only if* $rank(\alpha) = rank(\beta)$.

- $\alpha \mathcal{H} \beta$ *if and only if* $\alpha = \beta$.

We give some examples of the $\mathcal{R}$ and $\mathcal{L}$ relations in $POI(n)$.

Example 5. Let

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 4 & 5 & 6 & 7 & - & - \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & - & - \end{pmatrix}$$

in $POI(7)$. Since

$$\alpha \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & 2 & - & 3 & 4 & 5 & 6 \end{pmatrix} = \beta$$

and

$$\beta \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & 2 & 4 & 5 & 6 & 7 & - \end{pmatrix} = \alpha,$$

we see that $\alpha \mathcal{R} \beta$.

Example 6. Let

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 5 & 6 & - & - & 7 & 8 \end{pmatrix} \quad \text{and} \quad \gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & - & 4 & 5 & 6 & 7 & 8 & - \end{pmatrix}$$

in $POI(8)$. We have

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & - & 2 & 3 & 4 & 7 & 8 & - \end{pmatrix} \sigma = \gamma$$

and

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 3 & 4 & 5 & - & - & 6 & 7 \end{pmatrix} \gamma = \sigma,$$

so we conclude that $\sigma \, \mathcal{L} \, \gamma$.

Remark 1. Let $\alpha, \beta \in POI_{n-1}(n)$. Then $\alpha\beta \in POI_{n-1}(n)$ if and only if $Rng(\alpha) = Dom(\beta)$. In this case, $\alpha \, \mathcal{R} \, \alpha\beta$ and $\beta \, \mathcal{L} \, \alpha\beta$.

We will now turn our attention to the main problem in this paper, the identification of economical generating sets for $POI(n)$.

Definition 5. A **generating set** for an algebraic structure Z (such as a group, semigroup, monoid, vector space, and so on) is a collection $\mathcal{C}$ of elements in Z such that each element in Z can be expressed by applying a finite sequence of algebraic operations to elements in $\mathcal{C}$.

In linear algebra, a generating set for a vector space is commonly known as a spanning set.

Definition 6. A generating set for Z is called an **minimal generating set** if no proper subset of it generates Z.

The search for minimal generating sets of algebraic structures has led to a steady stream of work and has important applications in mathematics and computer science. In [2], these sets were classified for $SIM(n)$, while in [3], they were studied in the symmetric and alternating groups.

Returning to our study of $POI(n)$, observe that $POI_n(n)$ consists only of the identity element

$$id = \begin{pmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ 1 & 2 & 3 & \cdots & n-1 & n \end{pmatrix}.$$

This element is uninteresting with respect to the generating process, so we will focus our attention on generating the set $POI(n) - \{id\}$. For the remainder of this paper, when we speak of generating $POI(n)$, we will mean that we are generating $POI(n) - \{id\}$, which itself is a subsemigroup of $POI(n)$. The key to generating $POI(n) - \{id\}$ is to generate all of the rank $n-1$ elements first. In fact, it is customary to restrict generating sets of $POI(n)$ to consist of elements of rank $n-1$. By multiplying elements of $POI_{n-1}(n)$ together, we can obtain elements of lower rank (see Remark 1). We arrange the elements in $POI_{n-1}(n)$ in an $n \times n$ array as shown in Figure 1 for $POI(4)$. Elements in the same row have the same domain, elements in the same column have the same range, and elements on the main diagonal are partial identities. Let

$$\widehat{S}_i = \{1, 2, \ldots, i-1, i+1, \ldots, n-1, n\}$$

for $1 \le i \le n$. For example, for $n = 4$, if

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & - & 4 \end{pmatrix}, \tag{2}$$

then $Dom(\sigma) = \widehat{S}_3$ and $Rng(\sigma) = \widehat{S}_1$.

$$
\begin{array}{ccccc}
& \textit{Column 1} & \textit{Column 2} & \textit{Column 3} & \textit{Column 4} \\
\textit{Row 1} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & - \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & - \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & - \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & - \end{pmatrix} \widehat{S}_4 \\[2mm]
\textit{Row 2} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & - & 3 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & - & 4 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & - & 4 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & - & 4 \end{pmatrix} \widehat{S}_3 \\[2mm]
\textit{Row 3} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & - & 2 & 3 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & - & 2 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & - & 3 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & - & 3 & 4 \end{pmatrix} \widehat{S}_2 \\[2mm]
\textit{Row 4} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 3 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 2 & 3 & 4 \end{pmatrix} \widehat{S}_1 \\
& \widehat{S}_4 & \widehat{S}_3 & \widehat{S}_2 & \widehat{S}_1
\end{array}
$$

Figure 1: 4×4 array of the elements in $POI_3(4)$.

Note that if σ is an element of $POI_{n-1}(n)$ with $Dom(\sigma) = \widehat{S}_i$ and $Rng(\sigma) = \widehat{S}_j$, then we can uniquely denote σ as the pair of numbers (i, j). For example, the element σ in (2) can now be represented as $\sigma = (3, 1)$.

Notation: Let X be a subset of $POI_{n-1}(n)$. Let Ω_X denote a directed graph on $\{1, 2, \ldots, n\}$, where $\mathbf{a} \overset{\frown}{} \mathbf{b}$ is a directed edge in Ω_X if and only if $(a, b) \in X$.

If $\sigma = (a, b)$ and $\tau = (c, d)$ are two elements of $POI_{n-1}(n)$, then note that $\sigma\tau = (a, b)(c, d) = (a, d)$ if $b = c$. If $b \neq c$, by Remark 1 the rank of $\sigma\tau$ drops and we do not express $\sigma\tau$ as a pair. Multiplication can now be performed by following directed edges, as shown in our next examples.

Example 7. In $POI(4)$,

$$
\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & - & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & - & 3 \end{pmatrix}
$$

can be expressed as $(3, 1)(1, 4) = (3, 4)$, which is represented graphically as

(3)

In (3), the dotted arrow indicates the result of the composition of the two solid arrows.

Example 8. Again in $POI(4)$,

$$
\begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & - \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & - & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 3 & 4 \end{pmatrix}
$$

can be represented as $(1, 4)(4, 3)(3, 2) = (1, 2)$, which is represented graphically as

(4)

In (3) and (4), the solid directed edges form an example of a *walk* [4].

Definition 7. A **walk** is a sequence of m edges of the form

$$x_0 \to x_1 \to x_2 \to \cdots \to x_m,$$

where $x_0, x_1, x_2, \ldots, x_m$ are (not necessarily distinct) vertices, and $x_i \to x_j$ indicates that there is an edge from x_i to x_j. If $x_0 = x_m$, then the walk is **closed**.
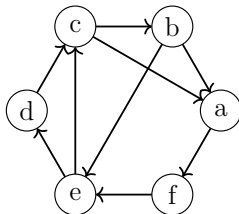

Figure 2:   An example of a directed graph

For example, in Figure 2 we have a walk $c \to b \to e \to d \to c \to a$. Note that if we remove the edge $c \to a$, we have a closed walk: $c \to b \to e \to d \to c$.

Definition 8. If a walk is composed of distinct edges and has distinct vertices, then it is called a **path**. That is, in a path we have $x_i \neq x_j$ for $0 \leq i < j \leq m$, except possibly for $x_0 = x_m$. In the case where $x_0 = x_m$, the path is closed and usually referred to as a **cycle** or **m-cycle**.

In Figure 2, we have the path $e \to d \to c \to b$ and the 4-cycle $e \to d \to c \to b \to e$.

Shortly, we will see how this basic graph theory can help us describe minimal generating sets of $POI(n)$. Our results were first motivated by studying output results from a program written using the programming language Groups, Algorithms, and Programming (GAP) [7]. The program explicitly computed all possible generating sets for $POI(n)$ for small integers n. Let us now describe the main results we obtained.

3   Main Theorems

Our first result provides a necessary condition for a subset $X \subseteq POI_{n-1}(n)$ to be a minimal generating set for $POI(n)$.

Lemma 2. *Let X be a subset of $POI_{n-1}(n)$ consisting of n elements that generate $POI(n)$. Then X contains (a) no two distinct $\mathcal{R}$-related elements, (b) no two distinct $\mathcal{L}$-related elements, and (c) no partial identities.*

Proof. We prove all three parts by way of contradiction. First, to prove part (a), we assume that there are two $\mathcal{R}$-related elements in X. Using Lemma 1, we deduce that for some $i \in \{1, 2, \ldots, n\}$, the subset $\widehat{S}_i$ is the domain for no elements in X. Let $\gamma \in POI_{n-1}(n)$ be any element such that $Dom(\gamma) = \widehat{S}_i$. Write $\gamma = (i, b)$. Since X is a generating set, then there exist $\xi_1, \xi_2, \ldots, \xi_k \in X$ such that

$$\gamma = \xi_1 \xi_2 \cdots \xi_k.$$

Clearly, each ξ_j must belong to $POI_{n-1}(n)$. Therefore, writing $\xi_j = (x_{j1}, x_{j2})$ for each j with $1 \leq j \leq k$, we have

$$\gamma = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{k1}, x_{k2}) = (i, b),$$

where $x_{j2} = x_{(j+1)1}$ for each j with $1 \leq j < k$. It follows that $i = x_{11}$, and thus, $Dom((x_{11}, x_{12})) = \widehat{S}_i$ which is a contradiction.

Similarly, for part (b), assume X contains two $\mathcal{L}$-related elements. Again by Lemma 1, there is some subset $\widehat{S}_i$ that no element in X has as its range. Let $\alpha \in POI_{n-1}(n)$ be any element such that $Rng(\alpha) = \widehat{S}_i$. Write $\alpha = (a, i)$. Then there exist $\xi_1, \xi_2, \ldots, \xi_k \in X$ such that $\alpha = \xi_1 \xi_2 \cdots \xi_k$. Writing $\xi_j = (x_{j1}, x_{j2})$ as before, we have

$$\alpha = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{k1}, x_{k2}) = (a, i),$$

where $x_{j2} = x_{(j+1)1}$ for each $1 \leq j < k$. It follows that $i = x_{k2}$, which implies $Rng((x_{k1}, x_{k2})) = \widehat{S}_i$, which is a contradiction.

Finally, we prove part (c). Assume X contains a partial identity element, say $\omega = (a, a)$, which fixes the elements of $\widehat{S}_a$. Choose any element $\gamma = (a, b) \in POI_{n-1}(n)$ with $\gamma \neq \omega$. Then, following the notation above, γ can be expressed as

$$\gamma = \xi_1 \xi_2 \cdots \xi_k = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{k1}, x_{k2}) = (a, b).$$

As above, we get $a = x_{11}$. By part (a), the only element of X of the form (a, y) is the partial identity (a, a). Therefore, it must be the case that $(x_{11}, x_{12}) = \omega$. Hence,

$$\gamma = (a, a)(x_{21}, x_{22}) \cdots (x_{k1}, x_{k2}).$$

Now we deduce that $x_{21} = a$, and we can repeat the same reasoning as above to conclude that $x_{22} = a$. Therefore,

$$\gamma = (a, a)(a, a) \cdots (x_{k1}, x_{k2}).$$

Proceeding in this way, we eventually have

$$\gamma = (a, a)(a, a) \cdots (a, a) = (a, a) = \omega,$$

a contradiction. $\square$

As a consequence of the proof of Lemma 2, observe that every $\widehat{S}_i$ (for $1 \leq i \leq n$) must be the domain of some generator in a generating set $POI(n)$. This therefore proves the following:

Corollary 1. *Every generating set of $POI(n)$ must contain at least n elements.*

We are now ready to state and prove the central result of this paper.

Theorem 3. *Let X be a subset of n elements of $POI_{n-1}(n)$ with no two distinct $\mathcal{R}$-related elements, no two distinct $\mathcal{L}$-related elements, and no partial identities. Then X is a minimal generating set for $POI(n)$ if and only if the edges of Ω_X form an n-cycle.*

Proof. First, assume X is a minimal generating set for $POI(n)$. By way of contradiction, suppose that the edges of Ω_X do not form an n-cycle. This implies we must have at least two smaller cycles, say C_1 and C_2, of lengths n_1 and n_2, respectively. Let us choose an element $(a, b) \in POI_{n-1}(n)$,

where a is a vertex of the n_1-cycle C_1 and b is a vertex of the n_2-cycle C_2. Since X is a generating set, we can express (a, b) as a product of generators:

$$(a, b) = (x_{11}, x_{12})(x_{21}, x_{22}) \cdots (x_{k1}, x_{k2}).$$

However, since $x_{11} = a$ is a vertex in C_1, then x_{12} is also a vertex in C_1, which implies (x_{11}, x_{12}) is a directed edge in $\Omega_X - C_2$. Similarly, since $x_{12} = x_{21}$ is a vertex in C_1 then $x_{22} = x_{31}$ is a vertex in C_1 which implies (x_{21}, x_{22}) is a directed edge in $\Omega_X - C_2$. Continuing this process, we find that (x_{k1}, x_{k2}) is a directed edge in $\Omega_X - C_2$, which is a contradiction to the fact that $x_{k2} = b$ is a vertex in C_2. Hence, Ω_X forms an n-cycle.

Now assume that the edges of Ω_X form an n-cycle. To show that X is a generating set, let (a, b) be an arbitrary element in $POI_{n-1}(n)$. Since the edges of Ω_X form an n-cycle, then we have a cycle of the form

$$x_1 \to x_2 \to \cdots \to x_{i-1} \to a \to x_{i+1} \to \cdots \to x_{j-1} \to b \to x_{j+1} \to \cdots \to x_n \to x_1.$$

Clearly, there is a path from a to b; namely,

$$a \to x_{i+1} \to \cdots \to x_{j-1} \to b.$$

Therefore, (a, b) can be generated by X as

$$(a, b) = (a, x_{i+1})(x_{i+1}, x_{i+2}) \cdots (x_{j-1}, b).$$

Once all the elements of $POI_{n-1}(n) - \{id\}$ are generated, there is a straightforward algorithm for generating all of the elements of $POI_k(n)$ for $k \leq n - 2$. We describe this algorithm here and Example 9 following the proof will illustrate it.

First, it is easy to generate all partial identity elements of all ranks by simply multiplying together the appropriate partial identities of rank $n - 1$. Now consider any element $\sigma \in POI_k(n)$ for $k \leq n - 2$. Let $\tau \in POI_k(n)$ be a partial identity element with $Dom(\tau) = Dom(\sigma)$. Let $Rng(\sigma) = \{a_1, a_2, \ldots, a_k\}$ and $Rng(\tau) = \{b_1, b_2, \ldots, b_k\}$ with $a_1 < a_2 < \cdots < a_k$ and $b_1 < b_2 < \cdots < b_k$. We now conduct a series of steps that correspond to right-multiplication of τ by elements of rank $n-1$ to replace the k-tuple $[b_1, b_2, \ldots, b_k]$ with $[a_1, a_2, \ldots, a_k]$ in the range. This is accomplished by changing exactly one entry of the k-tuple at a time by exactly one number in such a way that the increasing ordering of the entries of the k-tuple is maintained. To change b_i to $b_i \pm 1$, we use the element of rank $n - 1$ denoted in the pair notation $\gamma = (b_i \pm 1, b_i)$. A direct calculation verifies that $Dom(\tau\gamma) = Dom(\tau)$ and $Rng(\tau\gamma) = \{b_1, b_2, \ldots, b_{i-1}, b_i \pm 1, b_{i+1}, \ldots, b_k\}$. Therefore, we can construct a finite series of elements $\gamma_1, \gamma_2, \ldots, \gamma_r$ belonging to $POI_{n-1}(n)$ such that

$$Dom(\tau\gamma_1\gamma_2 \cdots \gamma_r) = Dom(\tau) = Dom(\sigma)$$

and

$$Rng(\tau\gamma_1\gamma_2 \cdots \gamma_r) = Rng(\sigma).$$

This implies that

$$\sigma = \tau\gamma_1\gamma_2 \cdots \gamma_r. \tag{5}$$

Since τ is a partial identity and each γ_i has rank $n - 1$, we can see from (5) that σ is generated by elements in $POI_{n-1}(n)$. This proves that X is a generating set for $POI(n)$, and since X contains exactly n elements by assumption, Corollary 1 implies that X is a minimal generating set. $\qquad \square$

Example 9. Let $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ - & - & 1 & 5 & 6 & - \end{pmatrix} \in POI(6)$. Here, we start with the partial identity

$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ - & - & 3 & 4 & 5 & - \end{pmatrix}$. Clearly, $Dom(\tau) = Dom(\sigma)$. We now carry out the series of steps described in the algorithm in the proof of Theorem 3 to replace $[3, 4, 5]$ with $[1, 5, 6]$. One way to do this is via the replacements

$$[3, 4, 5] \longrightarrow [2, 4, 5] \longrightarrow [1, 4, 5] \longrightarrow [1, 4, 6] \longrightarrow [1, 5, 6].$$

According to the algorithm, these four steps can be accomplished by right-multiplication by the elements

$$\gamma_1 = (2, 3), \qquad \gamma_2 = (1, 2), \qquad \gamma_3 = (6, 5), \qquad \gamma_4 = (5, 4).$$

We can easily compute that indeed we have

$$\tau \gamma_1 \gamma_2 \gamma_3 \gamma_4 = \sigma.$$

Let us now illustrate some examples of Theorem 3.

Example 10. The set

$$X = \left\{ \begin{array}{c} \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & 2 & 3 \end{pmatrix} \\ (1, 4) \end{array}, \begin{array}{c} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & - \end{pmatrix} \\ (4, 3) \end{array}, \begin{array}{c} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & - & 4 \end{pmatrix} \\ (3, 2) \end{array}, \begin{array}{c} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & - & 3 & 4 \end{pmatrix} \\ (2, 1) \end{array} \right\}$$

forms a generating set for $POI(4)$ since the edges of the graph Ω_X form a 4-cycle. By composing these edges, we can then generate all 16 edges of $POI_3(4)$ to get:
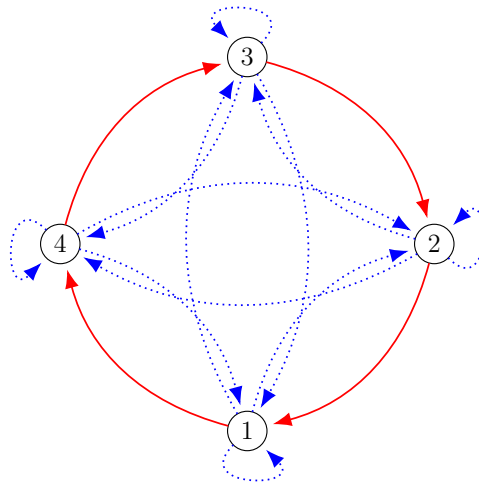

Figure 3: A graphical example of a generating set of $POI(4)$

Here, the solid directed arrows belong to the generating set, while the dotted directed arrows represent the remaining elements of $POI_3(4)$.

Example 11. The set X given by

$$\left\{ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & - \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & - & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & - & 4 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & - & 2 & 3 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & 1 & 3 & 4 & 5 \end{pmatrix} \right\}$$
$$\quad\;\; (5,3) \qquad\qquad (4,5) \qquad\qquad (3,1) \qquad\qquad (2,4) \qquad\qquad (1,2)$$

forms a generating set for $POI(5)$ since the edges of the graph Ω_X form a 5-cycle. From this 5-cycle, we can generate the remaining 20 edges that correspond to the remaining 20 elements of $POI_4(5)$:
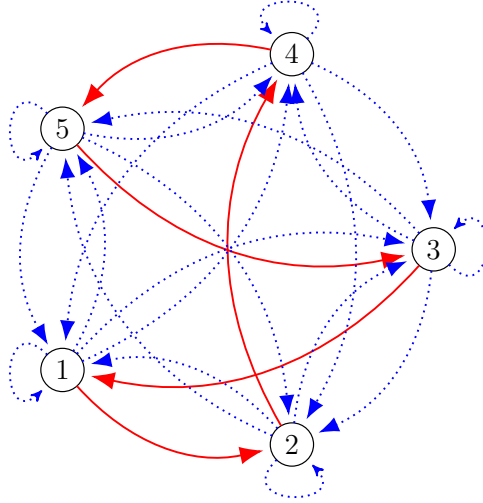

Figure 4: A graphical representation of a generating set of $POI(5)$

As in Example 10, the solid directed arrows belong to the generating set, while the dotted directed arrows represent the remaining elements of $POI_4(5)$.

