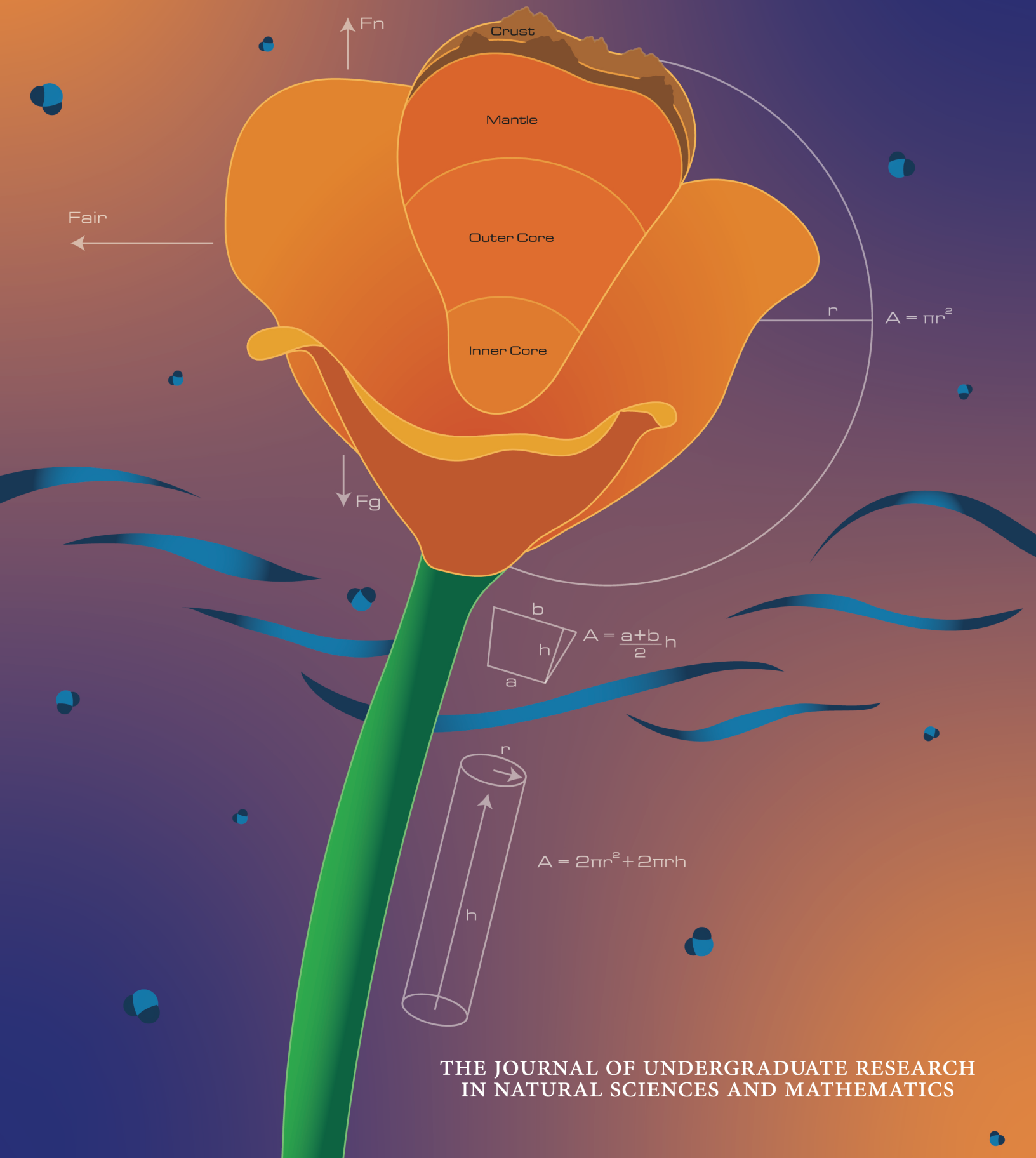


(ur) DIMENSIONS

VOLUME 24 | SPRING 2023



THE JOURNAL OF UNDERGRADUATE RESEARCH
IN NATURAL SCIENCES AND MATHEMATICS

(ur) DIMENSIONS

THE JOURNAL OF UNDERGRADUATE RESEARCH IN NATURAL SCIENCES AND MATHEMATICS

VOLUME 24 | SPRING 2023

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

DIMENSIONS: The Journal of Undergraduate Research in Natural Sciences and Mathematics is an official publication of California State University, Fullerton. DIMENSIONS is published annually by CSUF, 800 N. State College Blvd., Fullerton, CA 92834. Copyright ©2023 CSUF. Except as otherwise provided, DIMENSIONS grants permission for material in this publication to be copied for use by non-profit educational institutions for scholarly or instructional purposes only, provided that 1) copies are distributed at or below cost, 2) the author and DIMENSIONS are identified, and 3) proper notice of copyright appears on each copy. If the author retains the copyright, permission to copy must be obtained directly from the author.

ABOUT THE COVER

We would like to give a special thanks to Fernando Del Rosario and Group 4 of his graphic design class. The cover for Volume 24 was designed by Stephany Urzua Mendez. The artist's statement is as follows: "When designing the cover for this year's Dimensions Journal, I was inspired and knew I wanted to include every part of the department. Whether it was through text, equations, or visual illustrations, including every department within the College of Natural Sciences & Mathematics was an essential idea for my cover. The flower in the cover is a California Poppy, inspired by all the Poppies on campus, and is placed to visually represent the Biology department and all the natural processes that occur in plants. One of the petals of the poppy, was altered to represent and visually show the layers of the Earth as to represent the Geology department. When the flower is simplified to generic shapes, finding the area of those shapes is easier, representing the Mathematics department. To represent the Chemistry department I created various water molecules of different sizes throughout the cover to allude to the idea of it raining. The idea behind representing the Physics Department was to indicated the different forces acting on the flower. As for the background, I was inspired by the colors that appear in thermal readings."

DIMENSIONS Editorial Staff and NSM Leadership

EDITOR-IN-CHIEF

Jeremy Hansen – Assistant Director for Academic Support and Engagement

EDITORS

Ashleigh Quiroz – Geological Sciences

Christian Giron - Mathematics

Kristen Abe – Biological Science and Chemistry

Alyssa Ng – Biological Science

Marlene Lopez – Biological Science

GRAPHIC DESIGN

Stephany Urzua Mendez – Dimensions 2022-2023 Cover Art

COLLEGE OF NATURAL SCIENCES & MATHEMATICS

Dr. Marie Johnson - Dean

Dr. Sean Walker - Associate Dean

Dr. Merri Lynn Casem - Chair, Department of Biological Science

Dr. Nicholas Salzameda – 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

Table of Contents

7 - Biological Science

8 - Functional Assessment of an ExoS/ChvI Transcriptional Target Gene in *Sinorhizobium meliloti* During Free-Living Growth

- Author: Kristen Abe
- Advisor: Dr. Esther Chen

24 - Effects of grazing-related vegetation changes and moonlight on foraging behavior of shortgrass-prairie rodents

- Author: Marlene Lopez
- Advisor: Dr. Paul Stapp

42 - The Purification of the Mn Oxidizing Protein, MopA-hp, found in *Erythrobacter* sp. SD-21

- Author: Alyssa Ng
- Advisor: Dr. Hope A. Johnson

59 - Mathematics

60 - Combinatorial Identities Derived from the Monoid of Partial Order-preserving Injections

- Author: Tyler Kim
- Advisor: Dr. Scott Annin

79 - Efficiency of Horner's Method for Binomial Expansions

- Author: Cadence Pinkerton
- Advisor: Dr. Adam Glesser

88 - A Mathematical Investigation Of The Frontogenesis Function

- Author: Yulia Potyrina
- Advisor: Dr. Christopher Lyons

Biological Science

Functional Assessment of an ExoS/ChvI Transcriptional Target Gene in *Sinorhizobium meliloti* During Free-Living Growth

Kristen Abe

Advisor: Dr. Esther Chen

Abstract

Biological nitrogen fixation by microbes offers a sustainable alternative to synthetic nitrogen fertilizers that contribute to environmental pollution. The ExoS/ChvI two-component signaling pathway is critical for symbiotic nitrogen fixation by *Sinorhizobium meliloti*, a microbe that performs nitrogen fixation in an endosymbiotic relationship with legume plants. This pathway also regulates free-living bacterial phenotypes, including exopolysaccharide production, cell envelope integrity, and biofilm formation. This study aimed to determine the functions of small RNA (sRNA) genes that are transcriptionally regulated by ExoS/ChvI in *S. meliloti*. Strains overexpressing the sRNA gene *SmelC023* were successfully created, and preliminary results showed no differences in bacterial growth and free-living phenotypes compared to *S. meliloti* strains without *SmelC023* overexpressed. Overall, these results provide an initial contribution towards understanding the mechanisms of symbiotic nitrogen fixation and shed light on the potential role of ExoS/ChvI-regulated sRNA genes in *S. meliloti*. This work is supported by the Maximizing Access to Research Careers grant to CSUF from the National Institutes of Health [T34GM008612-26] and the National Institute of General Medical Sciences of the National Institutes of Health Award Number SC3GM144065 to E.J.C.

Introduction

Nitrogen is a crucial element in nucleic acids and amino acids and is therefore vital for the survival of all living organisms. Though nitrogen is one of the most abundant elements in the Earth's atmosphere, it is found in the chemically inert form of dinitrogen, which is not accessible for use by organisms until it is reduced to ammonia. This reduction relies on the cleavage of the molecule's triple bond during a process known as nitrogen fixation (Socolow, 1999). Biological nitrogen fixation is one of the main sources of fixed nitrogen on Earth, but it can only be performed by microbes called diazotrophs (Raymond, Siefert, Staples, & Blankenship, 2004). Rhizobia, such as the gram-negative, α -proteobacterium *Sinorhizobium meliloti*, are a type of diazotroph that can perform nitrogen fixation in an endosymbiosis with legume plants of the *Medicago* genus.

In nitrogen-limited conditions, the free-living bacteria evoke the formation of a nodule in the root of the legume and infect the plant through an infection thread (Gibson, et al., 2008; Long, 2016). The formation of this infection thread begins through hydrolysis of the plant cell wall and is dependent on the production of bacterial exopolysaccharides, such as succinoglycan or EPS-I (Cheng & Walker, 1998). Once they have successfully infected host cells, the free-living bacteria become surrounded by a plant-derived membrane in an environment known as the symbiosome where they undergo terminal differentiation into nitrogen-fixing bacteroids (Gibson et al., 2008; Long, 2016).

In *S. meliloti*, the ExoS/ChvI two-component signaling pathway transcriptionally regulates genes required for the bacteria to establish their symbiosis with the legume host. Two-

component signaling pathways are present in almost all bacteria and enable bacteria to sense and respond to environmental signals for survival and adaptation (Laub, 2011). The two conserved components in these pathways are the histidine protein kinase and the response regulator protein, which work together in signal transduction. The *exoS* and *chvI* genes are adjacent to one another in the *S. meliloti* chromosome and encode the two main components of the signaling pathway. The cytoplasmic membrane protein ExoS is the histidine kinase component with a periplasmic sensing domain and autophosphorylates in response to an unknown signal. ExoS is negatively regulated by the periplasmic inhibitor protein ExoR, which is believed to act by binding the sensing domain in ExoS, preventing its activity in the pathway (Chen et al., 2008). If ExoS is autophosphorylated, the phosphate group is transferred to the response regulator ChvI (Cheng & Walker, 1998). Once phosphorylated, ChvI activates transcription of *exo* genes, encoding enzymes for the synthesis of exopolysaccharides required for successful bacterial infection of the plant host (Belanger et al., 2009; Cheng & Walker, 1998). Notably, the ExoS/ChvI pathway regulates the synthesis of succinoglycan, which plays a vital role in the Rhizobia-legume symbiosis.