Example 12. The set X given by

$$\left\{ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 4 & 5 & - \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & - & 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & - & 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & 1 & 3 & 4 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & - & 2 & 3 & 5 \end{pmatrix} \right\}$$
$$\quad\;\; (5,3) \qquad\qquad (3,5) \qquad\qquad (4,1) \qquad\qquad (1,2) \qquad\qquad (2,4)$$

is not a generating set for $POI(5)$ since Ω_X forms two subcycles of the 5-cycle. By composing the edges in these subcycles we only generate 13 out of the 25 edges of $POI_4(5)$, as shown in Figure 5.
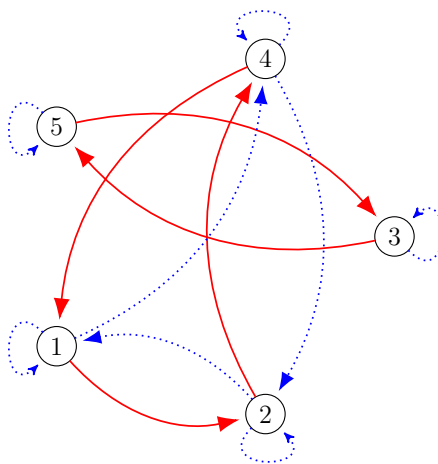
Figure 5: A graphical representation of a non-generating set of $POI(5)$

Theorem 4. *$POI(n)$ has $(n-1)!$ minimal generating sets.*

Proof. From Theorem 3, we see that the number of minimal generating sets of $POI(n)$ is equal to the number of directed n-cycles that can be put on n vertices. It is easy to see that $(n-1)!$ such directed n-cycles can be constructed. □

Acknowledgments

References

[1] U. Alvarez, On k^{th} roots in semigroups of order-preserving partial permutations, *Cal State Fullerton Dimensions*, (2015), 123–132.

[2] S. Annin, Hierarchy of efficient generators of the symmetric inverse monoid, *Semigroup Forum*, **55** (1997), 327–355.

[3] S. Annin, J. Maglione, Economical generating sets for the symmetric and alternating groups consisting of cycles of a fixed length, *J. Algebra Appl.*, **11** (2012).

[4] R.A. Brualdi, *Introductory Combinatorics*, Pearson, 2010.

[5] A.H. Clifford, G.B. Preston, *The Algebraic Theory of Semigroups*, American Mathematical Society, 1961.

[6] D.S. Dummit, R.M. Foote, *Abstract Algebra*, John Wiley and Sons, Inc., 2004.

[7] GAP 2017, The GAP Group, GAP-Groups, Algorithms, and Programming, Version 4.8.7, available online at the URL: `http://www.gap-system.org`

[8] W.T. Hungerford, *Abstract Algebra: An Introduction*, Brooks/Cole, Cengage Learning, 2013.

On Applications of The AM-GM Inequality

Author: Sushanth Sathish Kumar
Advisor: Bogdan Suceava

Abstract

The AM-GM inequality, which compares the arithmetic and geometric mean of a list of numbers, is one of the most foundational theorem in the subject of inequalities. Several complicated inequalities are consequences of this seemingly simple result. Not only does the AM-GM inequality aid in solving inequalities, it's conditions for equality can also make it adept in solving complicated equations which cannot be solved directly. During the last several issues of the American Mathematical Monthly (AMM), which is a prestigious math journal read and edited by scholars and professionals from around the world, there several inequality problems. Of those, four could be solved by applying only the AM-GM inequality. These four problems will be the focus of this presentation.

1. INTRODUCTION

The origins of the theory of inequalities are related to the work of Augustin-Louis Cauchy (August 21, 1789, Paris - May 23, 1857, Sceaux (near Paris)) In 1802, Cauchy entered the École Centrale du Panthéon where, following Lagrange's advice, he spent two years studying classical languages. He took the entrance examinations for the École Polytechnique in 1805 (at 16 years old) and the examiner Biot placed him second. At the École Polytechnique he attended courses by Lacroix, de Prony, Hachette, and was tutored in analysis by Ampère. Cauchy was one of the most knowledgeable mathematicians in his generation. He noticed that something is missing from the theory: besides his investigations of symmetric functions, he was much interested in developing a theory comparing certain symmetric expressions. During his lifetime, Cauchy wrote 789 mathematical papers, an incredible achievement.

We now present the main theorem.

Theorem 1.1 (AM-GM). *Let $a_1, ..., a_n$ be a list of n positive real numbers. Then,*

$$\frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 ... a_n}.$$

Equality holds if and only if all the variables are equal.

Throughout this paper, we make use of the shorthand $\sum_{cyc} f(a, b, c)$ to denote the sum $f(a, b, c) + f(b, c, a) + f(c, a, b)$.

2. SIMPLE APPLICATIONS

In this section we present some simple consequences of the AM-GM inequality to ease the reader to the material in the following sections.

Example 2.1. If $x_1, x_2, ..., x_n$ are positive real numbers, prove that

$$\frac{x_1}{x_2} + \frac{x_2}{x_3} + ... + \frac{x_{n-1}}{x_n} + \frac{x_n}{x_1} \geq n,$$

the equality being valid only when $x_1 = x_2 = ... = x_n$.

Proof. Since:

$$\frac{x_1}{x_2} \cdot \frac{x_2}{x_3} \cdot \ldots \cdot \frac{x_{n-1}}{x_n} \cdot \frac{x_n}{x_1} = 1,$$

then,

$$\frac{x_1}{x_2} + \frac{x_2}{x_3} + \ldots + \frac{x_{n-1}}{x_n} + \frac{x_n}{x_1} \geq n.$$

Equality holds only when

$$\frac{x_1}{x_2} = \frac{x_2}{x_3} = \ldots = \frac{x_{n-1}}{x_n} = \frac{x_n}{x_1} = 1,$$

which yields $x_1 = x_2 = \ldots = x_n$. $\qquad\square$

Although there are some arguably more beautiful ways to prove the following lemma, we present the AM-GM method here to show how versatile and indispensable the inequality is.

Lemma 2.2. *Let a, b, c be three positive reals. Then,*

$$bc + ca + ab \leq a^2 + b^2 + c^2.$$

Proof. By AM-GM,

$$bc \leq \frac{b^2 + c^2}{2}, \quad ca \leq \frac{c^2 + a^2}{2}, \quad ab \leq \frac{a^2 + b^2}{2}.$$

Adding up the inequalities yields the desired result. $\qquad\square$

Finally, to wrap up the section we present a result that seems to lend itself to a solution using calculus, but can be proved far more elegantly with the AM-GM inequality.

Example 2.3. Prove that for any $a > 1$

$$\ln a + \log_a e \geq 2.$$

Proof. Let us write

$$\log_a e = \frac{\ln e}{\ln a} = \frac{1}{\ln a}.$$

Then,

$$\ln a + \log_a e = \frac{\ln a}{1} + \frac{1}{\ln a}.$$

Now Example 2.1 with $n = 2$ proves that the latter term is at least 2. $\qquad\square$

3. An Inequality with Factorials

Problem 3.1. (AMM 12065) *Proposed by Hojoo Lee, Seoul National University, Seoul, South Korea.* Let n be a positive integer, and let $x_1, \ldots, x_n$ be a list of n positive real numbers. For $k \in \{1, \ldots, n\}$, let $y_k = x_k(n+1)/(n+1-k)$ and let

$$z_k = \frac{(k!)^{1/k}}{k+1} \left(\prod_{j=1}^{k} y_j \right)^{1/k}.$$

Prove that the arithmetic mean of $x_1, \ldots, x_n$ is greater than or equal to the arithmetic mean of $z_1, \ldots, z_n$, and determine when equality holds.

Proof. By AM-GM, we have

$$z_k = \frac{1}{k+1}\left(\prod_{j=1}^{k} jy_j\right)^{1/k} \leq \frac{1}{k(k+1)}\sum_{j=1}^{k} jy_j = \sum_{j=1}^{k}\frac{j(n+1)x_j}{(n+1-j)k(k+1)}.$$

Swapping the order of summation,

$$\sum_{k=1}^{n} z_k \leq \sum_{k=1}^{n}\sum_{j=1}^{k}\frac{j(n+1)x_j}{(n+1-j)k(k+1)} = \sum_{j=1}^{n}\sum_{k=j}^{n}\frac{j(n+1)x_j}{n+1-j}\left(\frac{1}{k}-\frac{1}{k+1}\right)$$

Therefore, we have

$$\sum_{k=1}^{n} z_k = \sum_{j=1}^{n}\frac{j(n+1)x_j}{n+1-j}\sum_{k=j}^{n}\left(\frac{1}{k}-\frac{1}{k+1}\right) = \sum_{j=1}^{n} x_j,$$

which was what we wanted. Equality, by AM-GM, occurs when $y_1 = 2y_2 = \cdots = ny_n$. Equivalently, we must have

$$x_k = \frac{C(n+1-k)}{k(n+1)}$$

for a constant C. $\qquad\square$

4. Inequality on the Altitudes and Exradii of a Triangle

Problem 4.1. (AMM 12068) *Proposed by D. M. Bătinețu-Giurgiu, "Matei Basarab" National College, Bucharest, Romania, and Neculai Stanciu, "George Emil Palade" School, Buzău, Romania.* Consider a triangle with altitudes $h_a, h_b,$ and h_c and corresponding exradii $r_a, r_b,$ and r_c. Let s, r, and R denote the triangle's semiperimeter, inradius, and circumradius, respectively.

(a) Prove

$$\frac{h_b + h_c}{h_a}r_a^2 + \frac{h_c + h_a}{h_b}r_b^2 + \frac{h_a + h_b}{h_c}r_c^2 \geq 2s^2.$$

(b) Prove

$$\frac{r_b + r_c}{r_a}h_a^2 + \frac{r_c + r_a}{r_b}h_b^2 + \frac{r_a + r_b}{r_c}h_c^2 \geq \frac{4s^2 r}{R}.$$

Proof. Letting K denote the area of the triangle, we see that $h_a = 2K/a$, and

$$r_a^2 = \frac{K^2}{(s-a)^2} = \frac{s(s-b)(s-c)}{s-a},$$

by Heron's formula. It follows that the left-hand-side of the inequality is

$$\sum_{cyc}\frac{\frac{1}{b}+\frac{1}{c}}{\frac{1}{a}}\frac{s(s-b)(s-c)}{s-a} = \sum_{cyc}a\left(\frac{1}{b}+\frac{1}{c}\right)\frac{s(s-b)(s-c)}{s-a}.$$

Thus, we wish to show

$$\sum_{cyc}\left(\frac{as(s-b)(s-c)}{b(s-a)} + \frac{as(s-b)(s-c)}{c(s-a)}\right) \geq 2s^2.$$

Since the sum is cyclic we may rewrite it as follows,

$$\sum_{cyc}\left(\frac{as(s-b)(s-c)}{b(s-a)} + \frac{bs(s-a)(s-c)}{a(s-b)}\right).$$

Applying AM-GM to the two terms in parentheses,

$$\sum_{cyc} \left(\frac{as(s-b)(s-c)}{b(s-a)} + \frac{bs(s-a)(s-c)}{a(s-b)} \right) \geq \sum_{cyc} 2s(s-c) = 2s^2,$$

which solves the first part. For **part (b)** note that

$$\frac{r_b + r_c}{r_a} h_a^2 = \frac{\frac{K}{s-b} + \frac{K}{s-c}}{\frac{K}{s-a}} h_a^2 = \frac{s-a}{s-b} h_a^2 + \frac{s-a}{s-c} h_a^2.$$

Therefore the left-hand-side is

$$\sum_{cyc} \left(\frac{s-a}{s-b} h_a^2 + \frac{s-a}{s-c} h_a^2 \right) = \sum_{cyc} \left(\frac{s-a}{s-b} h_a^2 + \frac{s-b}{s-a} h_b^2 \right).$$

Applying AM-GM to the terms in parentheses again,

$$\sum_{cyc} \left(\frac{s-a}{s-b} h_a^2 + \frac{s-b}{s-a} h_b^2 \right) \geq \sum_{cyc} 2h_a h_b = \sum_{cyc} \frac{8K^2}{ab}.$$

Recalling the formulas

$$R = \frac{abc}{4K} \quad \text{and} \quad rs = K,$$

we can simplify the right-hand-side as

$$\sum_{cyc} \frac{8K^2}{ab} = \frac{8K^2(a+b+c)}{abc} = \frac{2K(a+b+c)}{R} = \frac{4s^2 r}{R}.$$

which is precisely the right-hand-side of the original inequality. $\qquad \square$

5. ANOTHER GEOMETRIC INEQUALITY

Problem 5.1. (AMM 12103) *Proposed by George Apostolopoulos, Messolonghi, Greece. Let a, b, and c be the side lengths of a triangle with inradius r and circumradius R. Let r_a, r_b, and r_c be the exradii opposite the sides of length a, b, and c, respectively. Prove*

$$\frac{1}{2R^3} \leq \frac{r_a}{a^4} + \frac{r_b}{b^4} + \frac{r_c}{c^4} \leq \frac{1}{16r^3}.$$

Proof. Denote by K and s the area and semiperimeter of the triangle, respectively. By the Ravi Substitution, there are positive reals x, y, and z such that $a = y + z$, $b = z + x$, and $c = x + y$. It follows, $K = \sqrt{xyz(x+y+z)}$ and $s = x + y + z$.

Recalling the well-known formulas

$$R = \frac{abc}{4K}, \quad r = \frac{K}{s}, \quad r_a = \frac{K}{s-a},$$

we obtain,

$$R = \frac{(x+y)(y+z)(z+x)}{4\sqrt{xyz(x+y+z)}}, \quad r = \sqrt{\frac{xyz}{x+y+z}}, \quad r_a = \sqrt{\frac{yz(x+y+z)}{x}}.$$

Hence, the given inequality is equivalent to

$$\frac{32(xyz)^{3/2}(x+y+z)^{3/2}}{(x+y)^3(y+z)^3(z+x)^3} \leq \sum_{cyc} \frac{1}{(y+z)^4} \sqrt{\frac{yz(x+y+z)}{x}} \leq \frac{(x+y+z)^{3/2}}{16(xyz)^{3/2}}.$$

For the right hand side, note that by AM-GM, we have $(y+z)^4 \geq 16y^2z^2$. Thus,

$$\sum_{cyc} \frac{1}{(y+z)^4}\sqrt{\frac{yz(x+y+z)}{x}} \leq \sum_{cyc} \frac{1}{16y^2z^2}\sqrt{\frac{yz(x+y+z)}{x}}$$

$$= \frac{1}{16}\sum_{cyc}\sqrt{\frac{x+y+z}{xy^3z^3}}$$

$$= \frac{1}{16}\sum_{cyc}\frac{x\sqrt{x+y+z}}{\sqrt{x^3y^3z^3}}$$

$$= \frac{(x+y+z)^{3/2}}{16(xyz)^{3/2}}.$$

To prove the left hand side, rewrite the sum as follows

$$\sum_{cyc}\frac{1}{(y+z)^4}\sqrt{\frac{yz(x+y+z)}{x}} = \sum_{cyc}\frac{16}{(y+z)^4}\sqrt{\frac{x^3y^3z^3(x+y+z)}{256x^4y^2z^2}}.$$

By AM-GM, we may estimate $256x^4y^2z^2 = (16x^2y^2)(16x^2z^2) \leq (x+y)^4(x+z)^4$. Therefore, we have

$$\sum_{cyc}\frac{1}{(y+z)^4}\sqrt{\frac{yz(x+y+z)}{x}} \geq \sum_{cyc}\frac{16\sqrt{x^3y^3z^3(x+y+z)}}{(y+z)^4(x+y)^2(x+z)^2}$$

$$= \frac{16\sqrt{x^3y^3z^3(x+y+z)}}{(x+y)^2(y+z)^2(z+x)^2}\sum_{cyc}\frac{1}{(y+z)^2}$$

Recalling Lemma 2.2, we obtain

$$\sum_{cyc}\frac{1}{(y+z)^4}\sqrt{\frac{yz(x+y+z)}{x}} \geq \frac{16\sqrt{x^3y^3z^3(x+y+z)}}{(x+y)^2(y+z)^2(z+x)^2}\sum_{cyc}\frac{1}{(x+y)(x+z)}$$

$$= \frac{16\sqrt{x^3y^3z^3(x+y+z)}}{(x+y)^2(y+z)^2(z+x)^2}\sum_{cyc}\frac{y+z}{(x+y)(x+z)(y+z)}$$

$$= \frac{32(x+y+z)\sqrt{x^3y^3z^3(x+y+z)}}{(x+y)^3(y+z)^3(z+x)^3}$$

$$= \frac{32(xyz)^{3/2}(x+y+z)^{3/2}}{(x+y)^3(y+z)^3(z+x)^3},$$

which was what we wanted. $\qquad\square$

6. CONDITIONAL GEOMETRIC INEQUALITY

Problem 6.1. (AMM 12098) *Proposed by Leonard Giugiuc, Drobeta Turnu Severin, Romania, and Kadir Altintas, Emirdağ, Turkey.* Suppose that the centroid of a triangle with semiperimeter s and inradius r lies on its incircle. Prove $s \geq 3\sqrt{6}r$, and determine conditions for equality.

Proof. Let the side lengths of the triangle be $a \geq b \geq c$, and K its area. Denote by G and I the centroid and incenter of the triangle, respectively. By scaling, we may assume that $c = 1$. Note that this implies that $a + b \geq 2c = 2$.

From the barycentric coordinates

$$G = \frac{1}{3}\vec{A} + \frac{1}{3}\vec{B} + \frac{1}{3}\vec{B} \text{ and } I = \frac{a}{a+b+c}\vec{A} + \frac{b}{a+b+c}\vec{B} + \frac{c}{a+b+c}\vec{C},$$

it is tedious, but straightforward to derive

$$
\begin{aligned}
GI^2 &= \frac{1}{9(a+b+c)^2}\|(b+c-2a)\vec{A} + (c+a-2b)\vec{B} + (a+b-2c)\vec{C}\|^2 \\
&= \frac{-a^3 - b^3 - c^3 + 2a^2b + 2ab^2 + 2b^2c + 2bc^2 + 2c^2a + 2ca^2 - 9abc}{9(a+b+c)} \\
&= \frac{-a^3 - b^3 - 1 + 2a^2b + 2ab^2 + 2b^2 + 2b + 2a + 2a^2 - 9ab}{9(a+b+1)}.
\end{aligned}
$$

Additionally, from Heron's formula and the well-known $r = K/s$,

$$
\begin{aligned}
r^2 = \frac{K^2}{s^2} &= \frac{(a+b-1)(b+1-a)(1+a-b)}{4(a+b+1)} \\
&= \frac{-a^3 - b^3 - 1 + a^2b + ab^2 + a^2 + b^2 + a + b - 2ab}{4(a+b+1)}.
\end{aligned}
$$

Since G lies on the incircle, we have $GI^2 = r^2$. Setting the two expressions equal to each other, and simplifying yields

$$5a^3 + 5b^3 + 5 - 18ab = a^2b + ab^2 + a^2 + b^2 + a + b$$

$$\iff (a+b+1)(5a^2 + 5b^2 - 6ab - 6a - 6b + 5) = 0$$

(1) $$\iff 5a^2 + 5b^2 + 5 = 6ab + 6a + 6b.$$

We claim that $a + b \leq 5$. Indeed, adding $10ab$ to both sides and applying AM-GM gives

(2) $$5(a+b)^2 + 5 = 16ab + 6a + 6b \leq 4(a+b)^2 + 6(a+b).$$

Consequently, $(a+b-1)(a+b-5) = (a+b)^2 - 6(a+b) + 5 \leq 0$, which gives the desired $a + b \leq 5$. To show $s \geq 3\sqrt{6}r$, it suffices to show that

$$s^2 \geq 54r^2 = \frac{54K^2}{s^2} = \frac{54(s-a)(s-b)(s-c)}{s}.$$

Hence, it is enough to establish

$$(a+b+1)^3 \geq 54(a+b-1)(1-a+b)(a-b+a).$$

To see this, note that

$$
\begin{aligned}
(1-a+b)(1-b+a) &= 1 - a^2 - b^2 + 2ab \\
&= \frac{1}{5}(5 - (5a^2 + 5b^2) + 10ab) \\
&= \frac{1}{5}(10 + 4ab - 6a - 6b) \\
&\leq \frac{1}{5}(10 + (a+b)^2 - 6(a+b)),
\end{aligned}
$$

where we have use the AM-GM inequality and (1). For brevity, write $\mu = a + b$. Then, we wish to show that

$$(\mu+1)^3 \geq \frac{54}{5}(\mu-1)(\mu^2 - 6\mu + 10) \iff (5-\mu)(49\mu^2 - 148\mu + 109) \geq 0.$$

But since $2 \leq \mu \leq 5$, this is evident. Equality holds above if and only if $\mu = 5$. Therefore, (2) must actually be an equality, which implies $4ab = (a + b)^2$, or that $a = b = 5/2$. Rescaling the triangle, we see that equality holds if and only if the sides are in a $2 : 5 : 5$ ratio. $\square$

REFERENCES

[1] David R. Wilkins: *Getting started with Latex*,
    http://www.maths.tcd.ie/~dwilkins/LaTeXPrimer/
[2] http://en.wikibooks.org/wiki/LaTeX

PORTOLA HIGH SCHOOL, IRVINE CA 92620
Email address: brokard81517@gmail.com

An Interesting Series and The Geometric Mean of the Numbers in a Positive Interval

Author: Sushanth Sathish Kumar
Advisor: Bogdan Suceava

In the present note we discuss a problem inspired by the reference [2], we generalize it, and from this idea we derive a result on series. In our investigation we pursue first an elementary solution, similar to the argument from [2] (in the first problem, Solution 1). Interestingly enough, an argument by using Sterling formula is possible, and we present it below in the second solution. This discussion leads to an interesting interpretation: the geometric mean of all the real numbers in a positive interval $[m, m+1]$, for $m \in (0, \infty)$. The question which motivated our inquiry is the following.

Problem 1. *Define*

$$a_n = \prod_{k=1}^{2n} \left(1 + \frac{k}{2n}\right).$$

Evaluate $\lim_{n \longrightarrow \infty} a_n$ *and* $\lim_{n \longrightarrow \infty} \frac{a_{n+1}}{a_n}$

Solution 1: First, remark that $a_n > 2^n$. Indeed, by noting that

$$a_n^2 = \prod_{k=1}^{2n} \left(1 + \frac{k}{2n}\right)^2 = \prod_{k=1}^{2n} \left(1 + \frac{k}{2n}\right)\left(1 + \frac{2n+1-k}{2n}\right).$$

By using the estimate $\left(1 + \frac{k}{2n}\right)\left(1 + \frac{2n+1-k}{2n}\right) > 1 + \frac{k+2n+1-k}{2n} > 2$, it follows that $a_n^2 > 2^{2n}$, as claimed above. Therefore $\lim_{n \longrightarrow \infty} a_n > \lim_{n \longrightarrow \infty} 2^n = \infty$, and in consequence a_n approaches infinity. To evaluate the second limit, write $a_n = \frac{(4n)!}{(2n)!(2n)^{2n}}$, and note that

$$\frac{a_{n+1}}{a_n} = \frac{\frac{(4n+4)!}{(2n+2)!(2n+2)^{2n+2}}}{\frac{(4n)!}{(2n)!(2n)^{2n}}} = \frac{(4n+1)(4n+2)(4n+3)(4n+4)}{(2n+1)(2n+2)^3 \left(1 + \frac{1}{n}\right)^{2n}}.$$

When $n \longrightarrow \infty$, we have $\frac{(4n+1)(4n+2)(4n+3)(4n+4)}{(2n+1)(2n+2)^3} \longrightarrow \left(\frac{4}{2}\right)^4 = 16$. By using the properties of Euler's classical sequence, $\lim_{n \longrightarrow \infty} \left(\left(1 + \frac{1}{n}\right)^n\right)^2 = e^2$. Therefore, the second limit is $\frac{16}{e^2}$, as desired. □

Solution 2: The idea is to rewrite the limit as

$$a_n = \frac{\prod_{k=1}^{2n}(2n+k)}{(2n)^{2n}} = \frac{(4n)!}{(2n)!(2n)^{2n}}.$$

By Stirling's formula (see e.g. [3]), we have the approximations $(4n)! \sim \sqrt{8\pi n}\left(\frac{4n}{e}\right)^{4n}$ and $(2n)! \sim \sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}$. Therefore, we may write

$$a_n \sim \frac{\sqrt{8\pi n}\left(\frac{4n}{e}\right)^{4n}}{\sqrt{4\pi n}\left(\frac{2n}{e}\right)^{2n}(2n)^{2n}} = \sqrt{2}\left(\frac{4}{e}\right)^{2n}.$$

Observe that since $4/e > 1$, we have that $\lim_{n\longrightarrow\infty} a_n = \lim_{n\longrightarrow\infty}\sqrt{2}\left(\frac{4}{e}\right)^{2n} = \infty$. To evaluate the second limit, write

$$\lim_{n\longrightarrow\infty}\frac{a_{n+1}}{a_n} = \lim_{n\longrightarrow\infty}\frac{\sqrt{2}\left(\frac{4}{e}\right)^{2n+2}}{\sqrt{2}\left(\frac{4}{e}\right)^{2n}} = \frac{16}{e^2}.$$

$\square$

Problem 2. *Let m be a fixed positive integer, $m \geq 1$. Define*

$$a_n = \prod_{k=1}^{mn}\left(1 + \frac{k}{mn}\right).$$

Evaluate $\lim_{n\longrightarrow\infty} a_n$ and $\lim_{n\longrightarrow\infty}\frac{a_{n+1}}{a_n}$

Solution 1: We claim that $a_n > 2^{mn/2}$. Indeed, writing

$$a_n^2 = \prod_{k=1}^{mn}\left(1 + \frac{k}{mn}\right)\left(1 + \frac{mn-k+1}{mn}\right)$$

and making the estimate $\left(1 + \frac{k}{mn}\right)\left(1 + \frac{mn-k+1}{mn}\right) > 1 + \frac{k}{mn} + \frac{mn-k+1}{mn} > 2$ yields $a_n^2 > 2^{mn}$, as desired. Thus, $\lim_{n\longrightarrow\infty} a_n > \lim_{n\longrightarrow\infty} 2^{mn/2} = \infty$, which shows that a_n diverges. Observe that $a_n = \frac{(2mn)!}{(mn)!(mn)^{mn}}$, and write

$$\lim_{n\longrightarrow\infty}\frac{a_{n+1}}{a_n} = \lim_{n\longrightarrow\infty}\frac{\frac{(2mn+2m)!}{(mn+m)!(mn+m)^{mn+m}}}{\frac{(2mn)!}{(mn)!(mn)^{mn}}} = \lim_{n\longrightarrow\infty}\frac{(2mn+1)\cdots(2mn+2m)}{(mn+1)\cdots(mn+m)(mn+m)^m\left(\frac{mn+m}{mn}\right)^{mn}}.$$

As $n\longrightarrow\infty$, we have $\frac{(2mn+1)\cdots(2mn+2m)}{(mn+1)\cdots(mn+m)(mn+m)^m}\longrightarrow 2^{2m} = 4^m$. Euler's limit now gives $\lim_{n\longrightarrow\infty}\left(\frac{mn+m}{mn}\right)^{mn} = \lim_{n\longrightarrow\infty}\left(\left(1+\frac{1}{n}\right)^n\right)^m = e^m$. Therefore, the second limit is $\left(\frac{4}{e}\right)^m$.

$\square$

Solution 2: Let us rewrite a_n as

$$a_n = \frac{\displaystyle\prod_{k=1}^{mn}(mn+k)}{(mn)^{mn}} = \frac{(2mn)!}{(mn)!(mn)^{mn}}.$$

Stirling's Formula now gives $(2mn)! \sim \sqrt{4\pi mn}\left(\frac{2mn}{e}\right)^{2mn}$ and $(mn)! \sim \sqrt{2\pi mn}\left(\frac{mn}{e}\right)^{mn}$. Therefore,

$$a_n \sim \frac{\sqrt{4\pi mn}\left(\frac{2mn}{e}\right)^{2mn}}{\sqrt{2\pi mn}\left(\frac{mn}{e}\right)^{mn}(mn)^{mn}} = \sqrt{2}\left(\frac{4}{e}\right)^{mn}.$$

Since $4/e > 1$, we again derive that a_n diverges. For the second limit we obtain,

$$\lim_{n\longrightarrow\infty}\frac{a_{n+1}}{a_n} = \lim_{n\longrightarrow\infty}\frac{\sqrt{2}\left(\frac{4}{e}\right)^{m(n+1)}}{\sqrt{2}\left(\frac{4}{e}\right)^{mn}} = \left(\frac{4}{e}\right)^m.$$

$\square$

Proposition 1. *Let m be a fixed positive integer, $m \geq 1$, and let $a_n = \prod_{k=1}^{mn}\left(1+\frac{k}{mn}\right)$. Then the series defined by*

$$\sum_{n=1}^{\infty}a_n^{-1} = \sum_{n=1}^{\infty}\left\{\prod_{k=1}^{mn}\left(1+\frac{k}{mn}\right)\right\}^{-1}$$

is convergent.

Proof: It suffices to show that $\lim_{n\longrightarrow\infty}\frac{a_n}{a_{n+1}} < 1$, by the ratio test. From the second problem, we know that $\lim_{n\longrightarrow\infty}\frac{a_{n+1}}{a_n} = \left(\frac{4}{e}\right)^m$. Therefore, we have

$$\lim_{n\longrightarrow\infty}\frac{a_n}{a_{n+1}} = \left(\frac{e}{4}\right)^m < 1,$$

as desired. $\square$

With this preparation, we are ready now to provide the following interpretation.

Proposition 2. *Given a fixed constant $m > 0$, the geometric mean of all the numbers in the real interval $[m, m+1]$ is the following*

$$\lim_{n\longrightarrow\infty}\left(\prod_{k=1}^{n}\left(m+\frac{k}{n}\right)\right)^{1/n}.$$

Solution: In order to evaluate the above stated limit, taking logarithms yields,

$$\ln\lim_{n\longrightarrow\infty}\left(\prod_{k=1}^{n}\left(m+\frac{k}{n}\right)\right)^{1/n} = \lim_{n\longrightarrow\infty}\left(\frac{1}{n}\sum_{k=1}^{n}\ln\left(m+\frac{k}{n}\right)\right) = \lim_{n\longrightarrow\infty}\sum_{k=1}^{n}\frac{1}{n}\ln\left(m+\frac{k}{n}\right).$$

This is a Riemann integral, and we may write

$$\lim_{n \longrightarrow \infty} \sum_{k=1}^{n} \frac{1}{n} \ln \left(m + \frac{k}{n} \right) = \int_{m}^{m+1} \ln x \, dx$$

$$= x \ln x - x \Big|_{m}^{m+1}$$

$$= (m + 1) \ln(m + 1) - m \ln m - 1$$

Hence, the original limit evaluates to

$$\lim_{n \longrightarrow \infty} \left(\prod_{k=1}^{n} \left(m + \frac{k}{n} \right) \right)^{1/n} = e^{(m+1) \ln(m+1) - m \ln m - 1} = \frac{(m + 1)^{m+1}}{e \cdot m^m}.$$

$\square$

The next problem assesses when a product is convergent, and it naturally pertains to the context developed by our discussion.

Corollary 1. *[1] For a > 0, evaluate*

$$\lim_{n \longrightarrow \infty} \prod_{k=1}^{n} \left(a + \frac{k}{n} \right).$$

Solution: The key idea is to write the limit as

$$\lim_{n \longrightarrow \infty} \prod_{k=1}^{n} \left(a + \frac{k}{n} \right) = \lim_{n \longrightarrow \infty} \left[\left(\prod_{k=1}^{n} \left(a + \frac{k}{n} \right) \right)^{1/n} \right]^{n}.$$

Indeed, applying the result of Problem 3 yields

$$\lim_{n \longrightarrow \infty} \left[\left(\prod_{k=1}^{n} \left(a + \frac{k}{n} \right) \right)^{1/n} \right]^{n} = \lim_{n \longrightarrow \infty} \left(\frac{(a + 1)^{a+1}}{e \cdot a^a} \right)^{n}.$$

Thus, the limit is

$$\lim_{n \longrightarrow \infty} \prod_{k=1}^{n} \left(a + \frac{k}{n} \right) = \begin{cases} 0 & \text{if } \frac{(a+1)^{a+1}}{e \cdot a^a} < 1 \\ 1 & \text{if } \frac{(a+1)^{a+1}}{e \cdot a^a} = 1 \\ \infty & \text{if } \frac{(a+1)^{a+1}}{e \cdot a^a} > 1 \end{cases}$$

References

[1] H. Ohtsuka, Problem 12029, the American Mathematical Monthly.

[2] Liviu Pîrşan, 500 Practice Problems in Algebra and Analysis, Pan General Press, 1995.

[3] Teodora-Liliana T. Rădulescu, Vicenţiu D. Rădulescu, Titu Andreescu, *Problems in Real Analysis: Advanced Calculus on the Real Axis*, Springer, 2009.

MAXWELL'S CONJECTURE ON FOUR COLLINEAR POINT CHARGES

Author: Solomon Huang and Duy Nguyen
Faculty Advisor: Prof. T. Murphy
California State University, Fullerton
Department of Mathematics

Abstract

We study an old conjecture of Maxwell concerning the critical points of an electrostatic potential with finitely many point charges under the condition that the point charges are collinear. If the outermost charge has different sign to all others, we show there are finitely many critical points. In the case where there are four point charges, we investigate the conjecture in a special case.

1. INTRODUCTION

1.1. **Background.** A vector field $\mathbb{E}$ on a domain $\Omega \subseteq \mathbb{R}^3$ is said to be conservative if $\mathbb{E} = \nabla\phi$ for some function $\phi : \Omega \to \mathbb{R}$. These occur throughout physics and engineering. A point charge is defined as a charge $q \in \mathbb{R}\backslash\{0\}$ located at a point $\mathbf{x} \in \mathbb{R}^3$. Given charges q_i located at distinct $\mathbf{x}_i$, $i = 1, \ldots, n$, we define $\Omega := \mathbb{R}^3\backslash\{\mathbf{x}_i, \ i = 1, \ldots, n\}$ and the electrostatic (or Newtonian) potential $\phi : \Omega \to \mathbb{R}$ via

$$\phi(\mathbf{x}) := \sum_{i=1}^{n} \frac{q_i}{\|\mathbf{x} - \mathbf{x}_i\|}.$$

Its gradient $\mathbb{E} = \nabla\phi$ is the electrostatic field generated by the point charges.

Example 1.1. *A standard example is the vector field*

$$\mathbb{G}(\boldsymbol{x}) = \mathbb{G}(x, y, z) = \left\langle \frac{-x}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-y}{(x^2 + y^2 + z^2)^{3/2}}, \frac{-z}{(x^2 + y^2 + z^2)^{3/2}} \right\rangle,$$

which is conservative as $\mathbb{G} = \nabla\phi$, where $\phi : \mathbb{R}^3\backslash\{\mathbf{0}\}$ is the potential energy defined as

$$\phi(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}.$$

This paper is concerned with the following classical question, raised by Maxwell [3]. If there are finitely many distinct point charges, how many points are there in Ω where the vector field vanishes? At such points $\mathbf{x}$ no force is felt by an observer, and we have $\mathbb{E}(\mathbf{x}) = \nabla\phi(\mathbf{x}) = 0$. Thus $\mathbf{x}$ is a critical point of ϕ. Maxwell conjectured that if there are n point charges the number of critical points, if finite, is at most $(n-1)^2$.

This is still largely open. It is not even known in general under what conditions the number of critical points will be finite. Maxwell's conjecture has a long history, and generalizes work of Gauß on electrostatic potentials in $\mathbb{R}^2$. Recent progress on related problems is outlined in [2] and [5]. In $\mathbb{R}^3$, most work to date has focused on

the case where the potential is a Morse function, which means all the critical points are nondegenerate. In fact, in Section 32 of [4] Cairns and Morse study this situation as an application of Morse Theory. They prove a Theorem which applies when all the charges lie on a line (or more generally form a so-called minimal configuration) which realizes a lower bound on the number of critical points via the Morse Inequalities.

In [6], Tsai proves the conjecture for the case where $n = 3$ under some assumptions on the charges. Tsai proves that there are either two, three or four critical points if they are isolated, and describes exactly how all these cases occur. For the collinear case (i.e. all charges are on the same line in $\mathbb{R}^3$), he proves

Theorem 1.2. *Assume three point charges are located at (1,0,0),(-1,0,0) and $(u, v, 0)$ with charges $s^3, k^3, 1$ respectively. For $s = k = 1$, there are two critical points for ϕ if $v = 0$, $u \neq \pm 1$. For $s = k = -1$, there are two critical points for ϕ if $v = 0$, $\|u\| > 1$, and there is a circle of critical points for ϕ if $v = 0$ and $|u| < 1$.*

In the case where $v = 0$ and $|u| < 1$ this means there are infinitely many critical points. This result establishes the conjecture of Maxwell for the case of three collinear point charges with equal charges. In this paper we extend Tsai's work to consider the case where there are four or more collinear point charges, under the assumption one of the outermost point charges has a different sign to all the others. The assumption all point charges are collinear enable us to reduce this to a one-dimensional problem which is possible to solve using techniques from calculus and classical results from the study of roots of real polynomials. As pointed out to us by G. Jennings, S. Raianu and W. Horn, in fact there is an easy calculus argument (see Proposition 2.4) which establishes Maxwell's conjecture in the collinear case when the charges are all positive. This also covers the first case of Tsai's theorem with $n = 3$ and makes his theorem redundant in this case.

The main results are as follows.

Theorem A. *If there are n collinear point charges so that the outermost charge has opposite sign to all others, then ϕ has at most $2(n^2 - 1)$ critical points.*

In particular, this number is finite, but this is not enough to establish Maxwell's conjecture. When $n = 4$, this shows the number of critical points is at most 30. Our main theorem improves this significantly in a special case.

Theorem B. *Suppose there are four collinear point charges such that $-q_1 = q_2 = q_3 = q_4$. Then ϕ has at most twelve critical points.*

This is not optimal: the arguments presented here can be adapted to various other configurations of charges but the general conjecture remains intractable even in the collinear case.When $n = 4$, we have $(4 - 1)^2 = 9$ and so Theorem B does not establish Maxwell's conjecture in this case. If one additionally positions the charges in special locations along the line it is possible to verify the conjecture in many special cases. The reader might note that a natural situation to investigate include when $n = 3$ and the collinear charges satisfy the corresponding assumption $-q_1 = q_2 = q_3$. Then the

problem reduces to finding roots of quartics, which of course can be analysed exactly since there is an explicit formula for finding the roots of a quartic. Our goal in this work however was to investigate a test case where the degree of the polynomials is too large to directly find the roots.

2. PRELIMINARIES

2.1. **Initial Setup and Generalities.** We briefly review some material from multivariable calculus which will be required. Take a function $F : \mathbb{R}^n \to \mathbb{R}^m$, written in components as

$$F(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), f_2(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$$

for $f_i : \mathbb{R}^n \to \mathbb{R}$. Then the Jacobian is the $(m \times n)$ matrix

$$dF = \left(\frac{\partial f_i}{\partial x_j} \right)_{ij}$$

where $1 \le i \le m$, $1 \le j \le n$. Note that if $m = 1$, the Jacobian is a $1 \times n$ row matrix which is exactly ∇F.

Definition 2.1. A map $F : \mathbb{R}^3 \to \mathbb{R}^3$ is said to be a *change of coordinates* if the Jacobian dF has non-zero determinant.

Proposition 2.2. *Critical points are invariant under changes of coordinates.*

Proof. Letting $\mathbf{x} = (x, y, z)$ be the original coordinate system, and $\tilde{\mathbf{x}} = F(\mathbf{x})$ be the new coordinate system. Then $\phi(\tilde{x}) := \phi \circ F^{-1}(\tilde{x})$, so the chain rule states that

$$d\phi(\tilde{\mathbf{x}}) = (d\phi(\mathbf{x}))(dF^{-1})$$

Since F is a change of coordinates, $(dF^{-1}) = (dF)^{-1}$ and so is nonsingular (i.e. it has non-zero determinant). Hence $d\phi(\tilde{\mathbf{x}}) = 0$ if, and only if, $d\phi(\mathbf{x}) = 0$. $\qquad\square$

We assume throughout the point charges are located at distinct points collinear on the x-axis $\mathcal{L}$. Choosing a parametrization of $\mathcal{L}$ we label the point charges in the order we meet them as we vary the parameter. The first point charge, located at $\mathbf{x}_1$, is called the outermost point charge, and we change coordinates if necessary to ensure this happens at $(0, 0, 0)$ with all subsequent charges on the positive x-axis.