Bacterial small RNAs (sRNAs) are believed to play a role in microbe-host interactions, such as during the differentiation of free-living bacteria into their bacteroid form during symbiosis initiation. sRNAs are short, non-protein encoding RNAs that are found in abundance in all prokaryotes and regulate gene expression by base-pairing to a complementary mRNA target sequence (Storz et al. 2011). Genes directly regulated by the ExoS/ChvI pathway have their upstream region bound by ChvI, which acts as a transcriptional regulator. To date, numerous ExoS/ChvI direct target genes have been identified in *S. meliloti* through transcriptional profiling of strains with increased and reduced *chvI* activity to identify potential

target genes based on the reciprocal changes in expression in the *chvI* mutants compared to those in the wild-type strain (Chen et al., 2009). In a later study, more direct target genes of ChvI were identified by chromatin immunoprecipitation, followed by microarray analysis (ChIP-chip) and qRT-PCR; specifically, 489 ChvI-bound DNA regions were identified (Ratib et al., 2018). The sRNA gene *SmelC023* was found to be a potential candidate for direct regulation by the ExoS/ChvI pathway from previous RNA sequencing and ChIP-chip studies (del Val et al., 2007; Ratib et al., 2018; Schluter et al., 2010).

For this study, overexpression strains were constructed to observe any altered phenotypic effects. Plasmids with the sRNA gene *SmelC023* overexpressed were introduced into *S. meliloti* wild-type or *chvI* partial loss of function (LOF) strains. To observe growth phenotypes of free-living *S. meliloti*, these overexpression strains were grown on various media to test nutrient utilization, cell envelope integrity, survival in acidic conditions, and succinoglycan production.

Methods

Reagents and Equipment

All oligonucleotide primers used in this study were manufactured by Integrated DNA Technologies (IDT; Coralville, IA) and were diluted to 20 μ M for use (Table 1). Techne Progene (Techne; Staffordshire, UK) and Bio-Rad MJ Mini (Bio-Rad; Hercules, CA) thermocyclers were used for all PCR amplifications. Amplification of *SmelC023*, and its upstream and downstream regions, was carried out with Phusion polymerase (Thermo Fisher, Waltham, MA). All other PCR reactions used Taq polymerase (New England Biolabs [NEB], Ipswich, MA), following conditions in Appendix A and Appendix B with 10 mM dNTPs (NEB).

All reaction amplicons were verified with agarose gel electrophoresis using Tris-acetate-EDTA (TAE) buffer, and DNA was separated at 120 V for 30 minutes. Plasmids were generated

and maintained in *Escherichia coli* DH5- α cells, which were grown at 37 °C. Zymo ZR Plasmid Miniprep™- Classic kits (Zymo Research; Irvine, CA) were used in plasmid isolation, Zymo DNA Clean & Concentrator™-5 kits were used in the purification of PCR products, and Qiagen QIAEX II Gel Extraction Kits (Qiagen; Hilden, Germany) or Zymo Zymoclean Gel DNA Recovery Kits were used in the extraction of DNA.

The *E. coli* and *S. meliloti* strains listed in Table 2 were grown on media plates containing one or more antibiotics used at the following concentrations for all following sections: 500 μ g/mL streptomycin (Sm), 40 μ g/mL hygromycin (Hy), 50 μ g/mL chloramphenicol (Cm), and 50 μ g/mL spectinomycin (Sp).

Table 1. Oligonucleotide primers designed for this study.

5'				
Primer	Positio n	Restrictio n Site	F/ R	Primer Sequence
OCL 349	-36	PstI	F	TTTTctgcagCCTTGAAATGCCACATTTC AATCCA
OCL 350	177	PstI	R	TTTTctgcagGCTTTCGACCGTCGGGCAAAG
OCL 351	-400	SpeI	F	TTTTTactagtCGAAGATCACGGCCAGATGTG
OCL 352	-1	BamHI	R	TTTTTggatccGACATGATTTGTGGATTGAAATGT G
OCL 353	148	BamHI	F	TTTTTggatccTTTGTTCCTTTGCCCCGACGG
OCL 354	548	SacI	R	TTTTTgagctcCGAACGCTTGTCGTTCGGATG

Table 2. Bacterial strains used.

Strain	Antibiotic	Background	
Name	Resistance	d	Details
C10	Cm	<i>E. coli</i>	pMS4
C12	Sm	<i>S. meliloti</i>	Wild-type (WT) – Rm1021
C14	Cm	<i>E. coli</i>	Helper Strain
C26	Sm, Hy	<i>S. meliloti</i>	<i>chvI</i> K214T (Hy) marked with pDW181 at P-HisB
C27	Sm, Hy	<i>S. meliloti</i>	WT (Hy) – marked with pDW181 at P-HisB
C314	Cm	<i>E. coli</i>	<i>SmelC023</i> overexpression plasmid
C322	Sp, Sm, Hy	<i>S. meliloti</i>	<i>SmelC023</i> overexpression in <i>chvI</i> K214T mutant
C324	Sp, Sm, Hy	<i>S. meliloti</i>	<i>SmelC023</i> overexpression in WT

Overexpression Strain Construction

Primers OCL 349 and OCL 350 (Table 1) were used to PCR amplify *SmelC023*, along with 1 unit of Phusion polymerase (2 units/ μ L; Thermo Fisher) with and without dimethyl sulfoxide (DMSO), and 40 ng *S. meliloti* genomic DNA as a template. The plasmid vector pMS4 and *SmelC023* amplicon were both digested with 10 units PstI (NEB) in 1x NEB Buffer 3 for 2.5 hours at 37 °C. Five units of calf intestinal alkaline phosphatase (NEB) were added to the pMS4 digest and incubated for 5 additional minutes after the 2.5-hour digestion. Ligation of the digested amplicon and digested pMS4 vector was performed using 200 units of T4 DNA ligase (400 units/ μ L; NEB) and 1x T4 DNA Ligase Reaction Buffer (NEB) at room temperature for 3

hours. Five μL of the ligation mix was transformed into 50 μL competent DH5- α *E. coli* cells by heat shock at 42 °C for 45 seconds and transformants were selected using LB media plates with 50 $\mu\text{g/mL}$ Cm after growth at 37 °C for 25 hours. Colony PCR (Appendix B) was performed with 0.625 units of Taq DNA polymerase (5 units/ μL ; NEB) per reaction to check for successful ligation using T7 and T3 primers (IDT) with DNA from *E. coli* colonies lysed with boiling in 5mM Tris pH 8, 2 mM EDTA, 0.5% Triton X-100. The orientation of the *SmelC023* insert ligated into pMS4 was determined through PCR (Appendix B) with the T3 primer and primer OCL 349 or OCL 350 (Table 1). The overexpression plasmid was conjugated into the wild-type (C27) and *chvI* partial loss of function (C26) strains of *S. meliloti* through triparental mating, with an *E. coli* helper strain (C14) on LB media plates with Sm, Sp, Hy, and grown at 30 °C for about 24 hours.

Plate Assays

Overexpression and control strains were grown on LB media plates with Sm, Sp, Hy, as reported in Reagents and Equipment, at 30 °C for three days, then resuspended and serially diluted in Joel's minimal media (JMM2) until final optical densities of 1×10^{-4} OD/mL and 1×10^{-5} OD/mL were achieved. Five μL of these dilutions were spotted onto LB plates with Sm, Sp, Hy in addition to each of the following media plates. Calcofluor plates were prepared using LB media, Sm, Sp, Hy, and 0.02% calcofluor. Tryptone-yeast (TY) plates were prepared with Sm, Sp, Hy. Ethanol (EtOH) plates were prepared using LB media with Sm, Sp, Hy, and 4% EtOH. JMM2 plates were prepared using JMM2 salts, 0.3% sucrose, 1.5% agar, sterile double distilled water, and a mixture of magnesium sulfate, cobalt chloride, and biotin (MCB). Acidic plates were prepared using a 2x LB media, 80mM MES, and Sm, Sp, Hy, with NaOH added to pH

6.25. All plates were incubated for 1-2 days at 30 °C, then photographed using a Canon (Melville, NY) EOS Rebel camera. Calcofluor plates were also photographed with exposure to UV light.

Results

To investigate the role of *SmelC023*, overexpression studies were performed on free-living bacteria. ExoS/ChvI is known to regulate biofilm formation, motility, survival in acidic conditions, cell envelope integrity, succinoglycan production, and nutrient utilization in *S. meliloti* (Belanger et al., 2009; Wells et al., 2007; Yao et al., 2004, Zahran, 1999). Thus, the growth of *S. meliloti* strains on different types of media allows for the detection of differences between different strains under discrete conditions. While it was expected that the *chvI* partial LOF (K214T) strain will have less growth and smaller colony sizes than the wild-type strain on LB medium, the overexpression of *SmelC023* was investigated to potentially observe any additional phenotypes. *SmelC023* was overexpressed in both a wild-type background and in a *chvI* partial LOF background, to see if overexpression of *SmelC023* would rescue any phenotypes caused by the LOF mutation. The wild-type Rm1021 strain with empty vector and the *chvI* partial LOF strain with empty vector were used as control strains. Five microliters of each strain were spotted onto different media plates at either 10^{-4} OD₆₀₀ mL⁻¹ to observe overall growth or 10^{-5} OD₆₀₀ mL⁻¹ to be able to observe the growth of single colonies. Each phenotypic figure shows results from one trial, except for Figures 1, 3, and 4 which are representative of two trials.

S. meliloti strains were grown on LB and TY media plates to observe growth on two types of nutrient-rich media (Figs. 1 and 2). As expected, the *chvI* partial LOF strain showed a growth defect compared to the wild-type *S. meliloti*. Comparing the wild-type *S. meliloti* with an

empty vector to the wild-type overexpressing *SmelC023*, no difference in overall growth or single colony size was observed. Similar results were observed between the *chvI* partial LOF *S. meliloti* strain with an empty vector and *chvI* partial LOF strain with *SmelC023* overexpressed (Figs. 1 and 2).

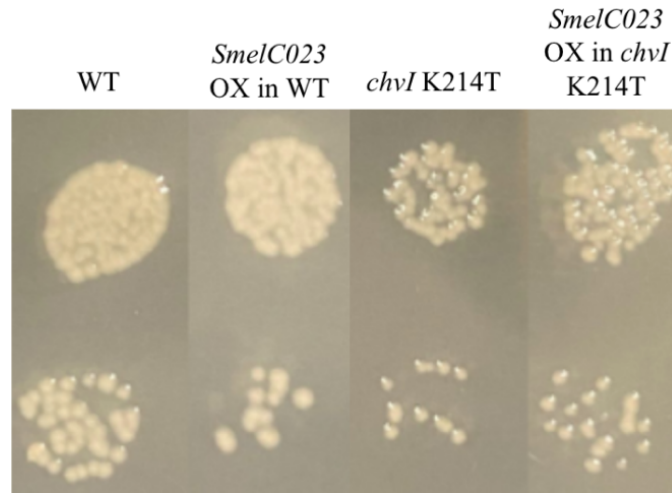


Figure 1. *SmelC023* overexpression did not affect growth on LB media. Five microliters of strains diluted to 10^{-4} OD₆₀₀ mL⁻¹ (top row) or 10^{-5} OD₆₀₀ mL⁻¹ (bottom row) then grown on LB media plates with Sm, Sp, and Hy for three days at 30°C. From left to right, strains shown are: WT (Rm1021), *SmelC023* overexpression in WT, *chvI* partial LOF, and *SmelC023* overexpression in *chvI* partial LOF. Results shown represent two trials.