Proposition 2.3. *Suppose that either (i) all q_i have the same sign or (ii) q_1 has the opposite sign to $q_2, \ldots, q_n$. Then the critical points of ϕ also lie on $\mathcal{L}$.*

Proof. Choosing suitable coordinates (i.e. scaling and rotating $\mathcal{L}$ appropriately), the point charges may be assumed to lie at the points $(x_1, 0, 0) = (0, 0, 0)$ and $(x_i, 0, 0)$, $i = 2, \ldots, n$. Thus the potential function is given by

$$\phi(\mathbf{x}) = \sum_{i=1}^{n} \frac{q_i}{\sqrt{(x - x_i)^2 + y^2 + z^2}}, .$$

To find the gradient of $\phi(\mathbf{x})$, taking the derivatives with respect to x,y and z yields :

$$\phi_x = \sum_{i=1}^{n} \frac{-q_i(x - x_i)}{[(x - x_i)^2 + y^2 + z^2]^{\frac{3}{2}}}$$

$$\phi_y = y \sum_{i=1}^{n} \frac{-q_i}{[(x - x_i)^2 + y^2 + z^2]^{\frac{3}{2}}}$$

$$\phi_z = z \sum_{i=1}^{n} \frac{-q_i}{[(x - x_i)^2 + y^2 + z^2]^{\frac{3}{2}}}.$$

(i) Setting $\phi_y = \phi_z = 0$, it is apparent that $y = z = 0$ and the result is established since the sum on the right-hand-side consists of non-zero terms each with the same sign.

(ii) Setting $\phi_y = \phi_z = 0$, it is apparent that y = z = 0 and the result is established unless

$$\sum_{i=1}^{n} \frac{-q_i}{[(x - x_i)^2 + y^2 + z^2]^{\frac{3}{2}}} = 0 \implies \sum_{i=1}^{n} \frac{-q_i x}{[(x - x_i)^2 + y^2 + z^2]^{\frac{3}{2}}} = 0.$$

Feeding this into the equation $\phi_x = 0$ and noting that $x_1 = 0$ yields

$$\sum_{i=2}^{n} \frac{q_i x_i}{[(x - x_i)^2 + y^2 + z^2]^{\frac{3}{2}}} = 0$$

Since $x_i \in (0, \infty)$ and all the q_i with $i \geq 2$ have the same sign, this is a contradiction. Hence $y = z = 0$ and the result is established. $\qquad\square$

In case (i), the following elementary argument establishes precisely the number of critical points and implies Maxwell's conjecture holds.

Proposition 2.4. *Suppose there are n collinear point charges with all charges having the same sign. Then ϕ has $n - 1$ critical points.*

Proof. Assume, without loss of generality, all the q_i are positive. Applying Proposition 2.3, the critical points of ϕ lie on the x-axis $\mathcal{L}$. Since $y = z = 0$ we can rewrite $\phi(\mathbf{x})$ as a function of x along, yielding

$$\phi(x) = \sum_{i=1}^{n} \frac{q_i}{|x - x_i|}.$$

Assume $x \in (x_{j-1}, x_j)$, where $j = 2, \ldots, n$. Then

$$\phi(x) = \sum_{i=1}^{j-1} \frac{q_i}{x - x_i} + \sum_{i=j}^{n} \frac{q_i}{x_i - x}$$

$$\phi'(x) = \sum_{i=1}^{j-1} \frac{-q_i}{(x - x_i)^2} + \sum_{i=j}^{n} \frac{q_i}{(x_i - x)^2}$$

$$\phi''(x) = \sum_{i=1}^{j-1} \frac{2q_i}{(x - x_i)^3} - \sum_{i=j}^{n} \frac{2q_i}{(x_i - x)^3} > 0.$$

Hence $\phi''(x) > 0$ on (x_{j-1}, x_j) and thus has at most one critical point on the interval (x_{j-1}, x_j). However,

$$\lim_{x \to x_{j-1}^-} \phi'(x) = - \lim_{x \to x_j^+} \phi'(x) = \infty$$

so the intermediate value theorem implies there is at least one critical point on this interval, and hence there is exactly one critical point on each interval between the two critical points. This yields $n - 1$ critical points. It is easy to see there are no critical points if $x < x_1 = 0$ or $x > x_n$ as $\phi' \neq 0$ on these intervals, and the claim is established. $\qquad \square$

Of course, since $n - 1 < (n-1)^2$, this implies Maxwell's conjecture.

2.2. **Preliminaries for the Proof of Theorem B.** A key ingredient will be the following classical result concerning the real roots of polynomials. For the purposes of this paper, a polynomial will always be assumed to have real coefficients and be written with the powers in descending order, e.g. $x^6 - 5x^4 + 2x - 3$.

Lemma 2.5. *(Descartes' Rule of Signs) The number of positive real roots of a polynomial $p(x)$ is either the number of sign changes between consecutive (nonzero) coefficients, or is less than this by an even natural number. The number of negative real roots can be found by applying this rule to $p(-x)$.*

The proof is via induction. See [1] for further details. Important special cases are when the number of sign changes are either zero or one. If it is zero, we know the polynomial has no roots, and if it is one then it follows that there is exactly one root.

3. Proof of Theorem A

Proof of Theorem A. We follow the proof from Proposition 2.4. There are $n + 1$ intervals. On each interval we solve $\phi'(x) = 0$ by taking common denominators. This results in a polynomial of degree $2(n - 1)$ whose roots are precisely the critical points of ϕ. So on each interval there are at most $2(n - 1)$ roots and the result follows.

4.1. Setup. In this section we will restrict to the case where $n = 4$ and assume the outermost charge q_1 is located at the origin with $-q_1 = q_2 = q_3 = q_4$. Moreover without loss of generality we assume $-q_1 = 1$ and all the other charges lie on the positive x-axis. Applying Proposition 2.3 (ii) we know that $\phi_x = 0$ and $y = z = 0$. Following Proposition 2.4, we know $\mathbf{x} = (x, 0, 0)$ is a critical point of $\phi(x)$ with $x \in (x_{j-1}, x_j)$ if, and only if,

$$\phi'(x) = \sum_{i=1}^{j-1} \frac{-q_i}{(x - x_i)^2} + \sum_{i=j}^{4} \frac{q_i}{(x_i - x)^2} = 0.$$

Here $j = 2, \ldots, 4$. When $x \in (-\infty, 0)$ and $x \in (x_4, \infty)$ this formula is also correct if suitably interpreted. To simplify the algebra further, we will rescale if necessary so that $x_2 = 1$. This rescaling is a change of coordinates in $\mathbb{R}^3$ and so does not affect the computation of the number of critical points. To make the polynomials which follow easier to read, we will relabel $x_3 = \alpha$ and $x_4 = \beta$.

Now, we notice that the sign of each term on the right-hand-side will depend on the position of x. As we have seen in the proof of Theorem A, the sign distribution will generate a different polynomial of degree at most six, yielding a bound of at most 30 roots. We now will break the argument into cases, depending on the position of $\mathbf{x}$ with respect to the charges.

(1) Suppose $x < 0$. Then x is a critical point of ϕ if, and only if,

$$\frac{-1}{x^2} + \frac{1}{(x - 1)^2} + \frac{1}{(x - \alpha)^2} + \frac{1}{(x - \beta)^2} = 0$$

with $x < 0$.

(2) Suppose $0 < x < 1$. Then x is a critical point of $\phi_x(x)$ if, and only if,

$$\frac{1}{x^2} + \frac{1}{(x - 1)^2} + \frac{1}{(x - \alpha)^2} + \frac{1}{(x - \beta)^2} = 0.$$

Hence there are no critical points of ϕ is this interval and the analysis of this case is finished.

(3) Suppose $1 < x < \alpha$. Then x is a critical point of ϕ if, and only if,

$$\frac{1}{x^2} - \frac{1}{(x - 1)^2} + \frac{1}{(x - \alpha)^2} + \frac{1}{(x - \beta)^2} = 0$$

with $\alpha < x < \beta$.

(4) Suppose $\alpha < x < \beta$. Then x is a critical point of ϕ if, and only if,

$$\frac{1}{x^2} - \frac{1}{(x - 1)^2} - \frac{1}{(x - \alpha)^2} + \frac{1}{(x - \beta)^2} = 0$$

with $\alpha < x < \beta$.

(5) Suppose $x > \beta$. Then x is a critical point of ϕ if, and only if,

$$\frac{1}{x^2} - \frac{1}{(x-1)^2} - \frac{1}{(x-\alpha)^2} - \frac{1}{(x-\beta)^2} = 0$$

with $x > \beta$.

4.2. **Analysis of Cases (1), (3), (4), and (5).** It is necessary to consider combinations of the following polynomials:

$$\begin{aligned}
f(x) &= (x-1)^2(x-\alpha)^2(x-\beta)^2 \\
g(x) &= x^2(x-\alpha)^2(x-\beta)^2 \\
r(x) &= x^2(x-1)^2(x-\alpha)^2 \\
s(x) &= x^2(x-1)^2(x-\beta)^2.
\end{aligned}$$

Expanding these polynomials yields

$$\begin{aligned}
f(x) &= x^6 - ux^5 + vx^4 - tx^3 + qx^2 - jx + \alpha^2\beta^2 \\
g(x) &= x^6 - (2\alpha + 2\beta)x^5 + (\alpha^2 + 4\alpha\beta + \beta^2)x^4 - (2\alpha^2\beta + 2\alpha\beta^2)x^3 + \alpha^2\beta^2 x^2 \\
r(x) &= x^6 - (2\alpha + 2)x^5 + (\alpha^2 + 4\alpha + 1)x^4 - (2\alpha^2 + 2\alpha)x^3 + \alpha^2 x^2 \\
s(x) &= x^6 - (2\beta + 2)x^5 + (\beta^2 + 4\beta + 1)x^4 - (2\beta^2 + 2\beta)x^3 + \beta^2 x^2,
\end{aligned}$$

where

$$\begin{aligned}
u &= 2\alpha + 2\beta + 2 \\
v &= \beta^2 + 4\alpha\beta + 4\beta + \alpha^2 + 4\alpha + 1 \\
t &= 2\alpha\beta^2 + 2\beta^2 + 2\alpha^2\beta + 8\alpha\beta + 2\beta + 2\alpha^2 + 2\alpha \\
q &= \alpha^2\beta^2 + 4\alpha\beta^2 + \beta^2 + 4\alpha^2\beta + 4\alpha\beta + \alpha^2 \\
j &= 2\alpha^2\beta^2 + 2\alpha\beta^2 + 2\alpha^2\beta.
\end{aligned}$$

With this notation, the polynomials are as follows. For Case (1), the rational function has a zero precisely when $(-f + g + r + s)(x) = 0$ with x negative, which simplifies to

$$2x^6 - (2\alpha + 2\beta + 2)x^5 + (\alpha^2 + \beta^2 + 1)x^4 + 8\alpha\beta x^3 - (4\alpha\beta^2 + 4\alpha^2\beta + 4\alpha\beta)x^2 + jx - \alpha^2\beta^2 = 0.$$

For Case (3), the rational function has a zero precisely when $f - g + r + s = 0$, which simplifies to

$$\begin{aligned}
0 = {} &2x^6 - (2\alpha + 2\beta + 6)x^5 + (\beta^2 + 8\beta + \alpha^2 + 8\alpha + 3)x^4 \\
&- (4\beta^2 + 8\alpha\beta + 4\beta + 4\alpha + 4\alpha^2)x^3 \\
&+ (4\alpha\beta^2 + 4\alpha^2\beta + 4\alpha\beta + 2\alpha^2 + 2\beta^2)x^2 \\
&- jx + \alpha^2\beta^2
\end{aligned}$$

with $1 < x < \alpha$.

For Case (4), the rational function has a zero precisely when $f - g - r + s = 0$, which simplifies to

$$0 = 2(\alpha - \beta - 1)x^5 + (\beta^2 - \alpha^2 + 8\beta + 1)x^4 - (4\beta^2 + 8\alpha\beta + 4\beta)x^3$$
$$+ (4\alpha\beta^2 + 2\beta^2 + 4\alpha^2\beta + 4\alpha\beta)x^2 - jx + \alpha^2\beta^2$$

with $\alpha < x < \beta$.

Finally in Case (5) the polynomial has a zero precisely when $f - g - r - s = 0$, which is (up to a minus) the same case as Case 1. Hence we seek to solve

$$2x^6 - (2\alpha + 2\beta + 2)x^5 + (\alpha^2 + \beta^2 + 1)x^4 + 8\alpha\beta x^3 - (4\alpha\beta^2 + 4\alpha^2\beta + 4\alpha\beta)x^2 + jx - \alpha^2\beta^2 = 0$$

but now with $x > \beta$.

4.3. Analysis of Case 1.
We claim there is one negative real root of the polynomial $z = -f + g + r + s$. This must have one negative root by the Intermediate Value Theorem, since $z(0) < 0$ and $z(x) \to \infty$ as $x \to -\infty$. In fact, if we apply Descartes rule of signs to estimate the number of negative zeros, we obtain that $z(-x)$ is

$$2x^6 + (2\alpha + 2\beta + 2)x^5 + (\alpha^2 + \beta^2 + 1)x^4 - 8\alpha\beta x^3 - (4\alpha\beta^2 - 4\alpha^2\beta + 4\alpha\beta)x^2 - jx - \alpha^2\beta^2.$$

This has one sign change so this proves there is exactly one negative root of z as required.

4.4. Analysis of Case (3).
Applying Descartes rule of signs to the polynomial $z(x) = (f - g + r + s)(x)$ does not yield any information in this case: it tells us there are either six, four, two or zero positive roots of the polynomial (and no negative roots). So we conclude there are at most six critical points of ϕ with $x \in (1, \alpha)$.

4.5. Analysis of Case (4).
Applying Descartes's rule of signs to

$$z(x) = 2(\alpha - \beta - 1)x^5 + (\beta^2 - \alpha^2 + 8\beta + 1)x^4 - (4\beta^2 + 8\alpha\beta + 4\beta)x^3 + (4\alpha\beta^2 + 2\beta^2$$
$$+ 4\alpha^2\beta + 4\alpha\beta)x^2 - jx + \alpha^2\beta^2$$

we obtain there are at either five, three or one positive real roots. Hence, in a worst-case scenario, there are five roots in $(1, \alpha)$.

4.6. Analysis of Case (5).
Here the argument presented in Case (1) does not work as Descartes' rule of signs is not immediately applicable, as the reader can check directly. To remedy this, we apply the change of coordinates

$$x \to \tilde{x} = \beta - (\beta - \alpha)x, \quad \tilde{y} = y, \quad \tilde{z} = z.$$

The reader may easily check the determinant of the corresponding Jacobian is $\alpha - \beta \neq 0$, so it is invertible and hence this transformation is indeed a change of coordinates. From Proposition 2.2, being a critical point is independent of coordinate system. Computing in the new coordinate system, $\phi(\tilde{\mathbf{x}})$ is the same polynomial as Case (1), but in the $\tilde{x}$ variable (with α replaced with $\tilde{\alpha}$ etc.). To see this, note that $\beta < x < \iff \tilde{x} < 0$. In these coordinates, ϕ has a critical point if, and only if, the polynomial $z(\tilde{x}) = (f + g + r - s)(\tilde{x}) = 0$ with $\tilde{x} < 0$. Notice the minus sign now lies before the last term

since the point charge with opposite sign now lies at $\tilde{\beta}$. Expanding this polynomial in the same fashion as above yields

$$\begin{aligned}
z(\tilde{x}) = {} & 2\tilde{x}^6 - (6\tilde{\alpha} + 2\tilde{\beta} + 2)\tilde{x}^5 + (3\tilde{\alpha}^2 + \tilde{\beta}^2 + 8\tilde{\alpha}\tilde{\beta} + 8\tilde{\alpha} + 1)\tilde{x}^4 \\
& - (4\tilde{\alpha}^2\tilde{\beta} + 4\tilde{\alpha}\tilde{\beta}^2 + 4\tilde{\alpha}^2 + 8\tilde{\alpha}\tilde{\beta} + 4\tilde{\alpha})\tilde{x}^3 \\
& + (2\tilde{\alpha}^2\tilde{\beta}^2 + 4\tilde{\alpha}\tilde{\beta}^2 + 4\tilde{\alpha}^2\tilde{\beta} + 4\tilde{\alpha}\tilde{\beta} + 2\tilde{\alpha}^2)\tilde{x}^2 \\
& - \tilde{j}\tilde{x} + \tilde{\alpha}^2\tilde{\beta}^2.
\end{aligned}$$

Solving $z(\tilde{x}) = 0$ with $\tilde{x} < 0$ yields, upon applying Descartes' rule of signs in the same fashion, that there is no critical point in this interval.

4.7. Conclusion of the Proof.

Proof of Theorem B. This now follows immediately from combining established facts. There is one critical point in $(-\infty, 0)$, none in $(0, 1)$ and (β, ∞), at most five in (α, β), and at most six in $(1, \alpha)$. Hence there are at most twelve critical points. $\square$

5. Final Remarks

We conclude with some final remarks. The obvious way to improve the bound of 13 we obtain is to apply Sturm's lemma to precisely calculate the number of roots inside each interval, but the algebra becomes very difficult. This is a project for the future.

At first glance, it is somewhat curious that Descartes' rule of signs does not work in Case (5), but after a change of coordinates it can be made to work. This can be explained by observing that the change of coordinates presented changes the direction of the positive x-axis. To illustrate, consider the case where there are two charges in $(0, \beta)$, and three in (β, ∞). Then there are five positive roots in the original coordinate system on the positive x-axis, but only two in the $\tilde{x}$ coordinates. This example illustrates that it is possible to change the number of positive and negative roots by applying changes of variables.

It is also natural to ask how much further the collinear case can be investigated. We hope to take this project up into the future. In particular, it is clear that the assumption all the charges lie in a line allow us to reduce the equations to finding roots of polynomials of higher degree, and Descartes' Rule of Signs is a useful tool to help us compute the number of roots.

References

[1] Anderson, B., Jackson, J., and Sitharam, M. *Descartes' Rule of Signs Revisited*, The American Mathematical Monthly 105, no. 5 (1998), 447-451.

[2] Gabrielov,A, Novikov, D. and Shapiro, B. *Mystery of point charges*, Proc. London Math. Soc. (3) vol 95, no. 2 (2007) 443–472.

[3] Maxwell, J.D. *A treatise on electricity and magnetism*, Vol. I. Oxford Classic Texts in the Physical Sciences, The Clarendon Press, Oxford University Press, 1998.

[4] Morse, M. and Cairns, S. *Critical Point Theory in Global Analysis and Differential Topology*, Acad. Press, 1969

[5] Shapiro, B. *Problems Around Polynomials: The Good, The Bad and The Ugly. . .*, Arnold. Math. J., 1, no. 1 (2015), 91–99.

[6] Tsai, Y.L. *Maxwell's conjecture on three point charges with equal magnitudes*, Phys. D 309 (2015), 86–98.

DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY FULLERTON, 800 N. STATE COLLEGE BLD., FULLERTON, CA 92831, USA.

Email address: solomonhuang0703@csu.fullerton.edu

Email address: duynguyen1993@csu.fullerton.edu

Physics
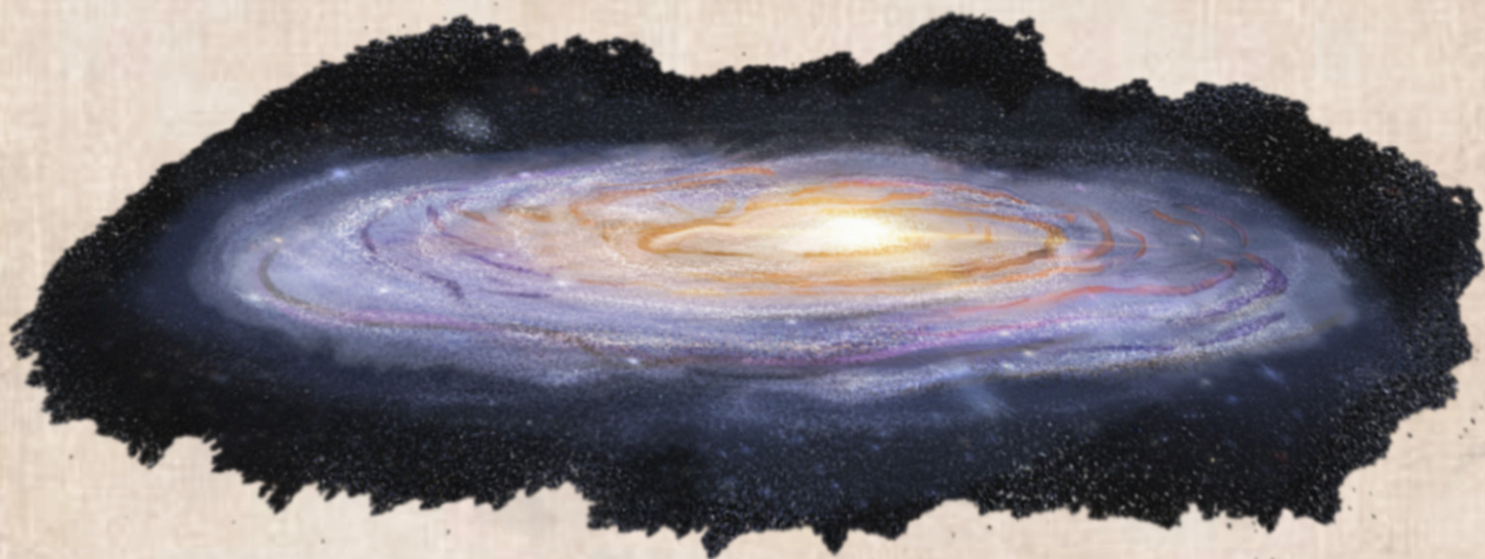
Ordinary Differential Equations in Upper-Division Physics: Identifying Productive Strategies via a Concept Image Framework

Author: Anderson T. Fung
Advisor: Dr. Michael E. Loverude
California State University, Fullerton
Department of Physics

Ordinary differential equations (ODEs) are an important mathematical tool for modelling in physics. In addition to being fluent with the symbols and structures of mathematics, physicists entering the advanced undergraduate sequence of courses are also expected to encode physical meaning into said symbols/structures. Through ungraded pre-instruction surveys in a math methods course, it is evident that this expectation is not an easy one. Previous research by Hyland has shown that while students may have been exposed to the theoretical aspects of ODEs in previous math courses (especially the mechanics of obtaining a solution), instructor intervention is needed to transform this theoretical knowledge into a form that is useful for physics. Through a theoretical model of student thinking developed by Tall and Vinner known as the concept image framework, we analyzed three semester's worth of these pre-instruction surveys and found that students who utilized a strategy that did not focus on the formal function definition of an ODE tended to be unproductive towards completing physics ODE tasks. This implies that an instructional shift towards a greater emphasis of general mathematical principles might encourage more productive strategies in using ODEs to model physical systems.

I. INTRODUCTION

Ordinary differential equations are ubiquitous throughout physics due to their ability to describe/model dynamic physical systems. This ubiquity means that any student taking advanced undergraduate physics courses (upper-division) should be comfortable with

(a) setting up an ODE for a given physical system

(b) analyzing the structure of an ODE

(c) determining a method of solution for the ODE.

This paper will explore how students at CSUF who are entering the upper-division use these three vital skills to solve novel physics problems.

A. Formal Definition of an ODE

For most students, an ODE and its resulting solution is defined rather succinctly as such:

"A DE is any equation that involves one or more derivatives of an unknown function. The variable which this function depends on is called the **independent variable** while the resulting outputted value is the **dependent variable**. The unknown function becomes a **solution** to the DE if it takes a form such that when substituted into the DE produces an identity."

B. Course Structure

Students at CSUF are initially exposed to ODEs in second-semester introductory calculus (MATH 150B). As is necessary for an introductory course, 150B focuses heavily on grouping ODEs into classes (separable, first-order linear, etc.) and developing/practicing an algorithmic method of solution for these classes.

ODEs are further explored in a hybrid differential equations/linear algebra semester-long course (MATH 250B) as well as a dedicated upper-division ODE course (MATH 310). However, these are not prerequisites and most CSUF physics majors have not taken either upon entering the upper division.

In order to augment and re-frame this initial exposure to ODEs in a fashion that is more useful to physics, physics majors are required to take a *math methods* course (PHYS 300) before embarking on

most upper-division courses. Unfortunately, ODEs are only a subset of the mathematics PHYS 300 attempts to review/introduce; instructional time typically ranges from 2-3 weeks for most semesters.

C. Different Presentations: An Example from Classical Mechanics

A thorough knowledge of ODEs is often not necessary to succeed in the introductory CSUF physics sequence. As an example, in the textbook used for introductory mechanics (PHYS 225), the position function $x(t)$ of a simple harmonic oscillator (SHO) is developed as such:

1. A particle of mass m that is subjected only to a linear restoring force (i.e., by Hooke's law) kx is described by the following equation of motion.

$$ma = -kx \tag{1}$$

where a is the acceleration of the block.

2. This gives rise to the following second order differential equation.

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -\omega^2 x \tag{2}$$

where ω is the angular frequency of the oscillating mass.

3. The solution to this differential equation is the following position function

$$x(t) = A\cos(\omega t + \phi) \tag{3}$$

which is commonly referred to as the *sinusoidal function*.

In the scope of an introductory course, elimination of the method of solution from which (2) is derived from (1) is necessary in order to ensure enough time is allocated to exploring the physical implications of the sinusoidal solution. However, in an upper-division classical mechanics course, it is important to spend some time on the method of solution. For example, the textbook used in CSUF's upper-division mechanics course (PHYS 320) spends nearly seven pages going through the method of solution as well as rewriting the sinusoidal solution into equivalent exponential and phase-shifted forms.

For many students entering the upper-division, such a transition is not trivial. This observation is consistent with recent work done by Hyland [1] who found that while students who have largely been exposed to the *theoretical aspects* of ODEs in previous courses, e.g., the mechanics of obtaining a solution, the process of using ODEs to build a physical model required instructor intervention.

II. THEORETICAL FRAMEWORK: CONCEPT IMAGE

In order to present student understanding of ODEs in a rigorous fashion, i.e., one that is intelligible to the larger physics education research (PER) community, it is often appropriate to do so with respect to a *theoretical framework*—an idealized model of student thinking. A framework which has been used to present student understanding on other mathematical topics such as unit vectors (Schermerhorn and Thompson) [2] and Fourier transforms (Mays and Loverude) is the **concept image framework**.

A. Overview of Framework

Initially developed as a model of student understanding in the mathematics education research community by Tall and Vinner the framework proposes that

– a student's collective understanding of a topic may be modeled as a **concept image**.

– the concept image, in its totality, is often abstract/difficult to describe and subject to change via outside sources.

– External events that task the student with applying a mathematical concept, e.g., an activity or exam question will cause a portion of the student's concept image to be activated or *evoked*.

– Ideas from the evoked concept image (**concept definitions**) are then used to complete the task at hand. If these ideas are incompatible with the task's definition (**formal definition**), then the student is unable to complete the task.

– Each student's concept image and the resulting concept image is unique, i.e., based on individual experiences both inside and out of the current curriculum.

B. Productive and Unproductive Definitions: A Simple Example

As a simple application of this framework, consider the topic of the limit of a function. A question about the long-term behavior of a RC circuit might evoke in a student the idea that a limit is "making t large for a given function". If the student were to

apply the same concept image towards a task that required the student to approximate an equation by allowing the mass to be infinitesimally small, i.e., $m \to -\infty$, the student would be unable to complete the task.

It is important to stress here that the student's concept definition of the limit is not "wrong". In fact, it is accurate as long as the task requires allowing the independent variable of a function to tends toward infinity. A concept definition that is not appropriate for a certain task is simply **unproductive**. However, this concept definition (while not strictly adhering to a rigid formal definition) might prove **productive** in many other scenarios.

III.   METHODOLOGY

Through the concept image framework, it is possible to not only identify productive student strategies for a particular task, but discover strategies that are more generalizable. More specifically, we were interested in whether or not students who used strategies that were more strongly associated with the formal function definition of an ODE were more productive.

A.   Data Set

The data for our project comes entirely *ungraded quizzes* (UQs) from three sections of PHYS 300: Spring 2018 ($N = 13$), Fall 2018 ($N = 13$), and Spring 2019 ($N = 14$). For all three sections, the UQs were administered with no formal instruction past some surface-level discussion about the importance of ODEs in physics. This is done as a countermeasure to prevent unnecessary biasing towards a particular strategy. The tasks in the actual UQs varied slightly between semester, so some tasks will have a larger population than others.

We manage to identify student strategies and classify them as productive or unproductive by developing an **emergent coding scheme**, i.e., assigning a label or *code* to a student response. This allows for the identification of how common or uncommon a certain strategy is relative to other data sets despite differences in task question.

B.   Survey Tasks

We grouped the tasks in these three sections of UQs into three broad categories.

1.   Identification of the Unknown Function

The following set of tasks asked students to identify the unknown function, i.e., the independent and dependent variables in a differential equation. In math classes, identifying these variables is usually not a difficult task as y typically represents a dependent variable and x or t the independent. These tasks are designed to see if students were able to carry out the same process when these symbols carried physical meaning or *contextualized* via a physical model.

For example, in the UQ for the Fall 2018 section, students were given the equation of motion for a spring-mass system that is oscillating in a medium that provides a linear drag force and the following equation of motion

$$m\ddot{x} = -kx - b\dot{x} \tag{4}$$

and were asked to identify an appropriate *de-contextualized* form out of three possible choices—the expected one being

$$a_2 y'' + a_1 y' + a_0 y = 0 \tag{5}$$

Alternatively, as in the case of the Spring 2019 section, students were required to rewrite a given physical model in terms of a de-contextualized DE. For example, in one task students were given a slightly modified version of (4) and asked to rewrite it into a DE of the form

$$\frac{d^2 y}{dt^2} + \alpha \frac{dy}{dt} + \beta y = 0 \tag{6}$$

Both of these tasks are similar as they required the student to identify the independent and dependent variable first before being able to take the requisite derivatives.

2.   Tasks Involving Functional Relationships

Another set of tasks revolved around using a given ODE to describe a plausible physical scenario. In all tasks, students were free to define what each variable meant physically. For example, in the 2018 UQ, students were given the following DE

$$\frac{dI}{dt} = -I\lambda \tag{7}$$

where we have deliberately chosen some unconventional variables. In order to construct a physical model, the student would've had to recognize the functional relationships shown by the DE. Specifically for (7), the student would've needed to recognize that the physical model needed to have a quantity that is negatively proportional to its rate of change.
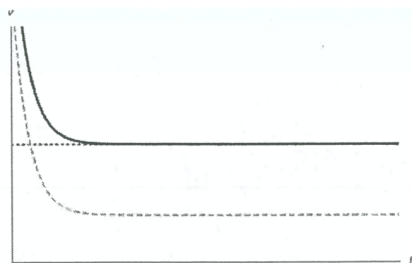
FIG. 1. Graph for the Solutions Task

3. Interpretation of Solutions

Our final set of tasks involved looking at student's strategies with interpretation the solution of an ODE. Here students were given a solution to a DE describing a particle falling downwards in a medium that provides a linear drag force.

$$v(t) = v_{\text{ter}} + Ae^{-t/\tau} \tag{8}$$

where v_{ter} is the terminal speed of the particle. Alongside (8), students were also presented with a graph of the solution (Figure 1) Students were then asked to comment on the long-time behavior of the particle and asked to explain their reasoning. The purpose of this task is to identify whether students would choose to explain physical phenomena via a function or via another method such as a graph when given the option.

IV.   RESULTS

Throughout all three tasks, we found that strategies that involved the usage of the formal function definition were the most productive. While strategies that relied on other means worked for a localized task, these strategies were often not generalizable. Through our emergent coding scheme we identified several examples where this localized strategy can cause student confusion.

A.   Pattern-Matching

A common strategy among student for the first set of tasks was to "pattern-match"—finding analogous mathematical structures such as similar-order derivatives and using these structures to classify DEs. This strategy proved productive for tasks where the equation of the physical model heavily resembled the decontextualized equation (for example, patter-matching worked well with relating (4)

and (5)) (Figure 2) This is not a controversial claim

FIG. 2. Pattern-matching works well for equations where constants and functions were clearly delineated relative to the decontextualized equation.

as in the timed environment the UQ was given in, pattern-matching is a significantly faster strategy to employ than identifying what the independent and dependent variables were. Students who employed this strategy for significantly more complicated equations, i.e., with more constants that required algebraic rearrangement found it to be unproductive (Figure 3). As figure 2 shows, pattern-

FIG. 3. Pattern-matching was unproductive when multiple constants.

matching results in the derivative of the independent variable with respect to the dependent y' being falsely identified as $L(\mathrm{d}I \,/\, \mathrm{d}t)$ with the additional constant being "mixed in".

In contrast, a function-based strategy would've first identified I as a variable that is dependent on t and probably would not have had led to the same error as shown in Figure 3.

B.   Familiar Symbolization

A common strategy we found regarding the second set of tasks was to judge the validity of the given ODE based on the presence of familiar symbols that were encountered in previous study. For example, many students responded that (7) was probably a DE describing "something to do with wavelength" without specifying further.

We conjecture that this response was probably not related to the actual functional relationships present in (7). Rather, the response was motivated by the presence of λ—a common symbol used for wavelength in the physical sciences. Another example

of the strategy of familiar symbolization being unproductive for these set of tasks is when students were asked to identify a plausible physical scenario for the following DE

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} = -\alpha\theta \qquad (9)$$

A common response among students was similar to that in Figure 4. The presence of θ—commonly as-

B. $\frac{d^2\theta}{dt^2} = -\alpha\theta$

Physical situation: Circular motion

FIG. 4. Familiar symbolization often prompted students to respond with a brief mention of circular motion.

sociated with rotational motion—seemed to prompt many of these responses. Again, a function-based approach would've probably seen that (9) is actually the same equation as the DE describing our SHO (1).

C. Interpretation of a Solution

For our final set of tasks, we found that when given the option to describe the long-term behavior of a falling particle, students overwhelmingly chose to use the provided graph over taking the limiting case of the provided solution (8). While not unproductive for the current task, the response rate could be an indication of a preference for using graphs and physical intuition to model a physical system rather than a mathematical function approach.

V. CONCLUSION: NOVICES VS. EXPERTS

The responses from our data set are in line with work done by PER researchers Chi, Feltovich, and Glaser [5] who identified that introductory students (**novices**) tend to interpret a problem based on more immediate salient structures (**surface features**). Most students in our data set have shown this by gravitating away from the formal function definition in favor of surface features such as the presence of certain derivatives or familiar symbols when approaching many of these tasks. More practiced physicists (**experts**) however tend to approach problems by reducing it to more fundamental principles. A question of further research would be whether or not expert physicists regularly invoke the formal function definition of the ODE when productively solving these tasks. In order to confirm this,

we will have to move beyond written responses in lieu of formal interviews.

VI. IMPLICATIONS FOR INSTRUCTION

Our data set implies that strategies that evoke the formal function definition of an ODE allow for the successful resolution of these tasks. In terms of instruction, these observations serve as an early indication that math method courses such as PHYS 300 should more greatly emphasize ODEs not as separate classes of decontextualized math problems that may be solved via a prerecorded algorithm but as the mathematical relationship between a function and its derivatives.

Further, upper-division courses should not assume that student mastery of solving decontextualized ODEs will translate well into problems requiring ODEs in a physical context. Allocating more time to demonstrating how ODEs are used in modeling a physical system might encourage more productive function-based strategies.

VII. REFERENCES

1. D. Hyland, Investigating students' learning of differential equations in physics, PhD Thesis, Dublin City University, 2018.

2. B.P. Schermerhorn and J.R. Thompson, presented at the Physics Education Research Conference 2017, Cincinnati, OH, 2017.

3. M. Mays and M.E. Loverude, presented at the Physics Education Research Conference 2018, Washington, DC, 2018.

4. D. Tall and S. Vinner, Educ. Stud. Math. 12, 151 (1981).

5. M. Chi, P.J. Feltovich, & R. Glaser (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5(2), 121–152.

VIII. ACKNOWLEDGEMENTS

Examining Student Understanding of Matrix Algebra and Eigentheory

Author: Pachi Her
Advisor: Dr. Michael E. Loverude
California State University, Fullerton
Department of Physics

As part of a project involving research on student use of mathematics in upper division physics, we conducted a study on the student understanding of matrix algebra in the context of a mathematical methods course for physics majors. We studied student written responses to three free-response exam questions to reach our conclusions. Each question was organized in accordance to applicable skills needed to solve these questions. Our data suggests that the root of student difficulty in constructing a matrix equation from a given set of systems of equations derived from their struggle to comprehend and interpret the mathematical language. Along with this difficulty, students often misused and misinterpreted the meanings of variables; for example confusing one variable such as the eigenvalue for another, the frequency. In conclusion, the misinterpretation of variables and difficulty translating physical systems to mathematical equations correlate to the student comprehension of matrix algebra and in a physics context.

I. INTRODUCTION

Mathematics is a major tool used in physics, and it assists in structuring the framework of which a physical system exists and operates. A commonly used mathematical branch is linear algebra, which encompasses concepts such as vectors, matrices, and linear equations. In this study, we examined student understanding of matrix algebra and eigentheory.

A. Previous Works

Several studies have consistently observed that linear algebra is a difficult subject for students to apply. In the Physics Education Research community, there is limited research on student understanding of linear algebra applications. A majority of the studies mainly focused on the context of quantum mechanics [1,2]. Therefore, we will be referencing most of the studies done in the math education community, RUME.

Drawing from other research, Sabella and Redish identified from various research papers that the difficulty in learning linear algebra is due to the fact that in these courses, students may "master the algorithmic skills. . . but lack a conceptual understanding. . . and (of) how to apply linear algebra concepts to physical systems"[3]. The Linear Algebra Curriculum Study Group, formed by Carlson et al, is a group that focuses on improving linear algebra courses to increase the student understanding of the subject [4].

Previous studies by Larson and Zandeih investigated multiple ways of interpreting matrix equations [5] and using the framework Larson created [6] of the different interpretations of matrix multiplication. The students in their study successfully solved matrix multiplication but utilized a different method in their solutions.

In addition to studies that investigated how students interpret matrix equations, studies have been conducted on how students form symbolic reasoning of eigentheory. Henderson et al compared student comprehension between numerical values and symbols in the form of the eigenequation. They identified three categories of student reasoning: (1) students who used superficial algebraic cancellation, (2) students who correctly solved the system and correctly interpret their result, and (3) students who correctly solved the system and incorrectly interpreted their results [7]. Another study showed that students tend to focus on the calculation procedures instead of grasping the underlying meaning of what was being calculated and modeled [8]. The results of these studies can foreshadow the struggle of applying matrix algebra and eigentheory for students in physics courses.

B. Motivation for Study

As stated in the previous section, there are only a few studies of linear algebra in various contexts of physics, other than quantum mechanics. Thus, further investigation of student understanding of matrix algebra and eigentheory in the context of physics can prove to be valuable to researchers, instructors, and students. It will be able to provide information on areas that students struggle the most and emphasize areas of instruction that should be improved given that the mathematics education community has provided through their research of student difficulty in linear algebra.

II. RESEARCH METHODS

The first step to our study was gathering the student written responses from an upper division physics course known as Survey of Mathematical Physics. This sophomore-junior level course teaches students various mathematical concepts observed in physics. Enrolled students have already taken and passed Calculus III and therefore, already have familiarity with some of the concepts taught in class. A majority of the students enrolled in the course were either physics majors or minors. An important factor to consider in this study was that the students were not required to take a linear algebra course before enrolling, thus their prior knowledge of linear algebra could be limited.

We investigated three exams questions from two different semesters, Fall 2017 (N = 14) and Spring 2019 (N = 14). Two of the questions came from a final exam given in Spring 2019, and the other in Fall 2017. The questions were chosen based on their similarity of tasks.

Skills	Characteristics	Example
Mathematization	• Translation into mathematical terms (*e.g., an equation*) • Mathematical sense making – provide an analogy and reasoning between mathematical and physical systems	$-m\omega^2 x_1 = -4kx_1 + kx_2$ $-m\omega^2 x_2 = 3kx_1 - 5kx_2$ $-m\omega^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} -4k \\ 3k \end{pmatrix} + x_2 \begin{pmatrix} k \\ -3k \end{pmatrix} = \begin{pmatrix} -4k & k \\ 3k & -3k \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $-m\omega^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -k \begin{pmatrix} 4 & -1 \\ -3 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
Interpretation	• "Reading" Equations • Evaluating the meaning of the expression, symbols, and values	The displacements from the equilibrium position for each mass. So, for $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1cm \\ -2cm \end{pmatrix}$, this means that mass 1 is 1 cm above its equilibrium position, and mass 2 is 2 cm below its equilibrium position.
Technical Operation	• Calculating for unknown values • Use of operations (multiplication, division, subtraction, and addition) and equations	$\det\begin{vmatrix} 4-\lambda & -1 \\ -3 & 3-\lambda \end{vmatrix} = (4-\lambda)(3-\lambda) - 3 = 12 - 7\lambda + \lambda^2 - 3 = \lambda^2 - 7\lambda + 9 = 0$

TABLE I. Description of each skill along with an example from a student's written response for Question 1, Fall 2017

Solving physics problems generally requires a series of steps and skills before reaching a solution. The Physical-Mathematical Model created by Uhden et al, displays three skills that students use when solving math- integrated physics problems [9]. These skills distinguish the relationship between structural skills and technical skills, hence the split of the two graphs as shown in Figure 1. Structural skills involve the conceptual understanding of the physics mathematically, whereas technical skills involve the application of mathematical operations.