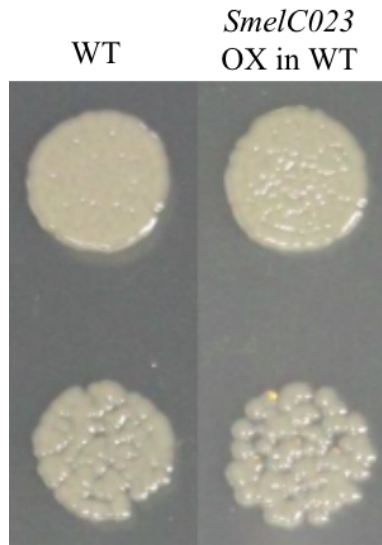


Figure 2. *SmelC023* overexpression did not affect growth on tryptone-yeast (TY) media.

Five microliters of strains diluted to 10^{-4} OD₆₀₀ mL⁻¹ (top row) or 10^{-5} OD₆₀₀ mL⁻¹ (bottom row) then grown on TY media plates with Sm, Sp, and Hy for three days at 30°C. From left to right, strains shown are: WT (Rm1021) and *SmelC023* overexpression in WT. Results shown represent one trial.

S. meliloti strains were grown on LB media plates with JMM2 (Fig. 3) and 4% EtOH (Fig. 4) and media plates to observe any effects on *S. meliloti* nutrient utilization and cell envelope integrity, respectively. JMM2 shows how the strains grow on a minimal medium, and EtOH plates allow for the detection of cell envelope integrity defects since EtOH disrupts cell membranes. On JMM2, the *chvI* partial LOF strain was expected to grow similarly to the wild-type strain, matching what was observed in these results. Again, there were no differences in overall growth or single colony size of the wild-type and partial LOF strains compared to the strains overexpressing *SmelC023* in those backgrounds. Together, these results indicate that *SmelC023* overexpression did not affect nutrient utilization or cell envelope integrity in *S. meliloti*.

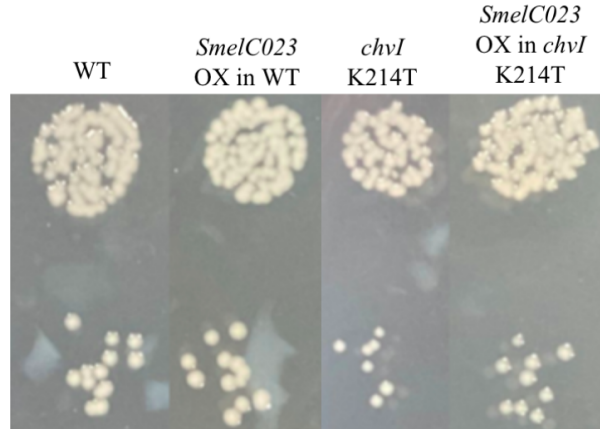


Figure 3. *SmelC023* overexpression did not affect growth on minimal media. Five microliters of strains diluted to 10^{-4} OD₆₀₀ mL⁻¹ (top row) or 10^{-5} OD₆₀₀ mL⁻¹ (bottom row), then grown on JMM2 media plates with Sm, Sp, and Hy for three days at 30°C. From left to right, strains shown are: WT (Rm1021), *SmelC023* overexpression in WT, *chvI* partial LOF, *SmelC023* overexpression in *chvI* partial LOF. Results shown represent two trials.

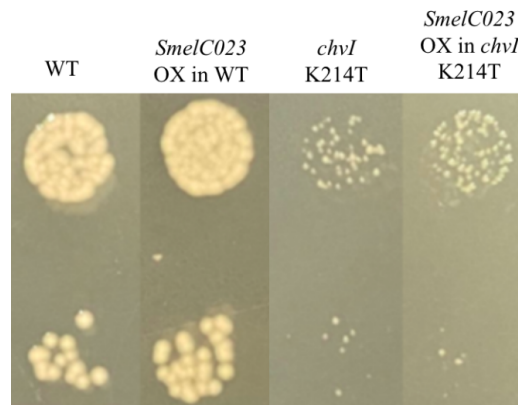


Figure 4. *SmelC023* overexpression did not affect cell envelope integrity. Five microliters of strains diluted to 10^{-4} OD₆₀₀ mL⁻¹ (top row) or 10^{-5} OD₆₀₀ mL⁻¹ (bottom row), grown on LB+EtOH media plates with Sm, Sp, and Hy for three days at 30°C. From left to right, strains shown are: WT (Rm1021), *SmelC023* overexpression in WT, *chvI* partial LOF, *SmelC023* overexpression in *chvI* partial LOF. Results shown represent one trial.

The ability to grow in acidic conditions is important for *S. meliloti* since the root nodule environment formed during host invasion is believed to be acidic (Zahran, 1999). To test this phenotype, *S. meliloti* strains were grown on LB media plates at pH 6.25. Preliminary results showed no difference in growth or single colony size between strains where *SmelC023* is overexpressed or not overexpressed, indicating that *SmelC023* has no effect on growth in acidic conditions (Fig. 5).

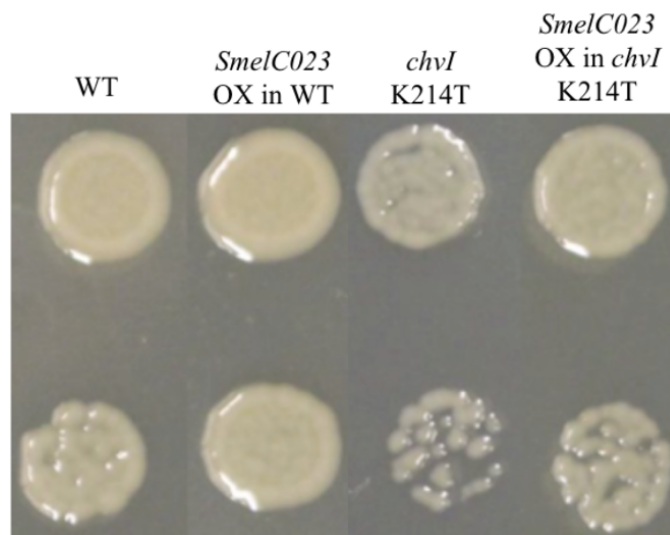


Figure 5. *SmelC023* overexpression did not affect growth in acidic conditions. Five microliters of strains diluted to 10^{-4} OD₆₀₀ mL⁻¹ (top row) or 10^{-5} OD₆₀₀ mL⁻¹ (bottom row) then grown on LB media plates at pH 6.25 with Sm, Sp, and Hy for three days at 30°C. From left to right, strains shown are: WT (Rm1021), *SmelC023* overexpression in WT, *chvI* partial LOF, and *SmelC023* overexpression in *chvI* partial LOF. Results shown represent one trial.

Calcofluor specifically binds to succinoglycan and fluoresces under UV light, making fluorescence a proxy for succinoglycan production. The resulting growth of the wild-type *S. meliloti* overexpressing *SmelC023* and the wild-type control strain on LB media with calcofluor

under UV light showed no differences in fluorescence between the two strains (Fig. 6). There was also no difference in fluorescence observed between the *chvI* partial LOF *S. meliloti* strain and the *SmelC023* overexpression in the *chvI* partial LOF background. Likewise, the strains with and without *SmelC023* overexpression show no noticeable differences in overall growth and single colony sizes. Together, these results indicate that overexpression of *SmelC023* had no noticeable effect on succinoglycan production in *S. meliloti*.

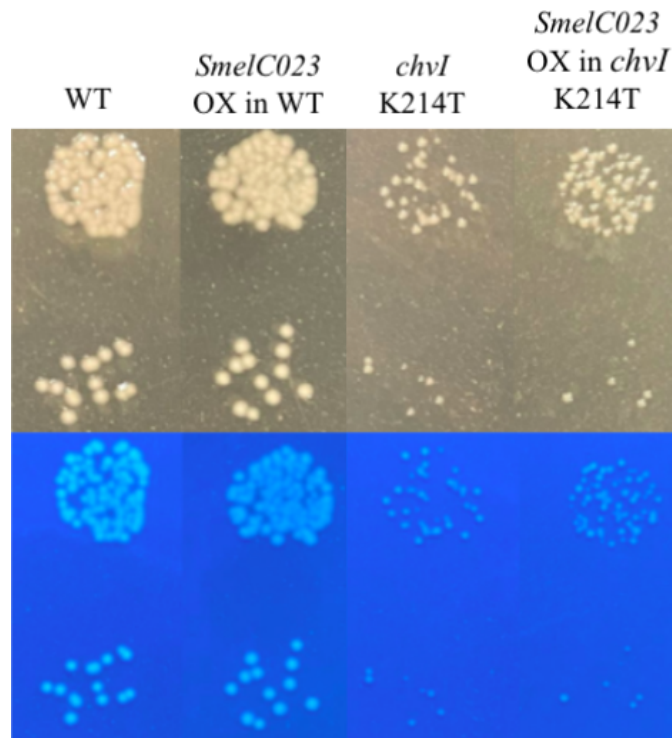


Figure 6. *SmelC023* overexpression did not affect succinoglycan production. Five microliters of strains diluted to 10^{-4} OD₆₀₀ mL⁻¹ (top row) or 10^{-5} OD₆₀₀ mL⁻¹ (bottom row) in each of the two panels were grown on LB+calcofluor media plates with Sm, Sp, and Hy for three days at 30°C, photographed under visible (upper panel) and UV light (lower panel). From left to right, strains shown are: WT (Rm1021), *SmelC023* overexpression in WT, *chvI* partial LOF, *SmelC023* overexpression in *chvI* partial LOF. Results represent shown one trial.

Discussion

This project aimed to characterize the functions of sRNA direct target genes of the ExoS/ChvI pathway, namely *SmelC023*. Overexpression strains were created by cloning each gene of interest behind a strong constitutive promoter, then introducing this construct as a plasmid into *S. meliloti* backgrounds via triparental mating.

The preliminary results presented here indicate that the overexpression of *SmelC023* does not affect succinoglycan production, growth on TY, growth in acidic conditions, cell envelope integrity, or growth on minimal media in free-living *S. meliloti*. Although overexpression of *SmelC023* had no effect on the phenotypes studied here, overexpression may show a distinct phenotype in other free-living bacterial functions known to be regulated by ExoS/ChvI, including biofilm formation. To further investigate the results of *SmelC023* overexpression presented in this study, it would be beneficial to conduct quantitative analyses of the succinoglycan production phenotype to determine the magnitude of difference in this function. This is important because small differences may be overlooked when conclusions are drawn solely from visual analysis. Additionally, a 2007 study by del Val et al. demonstrated that the expression of *SmelC023* is highly induced in endosymbiotic bacteria, and it was speculated that *SmelC023* may be involved in infection and/or bacterial differentiation (del Val et al., 2007). Therefore, the failure to detect a free-living phenotype in the assays performed here does not exclude a possible role for *SmelC023* during symbiosis. Furthermore, the effects of *SmelC023* deletion were not investigated in this study, and testing the phenotypes of a *SmelC023* knockout strain can provide new insights into both free-living and symbiotic potential functions. Together, these findings build upon previous research and provide new insights into the potential role of *SmelC023* in *S. meliloti* to guide future investigations.