The three skills, mathematization, interpretation, and technical operation, are described based on the degree of mathematization, or the level of math language the skill can acquire. For example, the mathematization skill is the ability of interpreting and categorizing a given physical system mathematically; in this study, it would be translating the systems of equations that describe the springs system of two masses into a matrix equation. This skill requires the ability to understand what the system of equations is describing and transform them into a matrix equation that also translates into the same system (normal modes). One important thing to note is that the students were not asked to construct the equations of motions instead, were already given them. Thus, the degree of mathematization for this skill had started in the middle, then increasing as the students perform the construction of the matrix equation. Table I provides a more in-depth description of each skill with a student example representing each one.

We used this model as a guidance to code our questions, classifying each question with a skill. This coding scheme helped us locate the students' strengths and weaknesses in their understanding of these mathematical tools.
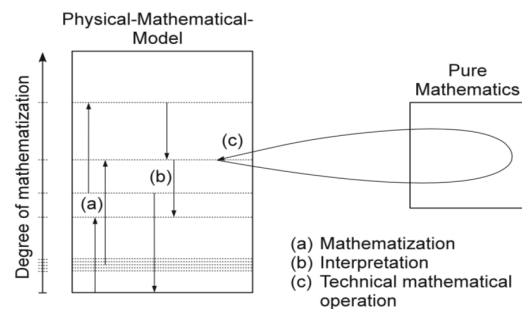

FIG. 1. Representation of the Physical-Mathematical Model created by Uhden et al.

A. Question Design

Before explaining the design of each question, we will introduce two equations that students were expected to use for solving these questions: (1) the eigenequation and (2) the characteristic equation. As discussed further below, the students were asked to construct a matrix equation and relate it to the eigenequation. This relationship demonstrates which elements from the matrix equation corresponds to the elements in the eigenequation, consequently, forming an understanding of how the physical system works mathematically. The characteristic equation involves several procedures such as finding

the determinant of the operation inside the parenthesis, that the students had to perform first, and then factor out lambda to determine its value(s).

$$\hat{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{1}$$

$$det(\hat{A} - \lambda\hat{I})\boldsymbol{x} = \boldsymbol{0} \tag{2}$$

1. Question 1: Two masses attached to springs

Two masses attached to springs is a common question given in a classical mechanics course. In this question, students were to observe the two normal modes system with a given a set of equations of motion. Question 1 was made up of three parts. In part (i), the students had to translate the system of equations into a matrix equation; in part (ii), they had to interpret the physical meaning of the column vector, $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, and in part (iii), solve for the eigenvalue problem for the frequencies of the two normal modes.

We identified that part (i) incorporated students' mathematization skills, part (ii) incorporating interpretation, and part (iii) both technical operation and interpretation. The mathematization in part (i) is activated when students have to interpret the given equations of motions mathematically, and then create a new mathematical equation (matrix) that would describe the same physical system. Interpretation plays a partial role when students apply mathematization, where they have to properly interpret the meaning of each element in the eigenequation and connect them to the elements given in the system of equations. Identifying the interpretation skill in part (ii) was easy as it directly asked students for the physical interpretation of the column vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

For part (iii), students had to apply technical operation when solving for the eigenvalue using the characteristic equation (Equation 2). After factoring out lambda, the students had to interpret this eigenvalue by recognizing what it described which was, $\frac{m\omega^2}{k}$, in order to find the frequencies, ω, of the two normal modes.

It is crucial for students to successfully answer part (i), as parts (ii) and (iii) could be difficult to solve and understand in

---

Fall 2017

Question 1:
Describe the position of the upper block as x_1 and the position of the lower block as x_2. Both x_1 and x_2 are measured so that positive values are up.

The equations of motion for the blocks are: $m\ddot{x}_1 = -4kx_1 + kx_2$
$$m\ddot{x}_2 = 3kx_1 - 3kx_2$$

Assume normal mode motion with the same frequency so that $x_1 = x_1 e^{-i\omega t}$ and $x_2 = e^{-i\omega t}$.
  i. Rewrite the system of equations for the blocks as a single matrix equation.
  ii. What is the physical meaning of the column vector $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$? (e.g., what does $\begin{bmatrix} 1\ cm \\ -2\ cm \end{bmatrix}$ mean?
  iii. Solve the resulting eigenvalue problem to determine the frequencies of the two normal modes.

FIG. 2. Original Question 1

---

the absence of correctly mathematizing the systems of equations into a matrix equation. This dependency between the parts led to the development of Questions 2 and 3, where parts (i) and (iii) were split into two separate questions.

2. Question 2: Solve eigenvalue problem to determine frequencies

Similar to Question 1 part (iii), Question 2 asked students to find the frequencies of the two normal modes. But instead of having students construct a matrix equation, they were given one. We also identified this question incorporating both *technical operation* and *interpretation* as it also addresses students to utilize the characteristic equation and decipher what the eigenvalue corresponds to.

Spring 2019

Question 2:
In a normal mode problem with two springs, the following eigenvalue equation arises:
$$\begin{pmatrix} -3k & k \\ k & -3k \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = m\omega^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Determine the frequencies of the two modes.

FIG. 3. Original Question 2

---

3. Question 3: Translate system of equations into a matrix equation

Question 3 is similar to Question 1 part (i), where it asked students to observe a physical system with a given system of equations, and transform it into a matrix equation. Instead of observing two masses connected by springs, the students had to observe a couple LC circuit. It is important to note that prior to this exam, the students had practiced applying matrix algebra and eigentheory exclusively on springs but not on other systems. Although the physics context is different, the procedure of translating the system of equations describing the LC circuit into a matrix equation is the same as for question 1.

Spring 2019

Question 3:
A complicated circuit coupled LC branches has the resulting second-order equations for I_1 and I_2, the currents through the two inductors:
$$\frac{d^2}{dt^2}I_1 = -\frac{1}{LC}I_1 + \frac{1}{LC}(I_2 - I_1)$$
$$\frac{d^2}{dt^2}I_2 = -\frac{1}{LC}I_2 + \frac{1}{LC}(I_1 - I_2)$$
Assume normal modes of the current such that $I_1 = I_{1,0}e^{i\omega t}$ and $I_2 = I_{2,0}e^{i\omega t}$.

Write a matrix equation for the system of equations. Describe how it fits the profile of an eigenvalue equation, but DO NOT SOLVE.

FIG. 4. Original Question 3

III. RESULTS AND DISCUSSION

Our data show that students tend to struggle more with the structural skills than the technical skills. Questions 1 and 3 provide evidence of students struggling in applying mathematization to a physical system. Classifying the elements of the physical system into mathematical terms can be difficult as it entails conceptual understanding of what each element in the

Question		Mathematization	Interpretation	Technical Operation
Fall 2017 (N=14)	Question 1 Part (i)	57%	50%	--
	Part (ii)	--	79%	--
	Part (iii)	--	57%	78%
Spring 2019 (N=14)	Question 2	--	79%	92%
	Question 3	43%	--	--

TABLE II. Description of each skill along with an example from a student's written response for Question 1, Fall 2017

system represents, physically and mathematically. We also saw student difficulty with problems that involve symbolic answers rather than numerical as seen in Questions 1 and 3.

We observed the similarities and differences between the tasks of each question and the students' responses. Question 1 part (i) was compared to Question 3, and Question 1 part (iii) to Question 2. Part (ii) of Question 1 was observed by itself since it was the only unique question out of the three.

The percentages in Table II define the number of students who successfully applied the corresponding skills to each question. The number of points that the students received for each question was partially considered in determining whether or not they applied the skills. The data from Table I implied that students struggle the most in applying *mathematization* and *interpretation* more so than *technical operation*. Comparing the percentages between interpretation and technical operation, Question 1 part (iii) has a 20 percent difference and Question 2, a 13 percent difference. These percent differences distinguish the students' performances of applying the two skills. The consistency of the technical operation having a higher percentage show that students were capable of applying the characteristic equation and solve for the eigenvalue but, struggled in interpreting that the eigenvalue was not the frequency. When students were given the matrix equation in Question 2, their *interpretation* increased by 12 percent as it was more recognizable to decipher the elements in the matrix equation with the elements in the eigenequation, thus resulted a higher success rate in finding the frequencies.

The results for Question 1 part (i) and Question 3 show how students implement the eigenequation for different types of physical systems. We mentioned in the previous section that these two questions have similar tasks but one dealing with normal modes and the other with LC circuits. Despite the procedure being the same for constructing the matrix equation, there was a decrease in performance from Question 1 part (i) to Question 3. We believe that the decrease is due to changing the physical system to one where students were not familiar with in terms of practice. Students have practiced similar problems to Question 1 through homework assignments and in-class activities, but not for LC circuits.

The responses in Question 3 showed that the students struggled in recognizing that the coupled LC circuit is a normal mode system thus, having a similar procedure as the problem for two masses on springs. The students' performance for Question 3 decreased by 14 percent in comparison to Question 1 part (i). This could possibly tell us that students tend to think that with a different physical system, there involves a different method of solving. Figure 5 displays two student written responses for both questions. These two responses in a way, are similar in such that they are both missing important elements from the eigenequation, but differ by how they constructed their answers.
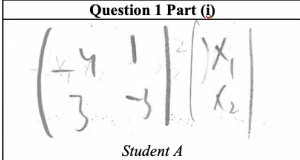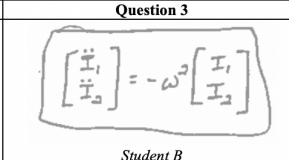

Question 1 Part (i)	Question 3
Student A	*Student B*

FIG. 5. A comparison of two student's written responses for Question 1 part (i) and Question 3, respectively.

For Question 1 part (ii), the success rate was high. More than half of the students were able to correctly interpret the physical meaning of the column vector. One student wrote, *"The motion of the blocks m_1 and m_2, $\begin{pmatrix} 1cm \\ -2cm \end{pmatrix}$ means that we moved block 1, 1 cm above its equilibrium and block 2, 2 cm below its equilibrium."* 79 percent of the students wrote similar interpretations.

Another issue we saw when examining the student responses was how students handled with the symbols in the problems. One student wrote for Question 2, " *Im not sure where to go from here. too many variables than im used to in an eigenvalue matrix problem."*, suggesting that the symbols were what hindered his performance in finding the eigenvalue. Figure 4 shows an example of a student misinterpreting the second order derivative as a variable rather than an operation.

FIG. 6. Student response for Question 3 (LC circuit). The student misinterpreted the second order derivative as a variable rather than an operation.

Problems that contain more symbols than numerical values in addition to asking questions that require mathematization

and interpretation, can potentially generate confusion for students on how to utilize and comprehend the information given to them, as seen in Questions 1 and 3. Students are fluent in technical operations, as we saw in Question 1 part (iii) and Question 2—where a majority of the students correctly used the characteristic equation to solve for the eigenvalue. However, it is important to note that the application of *technical operations* had been implanted in most of the students' education careers; using various mathematical operations to calculate values has consistently appear in a majority of math courses. This observation further validates the initial speculation presented. Our data show that the students in this study had a difficult time conceptually applying matrix algebra and eigentheory in physics.

IV. CONCLUSIONS

By utilizing the Physcial-Mathematical Model, we were able to identify and better correlate the student responses to their understanding of matrix algebra and eigentheory. Identi-

fying which skills students struggle applying the most helped us associate the challenges these concepts hold. The results insinuate that we need to look beyond the scope of student usage of these concepts and examine deeper into how students collect their understanding of these concepts. Their struggle in composing mathematical sense by constructing a matrix equation and interpreting what each symbol represents correlate to their understanding of matrix algebra and eigentheory in a physics context. Our study of student understanding of matrix algebra and eigentheory continues with more evaluations of student written responses and revising ungraded quizzes that gear towards mathematization and interpretation.

V. ACKNOWLEDGEMENTS

[1] C. Hillebrand-Viljoen and S. Wheaton, How students apply linear algebra to quantum mechanics, 2018 Physics Education Research Conference Proceedings 10.1119/perc.2018.pr.hillebrand-viljoen (2019).
[2] M. Wawro, K. Watson, and W. Christensen, Meta-representational competence with linear algebra in quantum mechanics, in *CERME 10* (Dublin, Ireland, 2017).
[3] M. S. Sabella and E. F. Redish, Student understanding of topics in linear algebra (2002).
[4] D. Carlson, C. R. Johnson, D. C. Lay, and A. D. Porter, The linear algebra curriculum study group recommendations for the first course in linear algebra, The College Mathematics Journal **24**, 41–46 (1993).
[5] C. LARSON and M. ZANDIEH, Three interpretations of the matrix equation ax=b, For the Learning of Mathematics **33**, 11

(2013).
[6] C. Andrews-Larson, Conceptualizing matrix multiplication: A framework for student thinking, an historical analysis, and a modeling perspective, ProQuest LLC (2010).
[7] F. Henderson, C. Rasmussen, M. Zandieh, M. Wawro, and G. Sweeney, Symbol sense in linear algebra: A start toward eigen theory (2010).
[8] M. Thomas and S. Stewart, Eigenvalues and eigenvectors: Embodied, symbolic and formal thinking, Mathematics Education Research Journal **23**, 275 (2011).
[9] O. Uhden, R. Karam, M. Pietrocola, and G. Pospiech, Modelling mathematical reasoning in physics education, Science Education **21** (2011).

Mass Distribution Comparison for GW190425

Author: Marc Daniell Penuliar
Advisor: Dr. Jocelyn Read
California State University, Fullerton
Department of Physics

On August 17, 2017 the Laser Interferometer Gravitational-Wave Observatory (LIGO), and its European counterpart Virgo, detected the emission of gravitational waves from two neutron stars as they were orbiting and spiraling into each other. Neutron stars (NS) are some of the most dense and compact objects ever discovered. Mass measurements of galactic neutron stars show a range of values between roughly 1 and 2.5 times the mass of the sun, and the population was fit to a double gaussian distribution in Alsing et al 2018. Most are at roughly 1.4 solar masses (Mø) LIGO's first detection of neutron stars also showed masses around 1.4 Mø. In 2019, LIGO discovered a second neutron star binary, GW190425 with a significantly larger total mass of about 3.4 Mø and mass of component neutron stars to be ranging from 1.12 Mø to 2.52 Mø. Here, we compare the range of masses compatible with this gravitational wave signal to a neutron star binary population generated following the galactic mass distribution of Alsing et al. This comparison suggests that, if the new binary comes from a population like we see in our galaxy, GW190425 had a larger mass ratio than suggested by gravitational wave data alone, with one component near the typical 1.4 Mø value and the other similar to the most massive neutron stars observed in our galaxy. Creating these joint plots opens the opportunity to compare our results of GW190425 quantitatively with the Alsing paper's probability density.

Optical scattering measurements for common off the shelf highly reflective 1064 nm mirrors

Authors: Daniel Martinez, Amy Gleckl, Gabriela Jaimes, Jazlyn Guerrero, Elenna Capote, Erick Engelby, and Joshua Smith

California State University, Fullerton
California Institute Of Technology
Massachusetts Institute Of Technology
Ligo Hanford Observatory
Ligo Livingston Observatory

1   Introduction

Highly reflective 1064nm mirrors are often used in experiments by the LIGO Scientific Collaboration and elsewhere. In Advanced LIGO, some optical setups such as the upper breadboards for the transmitted light monitors include such mirrors. While the transmission and reflectivity for these optics is readily available by vendor, scatter information is not. It is possible that there are wide variations in scatter between optics from different manufacturers. Identifying the best "off the shelf" optics (ie direct from manufacturer) for setups in which scattered light is critical could lead to improved optical performance. Here we look at the scatter versus angle of off the shelf 1064 nm laser mirrors from the vendors Newport, Thorlabs, and Edmunds Optics. We observed two things: i) the scatter from a majority of samples improved after cleaning with First Contact, indicating that they had arrived from the manufacturer with less than optimal cleaning and ii) after cleaning, the lowest scattering samples were from Newport, followed by Edmunds Optics.

2   Samples

Table 1 lists the mirrors measured in this study. There are six mirrors in total, two of the same type from Newport, two of the same type from Thorlabs, and two of differing types from Edmund Optics. All were purchased in June 2019. All of the optics have the same diameter and thickness. The information in this table and the subsections below was taken from the quotes and the company websites for each of these optics. Notably, scattered light specifications are not given. Because the Newport and Thorlabs samples are nominally identical, we will refer to their two samples as a and b.

Sample	Qty.	Coating and Substrate	Reflectivity 1064 nm	Surface Flatness	Surface Quality
Newport 10QM20HM.10	2	Dielectric coating on fused silica (front)	$R_s > 99.7\%$ $R_p > 99\%$	$\frac{\lambda}{10}$	10-5
Thorlabs NB1-K14	2	Dielectric coating on fused silica (front)	$R_{avg} > 99.8\%$ for 0° AOI	$\frac{\lambda}{10}$	10-5
Edmund 38905	1	Dielectric coating on fused silica (front)	$R_{abs} \leq 99.8\%$	$\frac{\lambda}{10}$	10-5
Edmund 89452	1	Dielectric coating on fused silica (front)	99.5% for S- and P-polarization	$\frac{\lambda}{8}$	20-10

Table 1: Specifications from manufacturer for each optic. Surface flatness is typically specified at 632.8 nm. Note that most of these optics have nearly the same optical specifications. However, optical scattering is not specified.
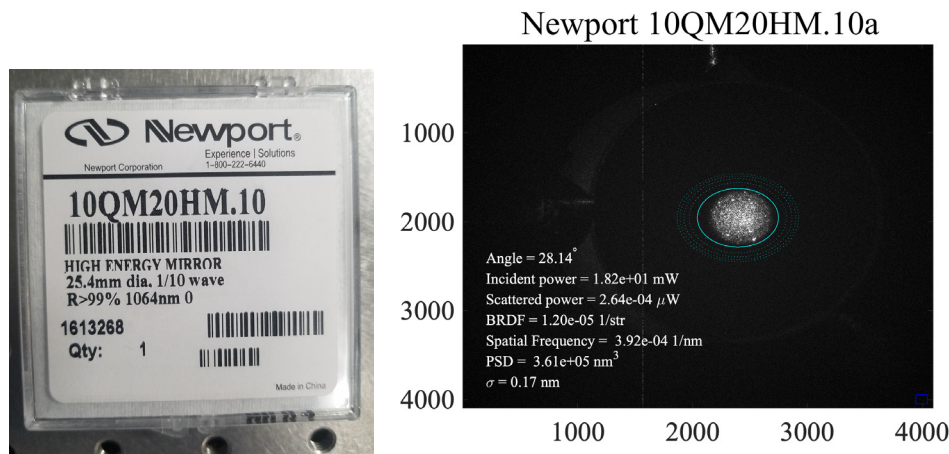
2.1 Newport 10QM20HM.10

Figure 1: Left: Sample Newport 10QM20HM.10.a Right: Scatterometer image at θ_s =28°.

The Newport mirrors have a $\geq 80\%$ clear aperture with a 1064 nm highly reflective coating Nd:YAG laser. It consists of a UV grade fused silica substrate with an ultra-hard dielectric coating. The AOI is 0° with a reflectivity of 99.7% for S-polarization and 99% for P-polarization. The surface flatness is $\frac{\lambda}{10}$, with a 10-5 scratch-dig. Two of the same sample were purchased to compared their similarities. Individually, they are sold for $258 on the Newport website, along with a more detailed list of specifications.
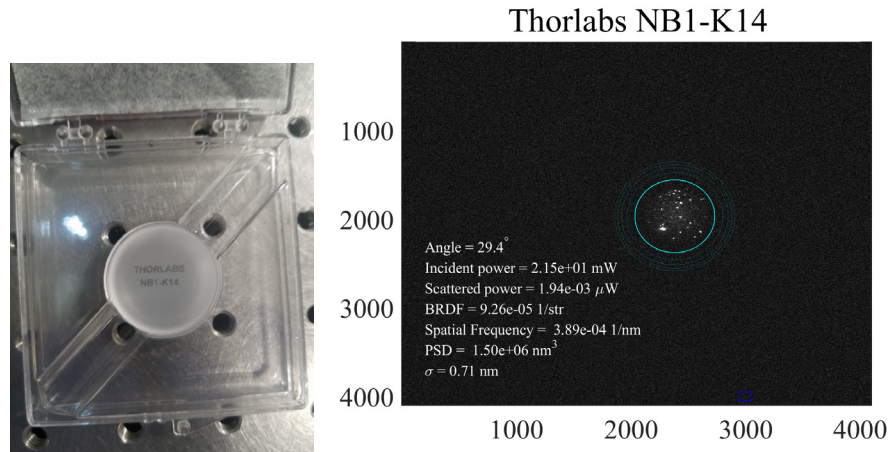
2.2 Thorlabs NB1-K14

Figure 2: Left: Sample Thorlabs NB1-K14. Right: Scatterometer image at $\theta_s = 29°$.