References

- Belanger, L., Dimmick, K. A., Fleming, J. S., & Charles, T. C. (2009). Null mutations in *Sinorhizobium meliloti* *exoS* and *chvI* demonstrate the importance of this two-component regulatory system for symbiosis. *Mol Microbiol*, 74(5), 1223-1237. doi:10.1111/j.1365-2958.2009.06931.x
- Chen, E. J., Fisher, R. F., Perovich, V. M., Sabio, E. A., & Long, S. R. (2009). Identification of direct transcriptional target genes of ExoS/ChvI two-component signaling in *Sinorhizobium meliloti*. *J Bacteriol*, 191(22), 6833-6842. doi:10.1128/JB.00734-09
- Chen, E. J., Sabio, E. A., & Long, S. R. (2008). The periplasmic regulator ExoR inhibits ExoS/ChvI two-component signalling in *Sinorhizobium meliloti*. *Mol Microbiol*, 69(5), 1290-1303. doi:10.1111/j.1365-2958.2008.06362.x
- Cheng, H. P., & Walker, G. C. (1998). Succinoglycan production by *Rhizobium meliloti* is regulated through the ExoS-ChvI two-component regulatory system. *J Bacteriol*, 180(1), 20-26. doi:10.1128/JB.180.1.20-26.1998
- del Val, C., Rivas, E., Torres-Quesada, O., Toro, N., & Jimenez-Zurdo, J. I. (2007). Identification of differentially expressed small non-coding RNAs in the legume endosymbiont *Sinorhizobium meliloti* by comparative genomics. *Mol Microbiol*, 66(5), 1080-1091. doi:10.1111/j.1365-2958.2007.05978.x
- Gibson, K. E., Kobayashi, H., & Walker, G. C. (2008). Molecular Determinants of a Symbiotic Chronic Infection. In *Annual Review of Genetics* (Vol. 42, pp. 413-441). Palo Alto: Annual Reviews.
- Laub, M. T. (2011). The Role of Two-Component Signal Transduction Systems in Bacterial Stress Responses. In G. Storz & R. Hengge (Eds.), *Bacterial Stress Responses* (2nd ed.). Washington D.C.: ASM Press.
- Long, S. R. (2016). SnapShot: Signaling in Symbiosis. *Cell*, 167(2), 582-582 e581. doi:10.1016/j.cell.2016.09.046
- Ratib, N. R., Sabio, E. Y., Mendoza, C., Barnett, M. J., Clover, S. B., Ortega, J. A., . . . Chen, E. J. (2018). Genome-wide identification of genes directly regulated by ChvI and a consensus sequence for ChvI binding in *Sinorhizobium meliloti*. *Molecular Microbiology*, 110(4), 596-615. doi:10.1111/mmi.14119
- Raymond, J., Siefert, J. L., Staples, C. R., & Blankenship, R. E. (2004). The Natural History of Nitrogen Fixation. *Molecular Biology and Evolution*, 21(3), 541-554. doi:10.1093/molbev/msh047
- Schluter, J. P., Reinkensmeier, J., Daschkey, S., Evguenieva-Hackenberg, E., Janssen, S., Janicke, S., . . . Becker, A. (2010). A genome-wide survey of sRNAs in the symbiotic nitrogen-fixing alpha-proteobacterium *Sinorhizobium meliloti*. *BMC Genomics*, 11, 245. doi:10.1186/1471-2164-11-245
- Socolow, R. H. (1999). Nitrogen management and the future of food: Lessons from the management of energy and carbon. *Proceedings of the National Academy of Sciences*, 96(11), 6001-6008. doi:doi:10.1073/pnas.96.11.6001
- Storz, G., Vogel, J., & Wassarman, K. M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell*, 43(6), 880-891. doi:10.1016/j.molcel.2011.08.022
- Wells, D. H., Chen, E. J., Fisher, R. F., & Long, S. R. (2007). ExoR is genetically coupled to the ExoS-ChvI two-component system and located in the periplasm of *Sinorhizobium meliloti*. *Mol Microbiol*, 64(3), 647-664. doi:10.1111/j.1365-2958.2007.05680.x

- Yao, S. Y., Luo, L., Har, K. J., Becker, A., Ruberg, S., Yu, G. Q., . . . Cheng, H. P. (2004). Sinorhizobium meliloti ExoR and ExoS proteins regulate both succinoglycan and flagellum production. *J Bacteriol*, 186(18), 6042-6049. doi:10.1128/JB.186.18.6042-6049.2004
- Zahran, H. H. (1999). Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol Rev*, 63(4), 968-989. doi:10.1128/MMBR.63.4.968-989.1999

APPENDIX A.

Thermocycler Conditions with Phusion Polymerase

1. 98°C for 3 minutes
2. 98°C for 10 seconds
3. Annealing temperature for 20 seconds (Skip if annealing temperature is greater than 72°C)
4. Repeat Steps 2-3 32 times (33 cycles total)
5. 72°C for 5 minutes
6. 4°C hold

APPENDIX B.

Thermocycler Conditions with Taq Polymerase

1. 94°C for 3 minutes
2. 94°C for 30 seconds
3. 55°C for 30 seconds
4. 72 °C 1 min per 1kb length
5. Repeat Steps 2-4 four times (5 cycles total)
6. 94°C for 30 seconds
7. 58°C for 30 seconds
8. 72°C 1 min per 1kb length
9. Repeat Steps 6-8 24 times (25 cycles total)
10. 72°C for 10 minutes
11. 4°C hold

Effects of grazing-related vegetation changes and moonlight on foraging behavior of shortgrass-prairie rodents

Marlene Lopez

Adviser: Dr. Paul Stapp

Abstract

Predation risk is a factor that highly influences rodent foraging behavior. Rodents alter their behavior in response to cues of predation risk, like levels of moonlight and changes in vegetation. I used artificial seed trays to analyze the effects of moonlight and grazing-related differences in vegetation on the foraging behavior of nocturnal rodents in the shortgrass prairie. Seed trays were placed in either open or shrub microhabitats for one night (new, quarter, full moon) in one of four different grazing treatments. Trays were recorded as either “visited” or “not visited” by a forager, and the amount of seed (g) removed from trays was recorded. I examined two response variables: the percentage of visited trays and amount of seed (g) removed from visited trays. Rodents visited trays most frequently during darker, new moon nights and visited treatments with lowest amount of grazing most frequently. The most seed was removed from trays during darker, new moon nights, but there wasn’t any difference between amount of seed removed from visited trays between grazing treatments. Rodents altered their foraging behavior in response to changes in moonlight and differences of grazing levels, presumably because of perceived predation risk.

Introduction

Behavioral decision-making by animals reflects tradeoffs between foraging efficiency and minimizing risk of predation (Lima and Dill 1990). For many small mammals, movements and microhabitat use reflect both the dispersion and availability of food as well as the need to avoid detection and capture by predators (Stapp and Lindquist 2007). In open environments such as deserts and grasslands, the amount of moonlight can influence predation risk and thus rodent foraging behavior. For example, predators are often more effective hunters during full moon nights (Clarke 1983, Longland and Price 1991), leading nocturnal rodents to alter their behavior to become more vigilant and more risk averse. On bright nights, desert rodents may move shorter distances, spend less time foraging, or use vegetation cover, e.g., shrubs, more (Longland and Price 1991). On darker nights, rodents may be more active aboveground and venture further from their burrows to forage or find mates (Longland and Price 1991). Moonlight thus may serve as a cue that rodents use to assess predation risk, ultimately causing them to alter their activity levels and aboveground behavior in response to ambient light caused by different phases of the moon.

Because foraging by rodents is sensitive to predation risk, and perceived or actual risk is mediated by the availability of protective cover, factors that influence the amount or type of vegetation can affect rodent foraging behavior and, ultimately, patterns of abundance and community structure. In grassland ecosystems, grazing by native herbivores and livestock are a key determinant of plant productivity, community composition, and vegetation structure (Hobbs 1996, Eldridge et al. 2016). Rodent communities in intensively grazed areas tend to differ from those in enclosures or in areas with less intensive grazing (Grant et al. 1982, Thompson and Gese 2013, Ellis and Cushman 2018), although the exact ecological mechanisms underlying these

differences have not been studied thoroughly. The influence of grazing on the activity and behavior small mammals could ultimately feed back to affect plant communities because rodents can be important seed predators, especially for large-seeded forbs and shrubs (Brown and Heske 1990, Hoffman et al. 1995, Maron et al. 2012, Larios et al. 2017).

I used artificial seed trays to examine the interactive effects of moonlight and grazing-related differences in vegetation on the foraging behavior of nocturnal rodents in the shortgrass prairie in north-central Colorado. I estimated overnight seed removal during foraging trials conducted during new, quarter- and full-moon phases, in plots where access to large and medium-sized herbivores were manipulated experimentally to examine the effects of herbivory on woody shrub encroachment. I expected that, in trays set out overnight, rodents would visit trays set out under shrubs than in open microhabitats, visit more trays on darker, new-moon nights than bright, full-moon nights, and visit more trays in ungrazed plots with taller vegetation than grazed plots with more bare ground and shorter vegetation. Of the trays that were visited, I expected that rodents would remove more seeds in trays beneath shrubs than in open microhabitats, remove more seeds during darker, new-moon nights than on bright, full-moon nights, and remove more seeds in ungrazed plots with taller vegetation than grazed plots with more bare ground and shorter vegetation.

Methods

I conducted my research between May and August 2022, at the Semi-arid Grasslands Research Center (SGRC), located at the USDA-ARS Central Plains Experimental Range (CPER) in north-central Colorado. Vegetation at the study site is classified as shortgrass prairie, with a great

diversity of grasses, shrubs, cacti, and forbs (Lauenroth and Burke 2008). The climate is semiarid, with cold winters and hot summers, and most precipitation falling as rain during the spring and summer growing season (Lauenroth and Burke 2008). Common nocturnal rodent species at the CPER include the Ord's kangaroo rat (*Dipodomys ordii*), northern grasshopper mouse (*Onychomys leucogaster*), deer mouse (*Peromyscus maniculatus*), and the western harvest mouse (*Reithrodontomys megalotis*; Lauenroth and Burke 2008).

I conducted my research at five CPER sites (19N, 19S, 24W, 11S, 5E; Figure 1) that were part of an ongoing study initiated in 2012 to investigate effects of different-sized herbivores on potential encroachment by four-wing saltbush (*Atriplex canescens*). Each site consisted of four experimental plots, each measuring 50 x 50 meters, and representing a different experimental treatment. "All herbivores" (AH) plots were not fenced and allowed access to all herbivores, including pronghorn (*Antilocapra americana*) and cattle (*Bos taurus*), the latter of which were stocked at a moderate grazing intensity during the growing season. "Cattle excluded" (CE) plots were fenced using 4 strands of barbed wire, which prevented cattle from entering, but still allowed smaller herbivores such as jackrabbits (*Lepus* sp.) to enter. "Rabbits excluded" (RE) plots were fenced with both barbed wire and chicken wire, but two steers were placed in the plot for two days (one day in June, one day in August) to simulate moderate-intensity grazing. Lastly, "cattle and rabbits excluded" (CRE) plots were fenced using barbed wire and chicken wire, permitting access to only small rodents and juvenile cottontail rabbits (*Sylvilagus audubonii*).

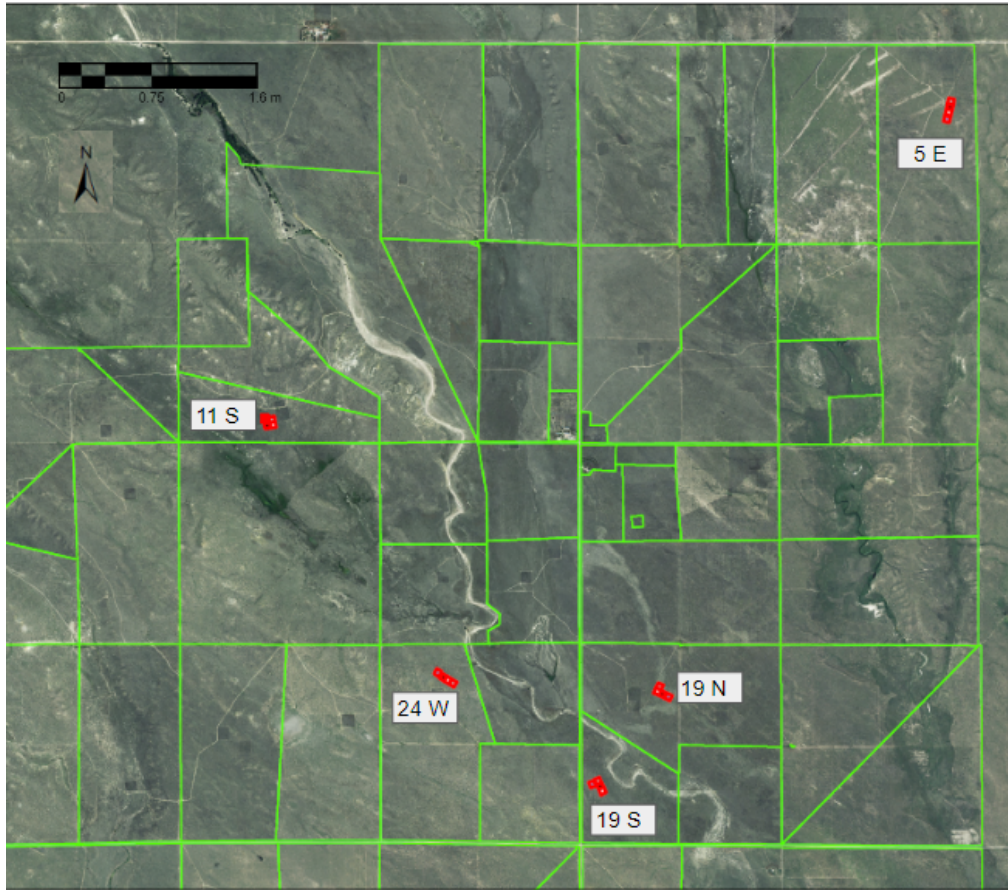


Figure 1. Map of the Semi-arid Grasslands Research Center (SGRC), located on the USDA-ARS Central Plains Experimental Range (CPER) in northern Colorado. The five study sites, each with three enclosures and one unfenced control plot, are labeled. Map provided by USDA-ARS.

To determine rates of seed removal by rodents, I set artificial seed trays in each treatment plot for one night during quarter-, full-, and new- moon phases. Quarter-moon trials were conducted between 21-23 June, full-moon trials were conducted from 14-16 July, and new-moon trials were conducted between 27-31 July. Because of the distance between and the need to check immediately at dawn, all four plots at a site (16 trays per site) were sampled on the same night.

Seed trays were circular, terra-cotta plant saucers constructed of stiff plastic, measuring 1,964 cm² in area and 8.9 cm deep. Two pairs of seed trays were placed into each treatment plot. One was placed beneath the canopy of a large shrub (usually *Atriplex canescens*), with the other placed 1-2 m away in the open, at least 1 m from any shrubs or dense vegetation (Figure 2). Each tray contained 8 g of commercial milo millet mixed with 2 L of sieved play sand. I set the trays at dusk and then checked them at dawn. I recorded any signs of foraging behavior (tail drags, disturbed sand, footprints) and labeled these trays as “visited” (Figure 2C). If no foraging signs were observed, the tray was considered “not visited”. After returning the seed to the lab, I sieved the remaining millet from the sand and weighed it to estimate the amount of seed removed.

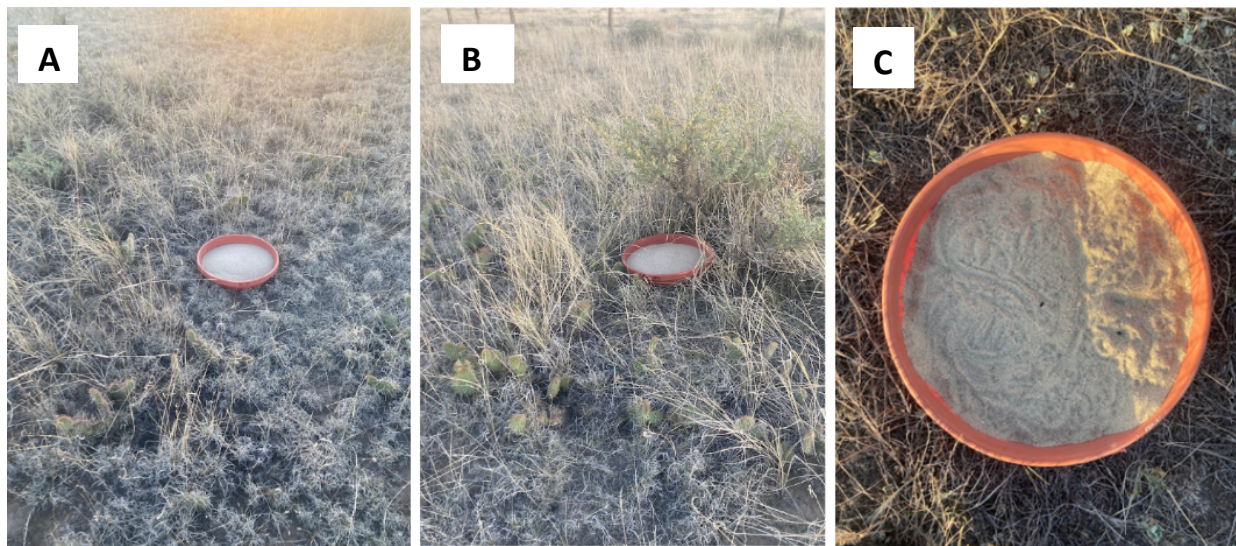


Figure 2. Images of artificial seed trays used in rodent foraging trials at the Central Plains Experimental Range, Colorado. A. Tray placed in open cover type. B. Tray placed in shrub cover type. C. Foraged seed tray showing evidence of visitation (scats, tail drags).

I analyzed two response variables: the proportion of trays in a given treatment type that were visited and the amount (g) of seed removed from visited trays. I used a three-way ANOVA to

test the effects of moonlight, grazing treatment (AH, CE, RE, CRE), and microhabitat (shrub, open) on the percentage of trays visited and the amount of seed removed from visited trays.

Results

Vegetation in plots was measured in most years of the study by USDA-ARS field crews. Table 1 shows the treatment means of some representative variables in 2022. CRE and RE plots had the highest canopy cover of mid grasses, followed CE plots. AH plots had lowest canopy percent cover of mid grasses.

Table 1. Mean (+ 1 SE) of vegetation measurements in treatment plots at the CPER in 2022 . Sample size was five plots per treatment.

Treatment	Canopy percent cover of "mid" grasses	Index of vegetation thickness/height	Density of Atriplex per m ²
AH	0.68 ± 1.0	16.75 ± 2.7	0.14 ± 0.01
CE	1.25 ± 1.8	28.72 ± 2.1	0.15 ± 0.01
CRE	1.90 ± 2.5	27.58 ± 4.0	0.16 ± 0.01
RE	1.40 ± 4.6	22.72 ± 3.3	0.11 ± 0.01

Based on live-trapping studies conducted concurrently by M.S. student Katie Biardi in 2022 (Table 2), the highest number of nocturnal rodents were captured in CRE plots (19), followed by RE plots(18). The fewest nocturnal rodents were captured in AH and CE (15). Ord's kangaroo

rat, a large granivore, was most common rodent caught in all treatments, making up 40-68% of captures. Omnivorous rodents (ONLE, PEMA, REME) were especially common on AH and CE plots (40% and 53% of total captured, respectively).

Table 2. Total number of nocturnal rodents live-trapped in four treatments (AH, CE, RE, CRE) at the Central Plains Experimental Range in north-central Colorado in June and July 2022.

DIOR = Dipodomys ordii,; CHHI = Chaetodipus hispidus, PEFL = Perognathus flavus, ONLE = Onychomys leucogaster, PEMA = Peromyscus maniculatus, REME = Reithrodontomys megalotis. Data from Katie Biardi, M.S. student, Department of Biological Science, CSUF.

	DIOR	CHHI	PEFL	ONLE	PEMA	REME	Total
AH	7	0	1	4	1	2	15
CE	6	1	0	2	2	4	15
RE	12	0	3	0	3	0	18
CRE	13	1	0	0	1	4	19

Based on results of a three-way ANOVA, there was no significant difference between the percentage of trays visited in open and shrub cover types ($p=0.263$, $F=1.37$, $df=1$; Figure 3).

There was also no interaction between treatment and cover type, nor between moon phase and cover type. Pooling across cover types, the percentage of trays visited differed between treatment ($p<0.0001$, $F=27.1$, $df=3$) and moon phases ($p<0.0001$, $F=23.2$, $df=2$; Figure 4). There was also a significant interaction between moon phase and treatment ($p=0.02$, $F=3.95$, $df=6$). Overall, during all moon phases, the fewest trays were visited in AH plots, whereas most of the trays in

CRE and RE plots were visited. The greatest difference in visitation was between AH and fenced plots on new-moon nights.