The Thorlabs mirrors have a clear aperture of $> 80\%$. The AOI is optimized from $0°$ to $45°$. The optic is coated with a -K14 dielectric coating to allow the reflectivity of 99.5% for S- and P-polarization. It has a surface roughness (RMS) of < 2 Å and a surface quality of 10-5 scratch dig. The backside and barrel of the optic are not polished to diffuse scatter from those surfaces. Two of the same sample were purchased as well, sold at \$142 on the Thorlabs website.
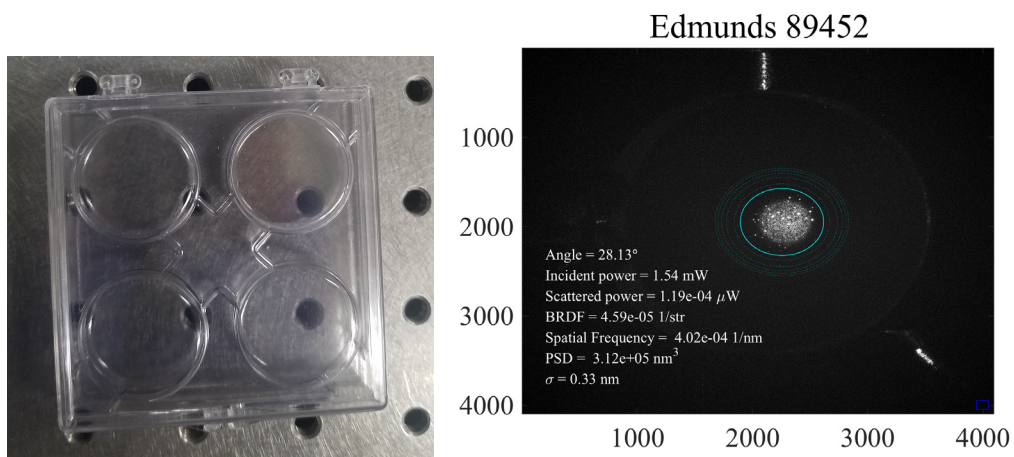
2.3 Edmund Optics 89452 and 38905

Figure 3: Left: Sample Edmund 89452. Right: Scatterometer image at $\theta_s = 28°0$.

Edmund 89452 has a clear aperture of 90%. It has a reflectivity of $> 99.95\%$ for both S- and P-polarization of the 1064 nm laser and an AOI of 0°. These mirrors use a fused silica substrate with a dielectric coating on the front surface only, with a clear back surface. The surface flatness is $\frac{\lambda}{8}$ with a surface quality of 20-10 scratch dig. It is sold for $159 on the Edmund Optics website.
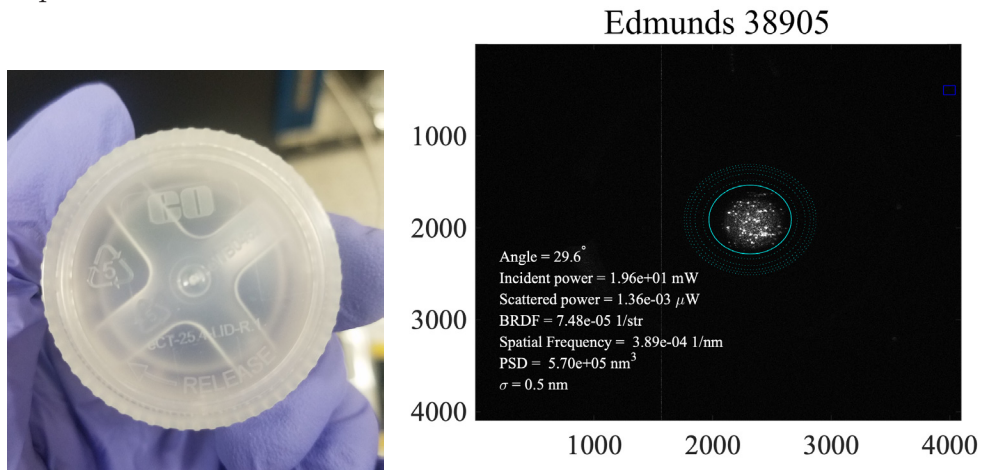

Figure 4: Left: Sample Edmund 38905. Right: Scatterometer image at $\theta_s = 29°$.

Edmund 38905 has a clear aperture of 90%. The optic again uses a fused silica substrate with a dielectric coating on the front surface and a commercial polish on the back surface. It has a 99.8% reflectivity at the 1064 nm wavelength and an AOI from 0° to 45°. The surface flatness is $\frac{\lambda}{10}$ with a surface quality of 10-5 scratch dig. It is sold for $130 on the Edmund Optics website.

3   Scatterometer Setup

To determine the angle-resolved scattering of these samples we conducted experiments with the Fullerton Imaging Scatterometer [1, 2, 3], with the setup shown in Fig 5.

A 1064 nm laser (Innolifht Mephisto S) is first channeled through a 90:10 fiber optic cable beam splitter. The 90% portion of the beam is sent through a reflective collimator, linear polarizer, and adjustable iris. This beam is then positioned at near normal incidents to the center of the optic sample. The other 10% is sent to an isolated power monitor which records power throughout the duration of the experiment. The 90% laser beam and optic are mounted on a rotation stage. Reflected and transmitted beams are attenuated by beam dumps made using black welders grade glass (also secured to the rotation stage).

Figure 5: The Fullerton Imaging Scatterometer diagram

The scattered light from the sample is imaged by a 4096x4096 CCD camera (Apogee Andor Alta F16) that utilizes a 100mm biconvex lens and an adjustable iris to focus the image at the CCD chip. The CCD camera images the scatter from the optic as the rotation stage (Thorlabs HS Nano Rotator)rotates the optic and laser in unison, from 0° to 80° in 1° increments. The entire experiment is automated by a Lab-VIEW Virtual Instrument. The the optic is first imaged with the laser on, this generates a set of "bright images". Next, the process is repeated, imaging from 0° to 80° in 1° increments but now with the laser off, generating a set of "dark images". These dark images are then subtracted out from the bright images to eliminate noise from the CCD chip and account for any possible ambient light in the room.

Incident power, reflected power, and transmitted power are all recorded at the beginning and end of the experiment to calculate an average incident power. Exposure times for images are selected such that the scatter does not saturate the CCD camera chip but is also strong enough to have a good signal-to-noise ratio for data analysis. The resulting data from the experiment is ran through a MATLAB script that calculates and plots the Bidirectional Reflectance Distribution Function (BRDF) and Total Integrated Scatter (TIS), as well as a movie of that sample as it rotates with the region of interest (ROI) shown.
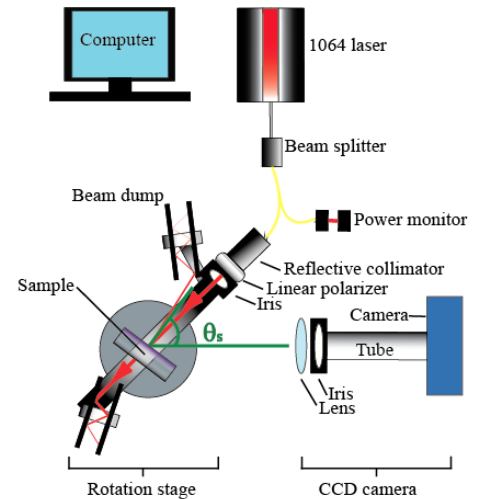
4   Cleaning Procedure

Mirrors were cleaned with Red First Contact from Photonic Cleaning Technologies following protocol from LIGO's cleaning procedure [4, 5, 6]. All cleaning and handleing of the optic takes place inside a softwall cleanroom. Since these are small optics (less than 2inch diameter), our lab utilized the pour and brush application protocol for First Contact alcohol-polymer solution [5]. In this protocol the optic is laid horizontally on an optical mount and red first contact is poured in the center of the optic, then gently brushed/pulled to the edges using a brush applicator and the surface tension of the polymer solution. After a thick layer of first contact has been poured the solution is left to dry for at least one hour. A piece of first contact removal floss is next placed on top of the first layer of first contact along the outer edge of the optic, then a second layer of first contact is poured to secure the floss between the two layers. The second layer is left to dry for a minimum of 1 hour, though out of precaution the second layer is usually left to dry overnight. After the two layers have been left to dry for a sufficiently long time, removal is done following LIGO procedure [4]. A Top Gun Ionizing Air Gun System is used during the removal process to account for excess charge left on the optic due to the ionizing nature of the polymer solution. As per LIGO's recommendation, the top gun is always operated using research grade nitrogen which was selected for its efficiency in cleaning as well as its rapid decay of built up static charge [6]. As soon as the first contact layer is pulled away from the optic, another labmate immediately beings spraying the optics surface to reduce the static charge on the optic before it can attract new dust and/or other particulates. For optics of this size the top gun is held approximately 3 inches away from the optics surface and sprayed sheer/parallel to the surface, at a pressure of 30psi [6].

5   Analysis

The Bidirectional Reflectance Distribution Function (BRDF) is used to quantify the reflected light from samples. BRDF is defined as the derivative of the scattered power, P_s, over the solid angle, Ω_s, divided by the incident power, P_i, and the cosine of incident angle, θ_s, which is defined as the angle from the normal of the camera to the normal of the optics surface [7].

$$BRDF = \frac{dP_s/d\Omega_s}{P_i \cos\theta_s} = \frac{P_s/\Omega_s}{P_i \cos\theta_s} \tag{1}$$

5 Analysis

The Bidirectional Reflectance Distribution Function (BRDF) is used to quantify the reflected light from samples. BRDF is defined as the derivative of the scattered power, P_s, over the solid angle, Ω_s, divided by the incident power, P_i, and the cosine of incident angle, θ_s, which is defined as the angle from the normal of the camera to the normal of the optics surface [7].

$$BRDF = \frac{dP_s/d\Omega_s}{P_i \cos \theta_s} = \frac{P_s/\Omega_s}{P_i \cos \theta_s} \tag{1}$$

A region of interest or ROI is used to define the region of the CCD image that encompasses significant scattered light and will be analyzed by the MatLab code. A CCD camera works by counting the number of photons that each well in the 4096x4096 CCD chip inside the camera housing. These counts encompassed by the ROI are summed together then normalized by exposure time and incident power for each image. This value is multiplied by a calibration factor determined at the time the system was set up. This value gives us a BRDF at the moment the image was taken for the scatter encompassed in that ROI. Using a set of 6 concentric ROI's with increasing diameter, the BRDF value for each ROI is calculated and then fit to a linear function. The y-intercept of this linear function gives us the most accurate value for the BRDF of the scattered light in our initial ROI.

Total scattered light from an optic can be estimated by integrating a measurement of BRDF multiplied by $\cos \theta_s$ over the entire range of observed angles. The scattering over the azimuthal angles is assumed to be constant, so the integral is only taken over the incident angle, θ_s. The resutling value is known as the total integrated scatter, or TIS, and is defined as shown below [7].

$$TIS = 2\pi(\cos \theta_1 - \cos \theta_2)BRDF(\theta_s) \cos \theta_s \tag{2}$$

For this system, TIS is estimated by integrating BRDF from 0° to 80°. TIS gives a good interpretation of the total scatter from a samples and is also often used to estimate rms surface roughness of a sample.

6   Results

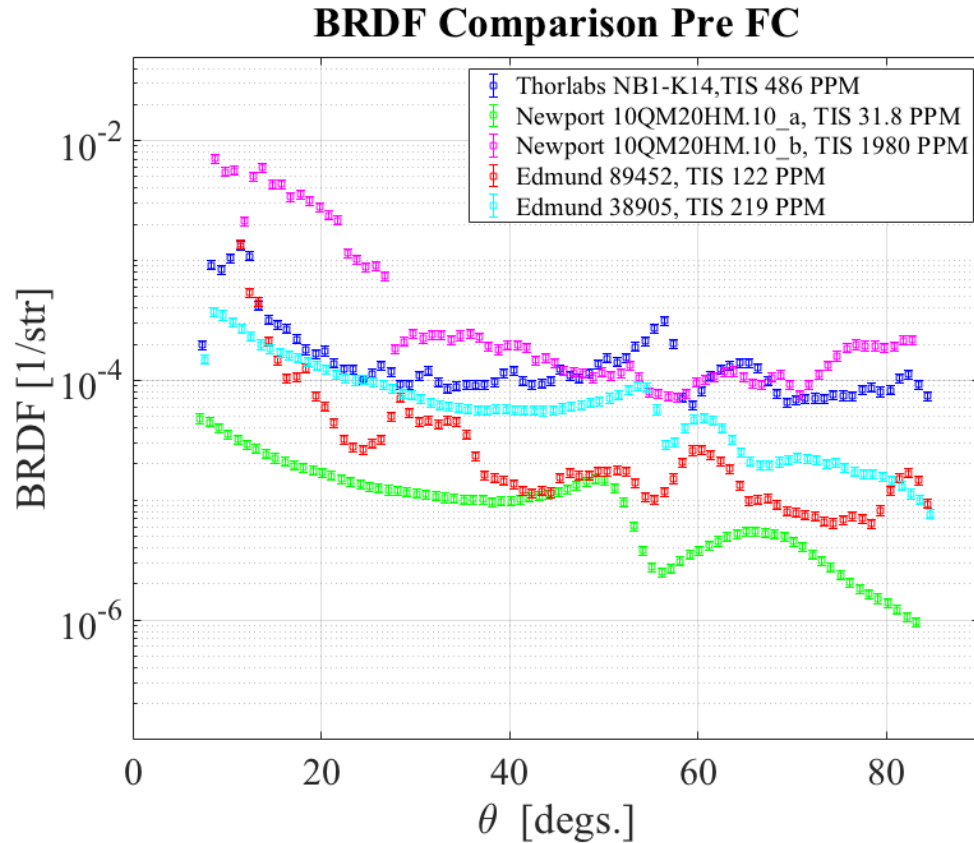## 6.1   Scatter Results Pre Cleaning with First Contact Solution

Figure 6: Measured normalized scatter (BRDF) versus scattering angle $\theta_s =$. These samples were measured directly from the manufacturer without cleaning.

Figure 6 shows the comparative plot for the normalized scatter (BRDF) versus scattering angle $\theta_s =$ of the samples before cleaning. Newport 10QM20HM.10_a had the lowest BRDF and TIS values directly from the manufacturer. However, Newport 10QM20HM.10_b which is meant to be near identical to Newport 10QM20HM.10_a, actually had the highest BRDF and TIS values at this stage. Assuming the samples have similar manufacturing, this indicates that there may be differing cleanliness between the two Newport samples. Edmunds 89452 had the second lowest BRDF and TIS values, followed by Edmunds 38905 as the third lowest. Thorlabs NB1-K14 had the second highest BRDF and TIS values, and as stated previously, Newport 10QM20HM.10_b had the highest overall BRDF and TIS values. While we did

acquire two Thorlabs samples, only one Thorlabs mirror was analyzed before cleaning due to a miscommunication in the lab.

Further data and movies for these samples is available at this link: `https://wiki.ligo.org/OPT/CSUFlog20190712`.

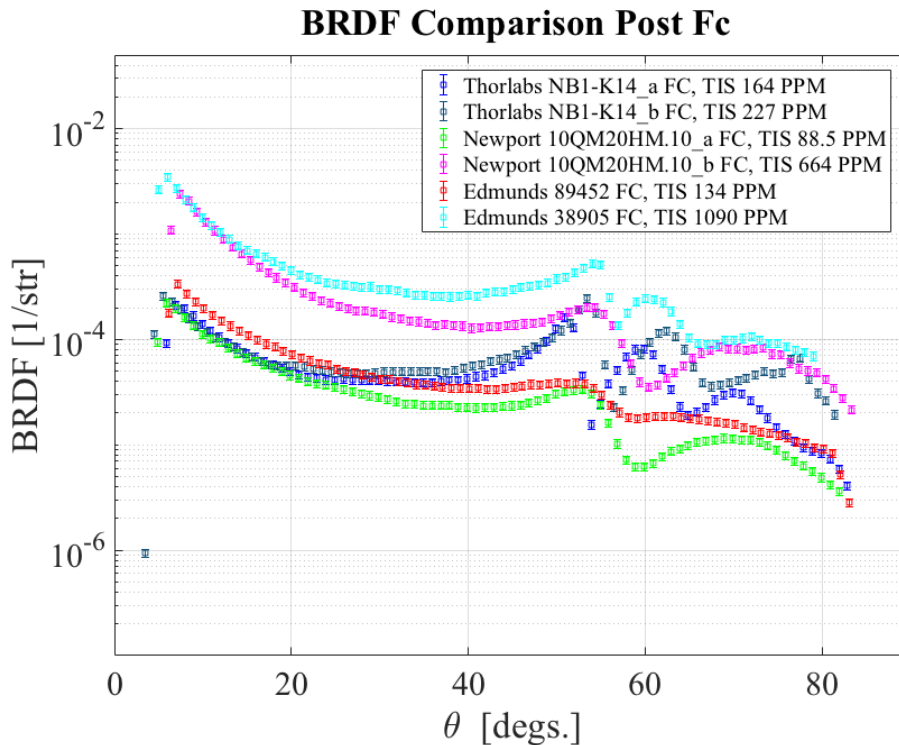### 6.2 Scatter Results Post Cleaning with First Contact Solution

Figure 7: Measured normalized scatter (BRDF) versus scattering angle $\theta_s =$. These measurements were taken after cleaning with first contact solution.

Figure 7 shows the comparative plot for the normalized scatter (BRDF) versus scattering angle θ_s of the samples after cleaning with first contact solution. Newport 10QM20HM.10_a exhibited a minor increase in TIS but overall was still found to have the lowest BRDF and TIS values. Edmunds 89452 also exhibited a minor increase in TIS but was again found to be the next lowest scattering sample. Thorlabs NB1-K14_a and Thorlabs NB1-K14_b were the third lowest, with NB1-K14_a being slightly lower than NB1-K14_b. Newport 10QM20HM.10_b, which was initially the highest scattering sample, did decrease in scatter slightly and was measured to be the second highest after cleaning. Edmunds 38905 increased in scatter after cleaning and was now found to be the highest scattering sample.

Sample	Pre TIS (ppm)	Post TIS (ppm)	Δ TIS
Newport 10QM20HM.10_a	31.8	88.5	+56.7
Newport 10QM20HM.10_b	1980	664	-1316
Thorlabs NB1-K14_a	486*	164	-322
Thorlabs NB1-K14_b	486*	227	-259
Edmund 38905	219	1090	+871
Edmund 89452	122	134	+12

Table 2: TIS values before and after cleaning for each sample as well as relative change. *Because Thorlabs NB1-K14 only had one sample analyzed before cleaning that value is being used as a pre TIS value for both a and b.

7   Conclusion

Overall Newport 10QM20HM.10 a was found to be the lowest scattering sample both be-fore and after cleaning. Contradictory, Newport 10QM20HM.10 b_ was found to be highest scattering sample before cleaning, and the second highest scattering sample after cleaning. Because of this we can say that the samples from Newport were found to have some of the largest variation amongst their samples. Thorlabs had the most consistent samples, having consistent decreases in TIS values and maintaining the same trend for BRDF both before and after cleaning. Edmunds 89452 had very little change in terms of TIS values, but did exhibit a noticeable change in the trend of the BRDF - smoothing out significantly after cleaning. Edmunds 38905 had a significant increase in TIS values but maintained the same BRDF trend before and after cleaning.

Directly from the manufacturer TIS values for these mirrors ranged between 30ppm and 2000ppm with no clear dependence on vendors. Excluding Edmunds 89452 which had a TIS increase of only 12ppm, first contact cleaning improved the TIS of all but two samples, which were initially the lowest and second lowest scattering samples. The lowest scattering sample, Newport 10QM20HM.10_a, had TIS of 32ppm before cleaning and TIS increased to 89ppm after cleaning. The second lowest scattering sample, Edmunds 89452, had a TIS of 122ppm before cleaning and TIS increased to 134ppm after cleaning. From this we can say that in samples with TIS below approximately 150ppm, cleaning did not significantly improve scatter and may have actually contributed to increased scatter. After cleaning, TIS values range of 100 to 1000ppm, though it should be noted that four of the six samples were near or below 200ppm. In conclusion we can say that a majority of the samples improved after cleaning with most of the samples having a final TIS near or below 200ppm.