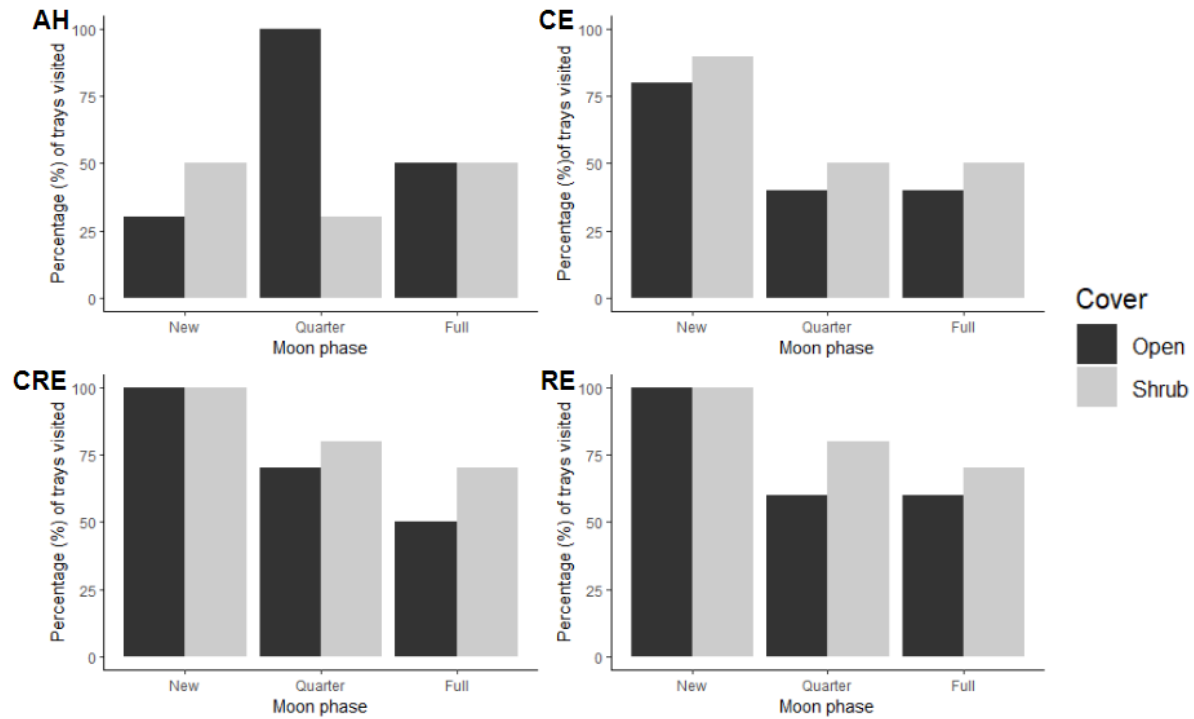


Figure 3. Percentage of trays visited in open (o) and shrub (s) cover types in four grazing treatments (AH, CE, RE, CRE) during three different moon phases (New, Quarter, Full) in shortgrass prairie in north-central Colorado in June and July 2022. There were five replicate sites in each treatment type and two trays per cover type, for a total of 10 trays per treatment and moon phase (trays were set for one night during each moon phase).

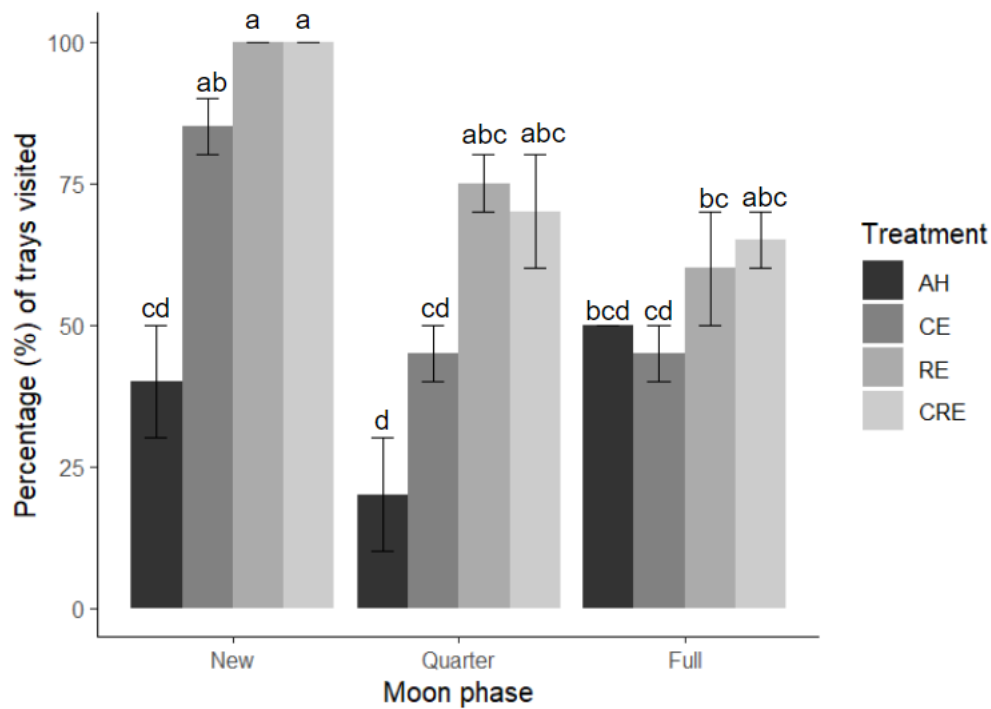


Figure 4. Pooling across cover types, the percentage of trays visited (mean + SE) in four grazing treatments (AH, CE, RE, CRE) during three different moon phases (new, quarter, full) in shortgrass prairie in north-central Colorado in June and July 2022. There were four trays per treatment type and five replicates of each treatment type for a total of 20 trays per treatment type in each moon phase (trays were set for one night during each moon phase). Bars with letters were not significantly different ($P > 0.05$, Tukey HSD tests).

Considering only trays that were visited, the amount of seed removed (g) from trays did not differ significantly between “open” and “shrub” cover types ($p=0.918$, $F=0.01$, $df=1$), nor between treatment types ($p=0.087$, $F=2.24$, $df=3$), and there was no significant interaction between moon phase and treatment or moon phase and cover. The mean amount of seeds removed from visited trays tended to be lowest in AH plots (Figure 5). Across all treatments and cover types, the most seeds were removed during new- moon nights and the lowest during

quarter- and full- moon nights (Figure 5). Combining across cover and treatment types, significantly more seed was removed from visited trays on new-moon nights than quarter- or full-moon nights ($p=0.0008$, $F=7.51$, $df=2$; Figure 6)

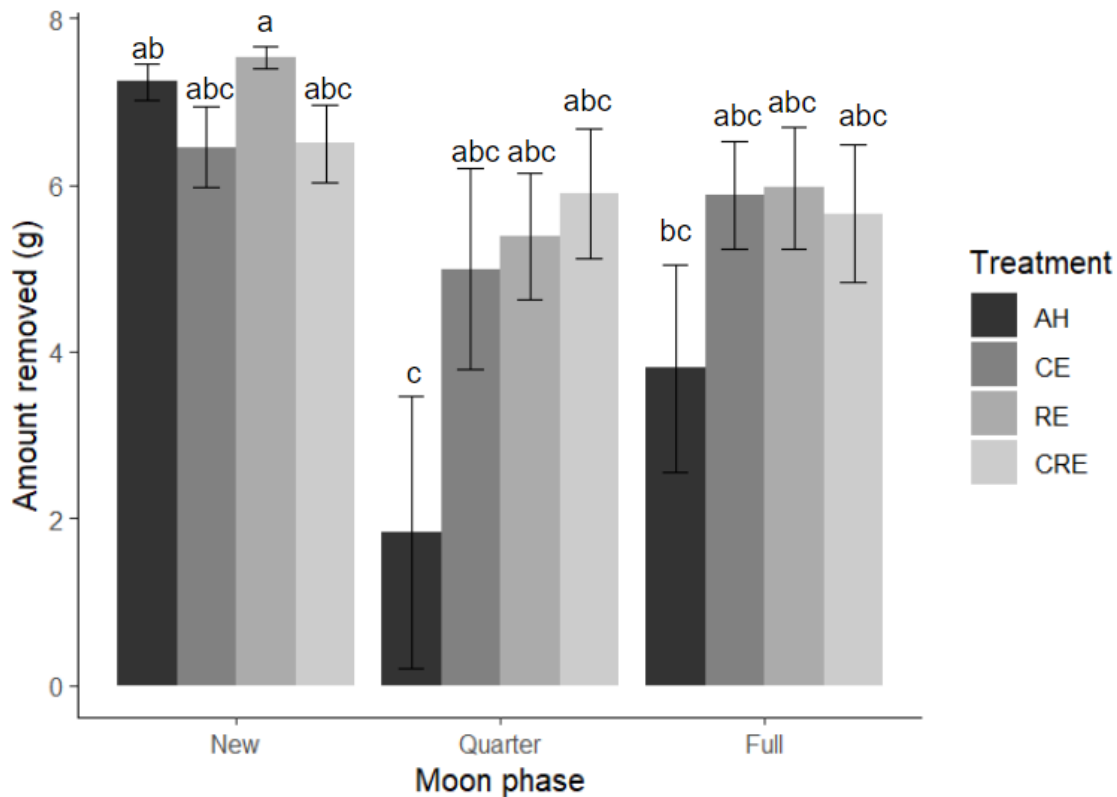


Figure 5. Mean (± 1 SE) amount of millet seed removed (g) from a total of 8 g of seed placed in one of the four grazing treatments (AH, CE, RE, CRE) during three different moon phases (New, Quarter, Full) in shortgrass prairie in north-central Colorado in June and July 2022. Only trays actually visited by rodents were included, so sample size varied from 4 to 20.

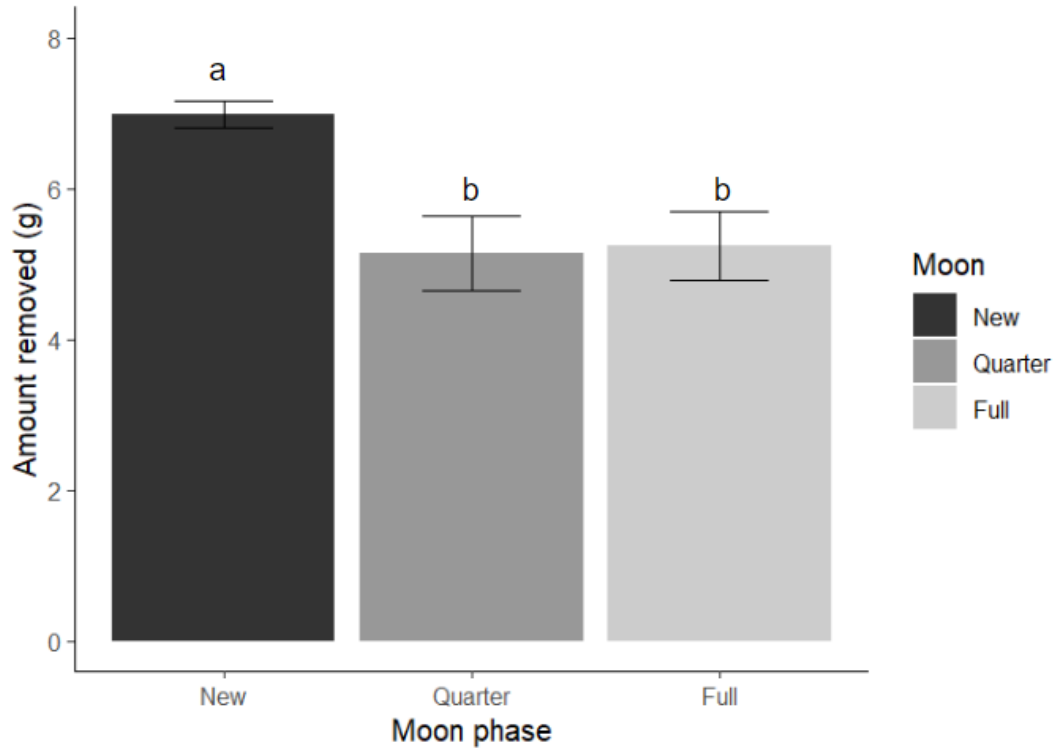


Figure 6. Considering only trays visited by rodents and pooling across treatments trays, mean (+1 SE) amount of millet seed removed (g) from a total of 8 grams of seed placed in trays during three moon phases (New, Quarter, Full) in shortgrass prairie in north-central Colorado in June and July 2022. Sample sizes were 65, 42, and 44 for New, Quarter and Full moon nights, respectively.