Acknowledgements

References

[1] Fabian Magana-Sandoval, Rana X. Adhikari, Valera Frolov, Jan Harms, Jacqueline Lee, Shannon Sankar, Peter R. Saulson, and Joshua R. Smith. Large-angle scattered light measurements for quantum-noise filter cavity design studies. *J. Opt. Soc. Am. A*, 29(8):1722–1727, Aug 2012.

[2] Cinthia Padilla, Fabian Magaa-Sandoval, Erik Muniz, Joshua R. Smith, Peter Fritschel, and Liyuan Zhang. Low scatter and ultra-low reflectivity measured in a fused silica window. *Appl. Opt.*, 53:1315, 2014.

[3] Daniel Vander-Hyde, Claude Amra, Michel Lequime, Fabian Magana-Sandoval, Joshua R Smith, and Myriam Zerrad. Optical scatter of quantum noise filter cavity optics. *Classical and Quantum Gravity*, 32(13):135019, 2015.

[4] Kate Gushwa and Calum Torrie. First contact spray application removal procedure. Technical report, LIGO, 2018.

[5] Margot Phelps, GariLynn Billingsley, Liyuan Zhang, Calum Torrie, and Charlie Taylor. First contact pour and brush application. Technical report, LIGO, 2018.

[6] Calum Torrie, GariLynn Billingsley, Eddie Sanchez, Bob Taylor, and Margot Phelps. Guidance on top gun ionizing air gun system - from gas to gun. Technical report, LIGO, 2015.

[7] J. C. Stover. *Optical Scattering*. SPIE Press, 3rd edition, 2012.

Calculating apparent-horizon quantities with SpECTRE, a next-generation numerical relativity code

Authors: Marlo Morales
Advisor: Dr. Geoffrey Lovelace
California State University, Fullerton
Department of Physics

Abstract

SpECTRE is a next-generation numerical-relativity code (currently under development) that will calculate the gravitational waves emitted by colliding black holes with unprecedented accuracy, by using novel techniques that enable it to scale to hundreds of thousands of compute cores. These high-accuracy calculations will help scientists interpret observations from next-generation gravitational-wave detectors. My research involves the completion of SpECTRE's computational infrastructure to measure the properties of black hole's horizons, such as the black hole's masses and spin angular momenta, which are especially important for connecting the calculations to observations. Then, I assess the accuracy that SpECTRE is able to achieve and compare the results to those from the Spectral Einstein Code (SpEC), a current-generation code.

Authors & Editors

Dimensions Editors

Sherif Ibraheem (Layout Editor)

Sherif Ibraheem is a first-year undergraduate student at California State University, Fullerton. He is majoring in Biology and working towards his Bachelor's Degree. One of his favorite hobbies is graphic designing. He is experienced working with professional designing softwares like Adobe Programs. His skills got him hired for the layout editor position for Dimensions 2021. After graduation, he wants to further his education and get into medical school. In the future, he plans to pursue a career in the medical field.

Sonali Vyas (Editor-In-Chief)

Sonali Vyas is a sixth year Applied Mathematics and English Major. She is wrapping up a new research project in which she applies machine learning to predict murder mystery novels. Sonali is in the process of applying to graduate school and hopes to earn a PhD in Applied Mathematics.

Krisha Marie Uy (Section Editor)

Krisha Marie Uy is a first-year biology major at California State University, Fullerton (CSUF) and plans to pursue a concentration in Molecular Biology and Biotechnology. She is a part of the President's Scholars Program and University Honors Program on campus. Upon graduating, Krisha wants to further her education by attending medical school as she aspires to become a pediatrician or pediatric surgeon in the future.
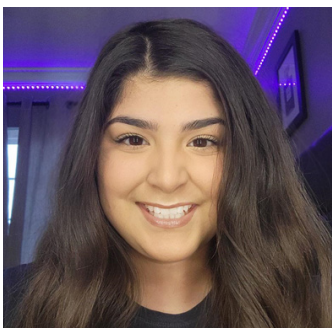
Melissa Fernandez (Section Editor)

Melissa Fernandez is currently a Junior at Cal State Fullerton and is majoring in Biological science, with a concentration in Cellular and Developmental Biology. She is a dedicated member of Dr. Shahrestani's Evolutionary and Genomics Research (EAGR) Lab, where she works with the rest of her team to establish Drosophila melanogaster as a model organism for the antibiotic resistance pathogen, Acinetobacter baumannii. She is a board member for AMSA and the editor for the Biology section of Dimensions. In her free time, she advocates for mental health awareness as a Robbie's Hope Organization ambassador. After graduation, she plans on earning her master's degree before pursuing a lifelong career as a doctor in the medical field.

Madeline Ceccia (Section Editor)

Madeline Ceccia is a senior majoring in applied mathematics with a concentration in modeling and computation. Under the mentorship of Dr. Martin Bonsangue, Madeline investigates the numerical negative space within the distinguished Fibonacci Sequence. She is passionately curious and inspired by the vast mysteries of the universe. Madeline was titled a NASA Community College Aerospace Scholar in 2018, and aspires to work for a space agency once she graduates in 2021.
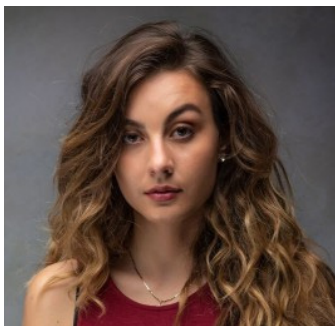
Anny Antunovich (Section Editor)

Anny Antunovich is a senior-level undergraduate student at Cal State Fullerton studying Physics. With the mentorship of Dr. Jocelyn Read in the Physics Department, Anny is working on astrophysical research with the Gravitational Waves Physics and Astronomy Center on campus. In the fall, she plans to further her education by pursuing a Master's degree in Astrophysics and Cosmology abroad.

Jessica Sherman (Section Editor)

Jessica Sherman graduated from California State University, Fullerton with a bachelor's degree in Biological Sciences with a concentration in Molecular and Biotechnology and a minor in Psychology. She is currently working towards her certification as a certified nurse assistant and set to leave on a medical mission in Kenya. She hopes to gain a different perspective of medicine in another country and help better prepare her for her future career as a travel Physician Assistant.

Mykayla Miller (Section Editor)

Mykayla Miller is an honors student majoring in Biochemistry. She is a McNair scholar at California State University, Fullerton. She has recently began working under Dr. Marcos Ortega in studying how to inhibit androgen receptors and lysine demethylase I using a molecule called sangaurine.
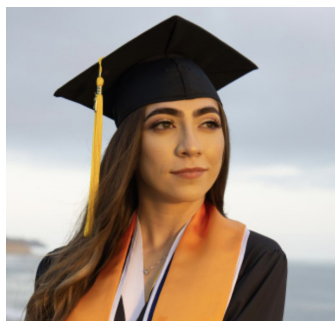
Mayra Silva (Section Editor)

Mayra Silva is a graduating Senior Majoring in biology and concentrating in ecology and evolutionary biology. She is a Scholar apart of the Southern California Ecosystems Research Program and conducts oyster conservation research with the Zacherl Lab. "This is my second year editing for Dimensions and it has been a blast! Thank you so much for the opportunity and the experience!"

Authors

Akçiz, Sinan

Dr. Sinan Akçiz is an Assistant Professor of Geology at California State University, Fullerton. Sinan's research areas include paleoseismology, geomorphology, Quaternary geochronology, structural geology, and geoscience education.

Albakri, Salma

Salma Albakri is a third-year undergraduate majoring in Mathematics with a Teaching Concentration at California State University, Fullerton. She plans to graduate in May 2022 and enroll in a credential program for the following fall semester. Her future career goals include becoming a high school mathematics teacher teaching subjects such as AP Statistics.

Alonso, Gustavo Sopena

Gustavo Sopena Alonso is a first year graduate student in the applied mathematics program at California State University, Fullerton (CSUF). He received his bachelor's degree in applied mathematics from CSUF in May 2020. His current research involves exploring Peg Solitaire's various game boards and generalizing the game to consider the use of colors. He is interested in modeling the disease dynamics among bird and human populations as well as employing a game-theoretic approach to studying the disease life cycle.

Anderson , Holly

Holly Anderson is an undergraduate junior at Cal State Fullerton. She is currently pursuing a bachelor's degree in mathematics with a concentration in teaching. After she graduates, she hopes to either obtain her teaching credential and become a high school mathematics teacher or pursue higher education and become a college professor.

Andrews, Summer

Summer Andrews is a graduate student majoring in Mathematics with a Teaching concentration and will earn her Master's degree from Cal State Fullerton in May 2021. She is also a Teacher's Associate teaching undergraduate mathematics courses at Cal State Fullerton. Upon graduation, she hopes to continue on with her passion of teaching mathematics and pursue a faculty position at a college or university.

Arreola, Seth

Seth Arreola recently received his Bachelor's degree in Mathematics, with a concentration in probability and statistics, in January 2021 from California State University, Fullerton. Seth worked with the modeling and analysis of data, including COVID-19 infection and mobility data under the guidance of Dr. Behseta. He plans on continuing his education with the pursuit of a Master's degree in Statistics.


Bursztyn, Natalie

Dr. Natalie Bursztyn is a Physical Sciences Tutor of Geology at Quest University Canada, as well as a former Assistant Professor of Geology at California State University, Fullerton. Natalie is primarily interested in college-level geoscience education research, including the integration of new technologies in the classroom to provide interactive learning experiences.

Cabrera, Juan

Juan Cabrera majored in mathematics with a focus on probability and statistics. His research interest includes data mining, data visualization, machine learning, neural networks, and deep learning. He will begin a master's degree at CSUF in the fall of 2020. Juan is a first-generation college student. His future plans are to finish his graduate studies and begin a career in data science and teaching.
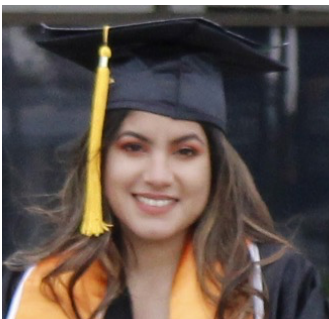
Carrilo, Heather

Heather Carrilo is a Biochemistry major and plans to work in the laboratory industry. Her career goal is to work in a crime lab for the county.

Ceccia, Madeline

Madeline Ceccia is a senior majoring in applied mathematics with a concentrationin modeling and computation. Under the mentorship of Dr. Martin Bonsangue, Madeline investigates the numerical negative space within the distinguished Fibonacci Sequence. She is passionately curious and inspired by the vast mysteries of the universe. Madeline was titled a NASA Community College Aerospace Scholar in 2018, and aspires to work for a space agency once she graduates in 2021.

Cugini, Brandon

Brandon Cugini is a graduating senior for the Summer 2021 semester with a major focus in Geology, and a minor in Geography. In addition, he is a member of the University Honors Program where he did his undergraduate thesis research in conjunction with his senior honors project. The thesis focused on using high resolution LiDAR images to take measurements of horizontal slippage of a segment of the Lavic Lake Fault, which was responsible for the Hector Mine Earthquake of 1999.

Culleton, Andrew

Andrew Culleton is a senior at California State University Fullerton, majoring in geological sciences. He is doing research with Dr. Valbone Memeti on the petrologic evolution of magma systems. After graduating, Andrew plans to start a career in the environmental consulting industry and work towards obtaining his Professional Geologist license.

Fung, Andy

Anderson Fung is an undergraduate physics major working in the physics education research group. His current project involves the improvement of ordinary differential equations (O.D.E.s) instruction in the department's math methods course (PHYS 300). He is in the final year of his undergraduate degree and intends to continue his involvement with the physics department as a graduate student.

Gamez, Christine

Christine Gamez is a senior pursuing a Bachelor of Arts degree in Mathematics with a concentration in teaching. She is currently doing research with Dr. Alison Marzocchi on gender representation in mathematics textbooks and is also on the executive council of her panhellenic sorority, Zeta Tau Alpha. After graduation, she will pursue her teaching credential and hopes to teach high school mathematics to at-risk youth.

Gutierrez, Joseph

Joseph Gutierrez is graduate student and teaching associate in the Department of Geological Sciences. His research focuses on educational materials assessment, with an emphasis on spatial visualization training in the geosciences. Having recently defended his thesis, Joseph will soon graduate with his M.S. in Geology, with plans to continue a career in geoscience education.

Her, Pachi

Pachi Her is the author of "Examining Student Understanding of Matrix Algebra and Eigentheory."

Huang, Solomon

Solomon Huang is a mathematics student at California State University, Fullerton. He is currently working on two research projects with Drs. Casper and Murphy. After graduation, Solomon plans to pursue a graduate degree to achieve his goal of teaching mathematics at the college/university level.

Jimenez, Jose Luis Gonzalez

Jose Luis Gonzalez Jimenez is an undergraduate majoring in chemistry at California State University, Fullerton. He is currently a junior working toward his bachelor degree. He is in his third semester of being a part of Dr. Fry-Petit research lab.

Kumar, Sushanth Sathish

Sushanth Sathish Kumar is a high school student at Portola High school who is very passionate about math. He enjoys learning about new math and solving problems in math journals such as the American Math Monthly, and the College Math Journal. He also enjoys math classes and loves to look beyond the traditional math taught in schools. In his free time, he watches anime and youtubers such as Vsauce.

Lind, Gwendolyn

Gwendolyn Lind is graduating in Mathematics with a concentration in Probability and Statistics in Spring 2021 and plans to continue at Cal State Fullerton to obtain a Master's of Science Degree in Statistics. As a BD3-REAP scholar, Gwen has participated in numerous public health research projects and presented her findings at the University of Southern California (Neuroimaging and Informatics) and the SACNAS Conference in Honolulu, Hawaii. Currently, she is managing a new team of scholars to continue last semesters CSUF COVID-19 research project by updating the existing statistical models and adding new data associated with the number of vaccines.

Lopez, Saul

Saul Lopez graduated from CSU Fullerton with BA in applied mathematics and graduated from the mathematics master's program with emphasis in teaching this past January. Saul was an Instructional Assistance for four years in the Math Study Hall at Santiago Canyon College. During his graduate studies Saul was a Teaching Assistant, TA, where he taught College Algebra and Pre-Calculus. Currently he is teaching mathematics for the Continuing Education program at Santiago Canyon College and Santiago Canyon College.

Martinez, Daniel

Daniel Martinez is an undergraduate studying physics. He transferred from Citrus College in 2020 after participating in a summer research program with GWPAC. After completing the summer project he wanted to return to continue working on research and take the courses CSUF provided. He has been a research assistant in Dr. Smith's GWPAC lab for the entirety of the 2020-2021 school year. After graduation he plans to attend graduate school to further his physics education.

Morales, Marlo

Marlo Morales is a junior physics student at California State University, Fullerton. Marlo is a currently working at CSUF's own Nicholas and Lee Begovich Center for Gravitational-Wave Physics and Astronomy, also known as GWPAC in Dr. Geoffrey Lovelace's numerical relativity group. His research involves contributing code that will be used to study the properties of extreme spacetime through numerical relativity research. His work will add new capabilities necessary to measure the properties of the black holes' horizon such as their surface areas and the measures of mass and spin angular momenta. His career goal is to pursue a Ph.D. in gravitational waves physics to contribute to the understanding of general relativity and the properties of Einstein's equation.

Nguyen, Duy

Duy Nguyen graduated with a Master's degree in Applied Mathematics at California State University of Fullerton. His goals are becoming a Modelling and Simulation engineer and a Math professor at a community college. Currently, he is a teaching intern at Goldenwest College, and plans to apply for a part-time teaching job after his internship.

Nichols, Kevin

Dr. Kevin Nichols is an Associate Professor of Mathematics at California State University, Fullerton. Kevin's research interests are spatial / temporal statistics and their applications to earthquake and wildfire data.

Nonora, Jesi

Jesi Nonora is an undergraduate research student pursuing a B.S. in Biological Science with a concentration in ecology and evolutionary biology. She is working under Dr. Paul Stapp to better understand coyote diet in rural areas of California and the comparability of different diet analysis techniques. After graduation, she plans to attend graduate school to continue her education and research on California wildlife.

Peña, Caleb

Caleb Peña graduated from Cal State Fullerton in 2020 with a bachelor's degree in Mathematics: Probability and Statistics. His interests include causal inference, data visualization, and sports analytics. He recently concluded an internship with the Seattle Kraken building models in anticipation of the upcoming draft. This fall, he plans to begin pursuing a Masters of Statistical Practice at Carnegie Mellon University.

Penuliar, Marc Daniell

Marc Penuliar is working on his B.S. in Physics and Minor in Mathematics at California State University, Fullerton (CSUF). His research work with advisor Dr. Jocelyn Read is carried out within the scientific collaboration of the Laser Interferometer Gravitational-Wave Observatory (LIGO), as a member of CSUF's Nick and Lee Begovich Center for Gravitational Wave Physics and Astronomy. With plans to graduate in the year 2022 and enter graduate school with physics or astronomy with the goal of a career in either astronomy or in industry.

Pilker, Matthew

Matthew Pilker is graduating *magna cum laude* with his Bachelor of Science in Geology. He completed his bachelor's thesis/honors project with Dr. Jeffrey R. Knott, worked as a Research/Field Assistant and Supplemental Instruction Leader for the Department of Geological Sciences, and served on Interclub Council for the College of Natural Sciences and Mathematics. Upon graduating, he will pursue his Master of Science from CSUF while continuing to work in environmental consulting.

Rhee, Chloe

Chloe Rhee is a third-year student at Cal State Fullerton, majoring in Mathematics with a concentration in Teaching. Chloe holds an executive position with SMART GIRLS. After graduation, she plans on getting her Masters and teaching credential. She plans on pursuing a career as a high school Math teacher.

Roque, Theodore

Theodore Roque was born and raised in Los Angeles county. He obtained his undergraduate degree in electrical engineering in 2013 and worked in the telecommunications industry for a few years before deciding to pursue a career in medicine. He completed the pre-health post-baccalaureate program at CSUF in 2019 and is working to enter medical school within the next few years.

Salvador, Stephanie Mariel

Stephanie is in her 4th year majoring in Biochemistry and a part of the University's Honors Program. While being a full time student, she also works part time as a secretary at her local church and as a volunteer for PIH Health Hospital in the Pathology Lab and Hospice Center. Since her freshmen year at CSUF, she always had a strong interest working in a lab whether it is in the medical field or the pharmaceutical industry. On her free time, she enjoys eating out with family and friends, walking her dog, shopping, hiking, playing with her ukulele, watching Forensic Files or any comedy sketch, working out, and yoga. Stephanie's research under the guidance of Dr. Pecic has deepen her knowledge of biochemistry as well as grow her skill set for the laboratory field and graduate school.

Slaven, Elijah

Elijah Slaven graduated from California State University, Fullerton as a biology major with a minor in chemistry. Working in Dr. Abraham's lab, he helped demonstrate the importance of environmental research and how we can impact and improve our communities. Shortly, he will be attending nursing school, where he plans to use these experiences to help improve his community with the aid of fellow healthcare providers.

Solomatov, Aleksei

Aleksei Solomatov is an undergraduate majoring in biochemistry at CSUF. He is a full-time student on campus but at the same time working towards getting his pharmacy technician license to begin his career in the pharmacy field. His goal is to work as a clinical pharmacist in order to help patients and to work directly with other medical professionals. Aleksei joined Dr. Pecic's research lab to gain a further understanding of medicinal chemistry and drug design.

Spalding, Leah

Leah Spalding is pursuing a Bachelor's of Science degree in Ecology and Evolutionary Biology at Cal State Fullerton. As an undergraduate researcher in Dr. Parvin Shahrestani's lab, she studies the gut microbiome within experimentally-evolved populations of the common fruit fly, Drosophila melanogaster. Following graduation in Fall 2020, she plans to pursue a graduate degree in ecology and evolutionary biology.

Tomich, Brandon

Brandon Tomich is a soon-to-be graduate with a Computer Science major at California State University, Fullerton. His interests include data science and machine learning and he is looking for a career in areas such as image processing and audio analysis. In terms of the future, Brandon is looking to further specialize his skills by attending graduate school.

Valenzuela, Alexander

Alexander Valenzuela is a senior at California State University, Fullerton expecting to graduate in the summer of 2021 with a Bachelor of Science Degree in Geological Sciences. He is currently working with Dr. Vali Memeti to research the concentric pattern of the Box Springs pluton in Riverside, California. Alex looks forward to graduating so he can join the workforce and pursue a career in geology.

Weaver, Tuesday

Tuesday Danielle is a Plant Science major with a minor in Anthropology, who completed her research with Dr. Der lab. Her main interests are in medicinal ethnobotany and horticulture. Tuesday spends her time reading historical non-fiction, brewing kombucha, and tending to her plant props. She plans to attend grad school soon, and is working towards becoming a horticulturalist at a commercial greenhouse.