Discussion

To understand the effects of moonlight and vegetation cover on rodent foraging behavior, I analyzed two variables: the percentage of trays that were visited by a nocturnal forager and the amount of seed removed from those visited trays. I was able to tell if rodents visited trays by the presence of footprints, tail drags, or feces and by the volume of sand moved. In many cases, the

most likely visitors were kangaroo rats, which are more than twice the size of mice, have large tails and hind feet, and are capable of moving large amounts of sand during foraging. The percentage of trays visited reflected the activity of nocturnal rodents and their willingness to forage in a variety of vegetation cover types and moon-light levels. The amount of seed removed from a tray was used as a measure of how long they spent in trays in different microhabitats in areas with different vegetation.

Surprisingly, microhabitat cover type did not seem to affect nocturnal rodent foraging behavior. I had hypothesized that rodents would prefer to forage in shrub microhabitats because of the increased cover that shrubs provide from predators, thus reducing predation risk (Kotler 1984, Rosenzweig 1973, Orrock et al. 2004). Despite my prediction that rodents would visit trays placed in shrub microhabitats more often than open ones, there was no significant difference in visitation of seed trays between microhabitats (Figure 3). I also predicted that of the visited trays, rodents would remove more seeds in shrub trays, but there were no significant differences in seed removal between microhabitats (Figure 4). This lack of preference for microhabitat shown was likely because the Ord's kangaroo rat, a bipedal granivore, was the most common nocturnal rodent across all treatments (Table 2). Differences in morphology between bipedal and quadrupedal rodent species can influence foraging behavior and preference. Bipedal rodents, such as kangaroo rats, have large auditory bullae and strong hindlimbs that allow them to detect predators easily and quickly evade them, while quadrupedal rodents, like the other rodents at the CPER, lack these traits (Brown et al. 1988). Consequently, quadrupedal rodent species tend to prefer shrub microhabitats, while bipedal species can prefer to forage in open microhabitats (Longland and Price 1991). For a bipedal rodent like a kangaroo rat, differences between the

perceived risk of visiting and potentially foraging in an open microhabitat compared to a shrub microhabitat might not be critical.

Grazing treatment affected whether or not rodents visited trays, but not the amount of seed removed. Foragers visited seed trays in CRE and RE treatments most frequently, followed by CE treatments (Figure 4). AH treatments had the lowest percentage of visits (Figure 3). This pattern closely follows the rodent trapping numbers of treatments in 2022 (Table 2). Where rodents were abundant (CRE, RE) seed trays were visited more often, and where there were fewer rodents (CE, AH), fewer seed trays were not visited. Rodent abundance therefore may determine the amount of foraging activity that occurs in a plot: the more rodents there are in a treatment, the higher the percentage of trays visited. Differences between treatments in rodent abundance may reflect differences in vegetation caused by removal of herbivores.

I had predicted that rodents would remove more seeds in ungrazed plots with taller vegetation than grazed plots with more bare ground and shorter vegetation because of the protective cover from predators that vegetation would provide (CRE>RE>CE>AH; Table 1). However, the amount of seed removed from visited trays did not differ significantly between treatment (Figure 5). This was most likely because of the rodent population effect on the percentage of seed trays visited. Once a forager encountered a seed tray and made the decision to forage in it, the type and height of the vegetation around them had less of an effect on seed removal than the level of moonlight.

Moonlight was a significant determinant of foraging preference and activity. Foragers visited seed trays more frequently during the new moon than the quarter and full moon, with little difference in number of visits between quarter and full moon (Figure 4). Rodents also removed the most seeds from visited trays during new moon, while there was little difference

between amount of seed removed during quarter and full moon (Figure 5). During the new moon, rodents removed almost all seeds from the few trays visited in AH plots, the most “risky” plot due to the low vegetation cover caused by grazing, at the same level as the other three treatments (Figure 5). But on quarter and full-moon nights, rodents removed less seeds in AH than in the other three treatments (Figure 6).

This finding is consistent with previous studies showing higher levels of nocturnal activity during darker nights compared to bright nights (Kotler 2010 et al., Clarke 1983). Rodents significantly altered their behavior in response to a decrease in light because predation risk is presumably lower, allowing them to forage for more seeds (Prugh 2014). When illumination levels decreased during the new moon, the perceived risk of predation decreased, allowing rodents to forage in AH treatments at the same level as CE, CRE, and RE treatments. When illumination increased during the full moon, foragers did not remove seeds in AH treatments at the same level as in the other three treatments, perhaps because the perceived risk of predation in the grazed AH treatments outweighed the benefit of millet seed.

I had also expected that more seed would be removed during the quarter moon than the full moon, but that was not the case (Figure 5). Seeds in artificial trays represent a rich but unfamiliar resource to foraging rodents. Because the first round of seed trays set out were during the quarter moon, it is possible that rodents might have shown some neophobia towards the trays (Barnett 1958), resulting in low seed removal during the first round of seed trays set out during the quarter moon.

Both moonlight and the intensity of grazing by herbivores affect the foraging habits of nocturnal rodents. Intensity of grazing in each treatment highly influences the type and amount of vegetation available, which in turn influences the type and numbers of rodents present.

Rodents were most numerous in treatments that have low grazing intensity by cattle (CRE) and lowest population in treatments that experience high levels of grazing (AH). The number of rodents in a treatment type was a good indicator of the percentage of seed trays that were visited, suggesting that rodents may prefer ungrazed areas with more vegetation that provide cover from predators, such as owls (Zimmerman et al. 1996). Once a rodent visited a tray, decisions about how much seed to remove and how much time to spend potentially exposed to predators was most influenced by moon phase, with the greatest amount of seeds removed on dark, new-moon nights.

Predation risk due to levels of illumination and grazing-related vegetation changes highly influences the foraging behavior of rodents and in turn, the abundance of rodents in the shortgrass prairie. Rodents, especially kangaroo rats, are important seed predators that influence the vegetation around them (Brown and Heske 1990, Larios et al. 2017). The influence of livestock grazing on the foraging behavior of rodents could potentially affect the plant communities they live in because granivorous rodents affect seed distribution and plant invasion (Maron et al. 2012, Larios et al. 2017). Future studies at the CPER should focus on the effects of the rodent population on the local plant community assembly.

Literature Cited

- Orrock J.L., Danielson B.J., Brinkerhoff R.J. 2004. Rodent foraging is affected by indirect, but not by direct, cues of predation risk. *Behavioral Ecology* 15:433-437.
<https://doi.org/10.1093/beheco/arh031>
- Brown J.S., Kotler B.P., Smith R.J., Wirtz W.O. 1988. The effects of owl predation on the foraging behavior of heteromyid rodents. *Oecologia* 76:408–415.
- Rosenzweig, M.L. 1973. Habitat selection experiments with a pair of coexisting heteromyid rodent species. *Ecology* 54:111-117.
- Kotler B. P., Brown J., Mukherjee S., Berger-Tal O., Baouskila A. 2010. Moonlight avoidance in gerbils reveals a sophisticated interplay among time allocation, vigilance and state-dependent foraging. *Proceedings of the Royal Society B: Biological Sciences* 277:1469-1474.
- Kotler B. P. 1984. Risk of predation and the structure of desert rodent communities. *Ecology* 65:689-701.
- Clarke J. A. 1983. Moonlight's influence on predator/prey interactions between short-eared owls (*Asio flammeus*) and deer mice (*Peromyscus maniculatus*). *Behavioral Ecology and Sociobiology* 13:205-209.
- Lima S. L., Dill L.M. 1990. Behavioral decisions made under the risk of predation: a review and prospectus. *Canadian journal of zoology* 68:619-640.
- Stapp P., Lindquist M. D. 2007. Roadside foraging by kangaroo rats in a grazed short-grass prairie landscape. *Western North American Naturalist* 67:368–377.
- Longland W. S., Price M. V. 1991. Direct Observations of Owls and Heteromyid Rodents: Can Predation Risk Explain Microhabitat Use? *Ecology* 72:2261–2273.
<https://doi.org/10.2307/1941576>
- Grant W. E., Birney E.C., French N. R., and Swift D. M. 1982. Structure and Productivity of Grassland Small Mammal Communities Related to Grazing-Induced Changes in Vegetative Cover. *Journal of Mammalogy* 63:248–260.
- Ellis T. D., Cushman J.H. 2018. Indirect effects of a large mammalian herbivore on small mammal populations: Context-dependent variation across habitat types, mammal species, and seasons. *Ecology and evolution* 8:12115-12125.
- Brown J. H., Heske E. J. 1990. Control of a Desert-Grassland Transition by a Keystone Rodent Guild. *Science* 250:1705–1707. <http://www.jstor.org/stable/2878546>
- Maron J.L., Pearson D. E., Potter T., Ortega Y.K. 2012. Seed size and provenance mediate the joint effects of disturbance and seed predation on community assembly. *Journal of Ecology* 100:1492-1500. <https://doi.org/10.1111/j.1365-2745.2012.02027.x>
- Larios L., Pearson D. E., and Maron J.L. 2017. Incorporating the effects of generalist seed predators into plant community theory. *Functional Ecology* 31:1856-1867.
<https://doi.org/10.1111/1365-2435.12905>

- Lauenroth W. K., and Burke I.C. 2008. Ecology of the Shortgrass Steppe : A Long-Term Perspective. 17th ed. Oxford University Press.
- Barnett S. A. 1958. Experiments on 'neophobia' in wild and laboratory rats. British journal of psychology 49:195-201.
- Zimmerman G., Stapp P., Van Horne B. 1996. Seasonal variation in the diet of Great Horned Owls (*Bubo virginianus*) on shortgrass prairie. American Midland Naturalist 176:149-156.
- Prugh L. R., Golden C.D. 2014. Does moonlight increase predation risk? Meta-analysis reveals divergent responses of nocturnal mammals to lunar cycles. Journal of animal ecology 83:504-514.
- Hobbs N. T. 1996. Modification of Ecosystems by Ungulates. The Journal of Wildlife Management 60:695–713. <https://doi.org/10.2307/3802368>
- Eldridge D. J., Poore A.G., Ruiz-Colmenero M., Letnic M. and Soliveres S. 2016. Ecosystem structure, function, and composition in rangelands are negatively affected by livestock grazing. Ecological Applications 26:1273-1283.
- Thompson C.M., Gese E.M. 2013. Influence of vegetation structure on the small mammal community in a shortgrass prairie ecosystem. Acta Theriol 58:55–61.

The Purification of the Manganese Oxidizing Protein, MopA-hp, found in *Erythrobacter* sp. SD-21

Alyssa Ng

Advisor: Dr. Hope A. Johnson

Abstract

Microbes play an important role in the health and nutrient cycling of our planet. The transition metal manganese (Mn) found in terrestrial and marine environments can be oxidized from soluble Mn-II into insoluble Mn-III/IV by manganese oxidizing bacteria (MOB). The produced Mn-III/IV oxides are some of the most reactive compounds in nature and have the ability to oxidize many other elements. The marine bacterium *Erythrobacter* sp. SD-21 is known to produce a Mn-oxidizing protein, MopA (238-kDa). The purpose of this project is to obtain an active and pure sample of MopA-hp through different purification methods. The current protocol involves gravity nickel affinity chromatography (GNAC) and anion exchange chromatography (AEC). Currently, we are obtaining more AEC samples to test the activity and analyze the purity of the protein. A pure and active sample of MopA-hp will allow future studies on the mechanism of Mn oxidation by MopA-hp.

Introduction

The transition metal manganese (Mn) is an important element and is required for the activity of many enzymes involved in metabolic pathways, such as the oxidation of organic matter. Manganese in both terrestrial and marine environments can be oxidized from soluble Mn-II into insoluble Mn-III/IV by manganese oxidizing bacteria (MOB) (Tebo et al., 2005). The produced Mn-III/IV oxides are among the strongest oxidants in nature and have the ability to oxidize many other elements (Tebo et al., 2004). Mn-III/IV oxides play an important role in the fate and biogeochemical cycling of vital elements such as carbon, sulfur, and iron (Burdige, D. J., 1993)

Manganese is involved in the decomposition of organic matter in soil and water, which are two essential natural resources on Earth. Specifically on Earth, characteristics of Mn-III/IV oxides such as low solubility under natural environmental conditions and the ability to efficiently oxidize Fe and other metals (Tebo et al., 2018; Plathe et al., 2013). Mn (III/IV) oxides can also readily absorb pollutants such as lead, mercury, and arsenic due to their porous and highly reactive nature (Mensah et al., 2021) Coupled with Mn(II)-oxidizing microorganisms, biotic production of Mn(III/IV) oxides provides a natural solution for remediation of pollution sites. These oxides have not only been added directly to water and sediments for toxin removal but are generated by the environmental bacterial colonization of wastewater filters, a potentially cost saving means for safe water (Maurya et al., 2020). Despite the powerful geochemical reactivity of Mn and its widespread use by organisms, the mechanism of manganese oxidation by bacterial enzymes is unclear. Studying the natural mechanism of these enzymes has the potential to be utilized in the field of bioremediation.

The marine bacterium *Erythrobacter* sp. SD-21 is known to produce a Mn-oxidizing protein, MopA (238-kDa) – a peroxidase cyclooxygenase. Prior purification attempts of the active peroxidase domain, MopA-hp (105-kDa), have provided heterologously expressed samples with incomplete purification or no activity. Purifying a protein into a homogenous solution allows for further defined studies, however Mn(II) oxidizing enzymes are known to be difficult to purify. The purpose of this project is to obtain an active and pure sample of MopA-hp through different purification methods. This research focuses on the heterologous expression of MopA-hp in *Escherichia coli*, followed by exploring further purification of an MopA-hp using additional purification protocols. The current protocol involves gravity nickel affinity chromatography (GNAC) and anion exchange chromatography (AEC). Mn oxidation activity is quantified through a colorimetric leucoberberlin blue assay. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) is used to confirm the purity of the desired protein samples. Increased purification has been obtained through AEC, but a high amount of protein is lost throughout this process. A pure and active sample of MopA-hp will allow future studies on the mechanism of Mn oxidation by MopA-hp. Understanding the bacterial Mn oxidation mechanism has the potential to help us clean contaminated sites and better understand the biogeochemical cycles of Earth that keep the planet and its inhabitants healthy.

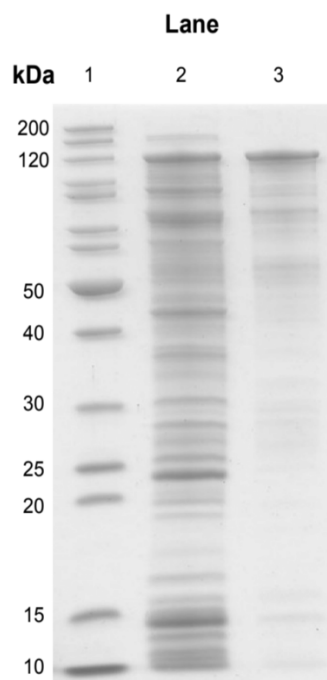


Figure 1. SDS-PAGE gel displays the current purification of MopA-hp. Note the contaminants present following affinity chromatography. Lane 1: Ladder, Lane 2: Cell free extract, Lane 3: MopA-hp purified by NAC located at 130kDa.¹

Methods

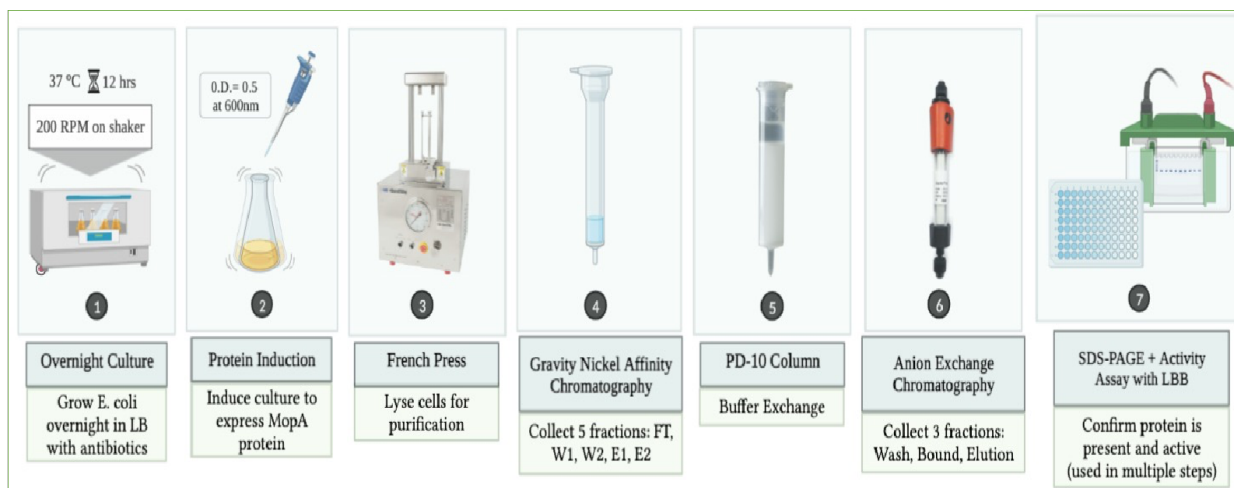


Figure 2. Protein purification methods displayed.

Heterologous Expression of MopA-hp

Escherichia coli strain Rosetta 2 cells (Novagen) transformed with pSpeedET-HP10 (MopA-hp) are streaked monthly on LB agar plates containing 60 $\mu\text{g mL}^{-1}$ kanamycin (KAN) and 30 $\mu\text{g mL}^{-1}$ chloramphenicol (CHL) for culture maintenance. A colony of this strain was inoculated into 25 mL of Luria broth (LB) with 60 $\mu\text{g mL}^{-1}$ kanamycin (KAN) and 30 $\mu\text{g mL}^{-1}$ chloramphenicol (CHL) and were incubated overnight at 37°C with constant shaking at 200 rpm. After approximately 12 hours, 2 mL of culture was aseptically transferred into 100 mL of LB media with the same concentration of KAN and CHL. Cells were incubated under the same conditions until an optical density of 0.5 was achieved at 600 nm ($\text{OD}_{600\text{nm}}$) measured with a Varian Cary 50 Bio UV-Visible Spectrophotometer (approximately 2 hours). At this point, protein expression was induced under the same conditions with the aseptic addition of 200 μL of 10% L-arabinose. After 4 hours, the cultures were harvested by centrifugation at 5000 x g for 15 minutes at 4°C using a Sorvall Legend X1R Centrifuge (Thermo Scientific). Cell pellets were re-suspended in 1.5 mL of EQ buffer (50 mM 4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid (HEPES) buffer pH 8, 50 mM NaCl, and 10 mM Imidazole). Re-suspended pellets were stored at -20°C for later use.

French Press: Acquisition of Cell Free Extract

Re-suspended cell pellets were thawed and lysed by 4-5 passages through a French press cell at 16000 psi. Cell lysate was centrifuged at 5000 x g for 15 minutes at 4°C. Cell-free extract (CFE) was collected and used for purification through Ni^{2+} affinity purification by gravity flow. CFE saved for later use was stored in the -20°C freezer.

Purification by Gravity Nickel Immobilized Metal Affinity Column Chromatography

Nickel affinity chromatography (NAC) was carried out using a gravity-flow column (Bio-Rad, Irvine, CA) and nickel nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen, Valencia, CA) at 4°C. This column was washed with 5mL of MQ water then 5 mL of EQ buffer.

Approximately 4 mL of CFE was loaded onto the column and incubated, horizontally and gently shaken for 15 mins at 4°C. After incubation, the resin was allowed to settle and unbound proteins are collected by flow through (FT). After collecting the FT fraction, protein was eluted with equilibrium (EQ) buffer at the following volumes with increasing imidazole concentrations: 10 mM (5 mL) wash 1 of unbound proteins, 20 mM (3 mL) wash 2, 40 mM (5 mL) elution 1, and 300 mM (5 mL) elution 2. After the collection of 5 fractions, the column was washed with 5 mL of Milli-Q (MQ) water, then 5 mL of 20% ethanol and stored in the 20°C deli fridge in 20% ethanol.

Protein Concentration Determination

Protein concentrations were measured on a Nanodrop One^c (Thermo Scientific) at an absorbance of 280 nm of 1 equal to 1 mg/mL⁻¹ of protein. The baseline correction was 340 nm. The EQ buffer was utilized as a blank.

Centrifugal Protein Concentration

The elution 1 (E1) fraction containing the highest protein concentration by absorbance at 280 nm (A₂₈₀) was retained for concentration by centrifugation in Amicon® Ultra-4 filter centrifuge tubes (Merck Millipore Ltd.) at 5,000×g to a volume of approximately 2.5 mL for the PD-10 column.

Buffer Exchange: Desalting Column

A GE PD-10 Desalting column with SephadexTM G-25 resin (GE Healthcare, Chicago, IL) was equilibrated with 25 mL of 20 mM HEPES pH 8, 100 mM NaCl, and 10% glycerol buffer (dialysis buffer). 2.5 mL concentrated protein was added then eluted from the column with 3.5 mL of the same equilibration buffer and collected as a 1 and 2.5 mL fraction. The 1 mL fraction was discarded, while the 2.5 mL protein fraction was quantified using the Nanodrop One^c. After obtaining concentration the protein was then assayed.

Anion Exchange Chromatography

Further protein purification included anion exchange chromatography using a HiTrapTM Q XL 1 mL (Cytiva) column. This technique was conducted by hand. First the column was washed to remove preservatives with a syringe containing 5 mL of binding buffer (BB) [50 mM Tris-HCl pH 8 and 20% glycerol] followed by 5 mL of elution buffer (EB) [50 mM Tris-HCl pH 8, 20% glycerol, and 1 M NaCl]. Both buffers were filtered through a 0.45 µm syringe filter. 5 mL of BB was utilized for column equilibration before passing the MopA-hp containing sample through the column. The MopA-hp-containing sample was filtered through a 0.22 µm syringe filter before placed in column. The column was then washed with 5 mL of BB to remove unbound proteins. Followed by a 5 mL wash of EB for analysis.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Non-native sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS - PAGE) was performed at a constant 200 V for approximately 1.5 h using 12% acrylamide separating and 4% acrylamide stacking gels. The 2X stock solution (0.5 M TRIS pH 8.6, 10% SDS, 10% glycerol, 0.05% bromophenol blue) mixed with 2-mercaptoethanol (1:20 dilution 2-mercaptoethanol) was utilized in a 1:1 ratio with the protein sample and heated for 5 mins. Protein ladder Broad Range from Bio - Rad Laboratories (Irvine, CA) or Thermo Scientific PageRuler Plus pre - stained protein ladder (Waltham, MA) were used as molecular weight (MW) standards. Proteins were stained in a solution of 0.05% Coomassie brilliant blue (Bio - Rad Laboratories), 10% acetic acid, and 40% ethanol. The gels were destained in a solution of 10% acetic acid and 10% ethanol.

Mn(II) Oxidation Activity: Leucoberbelin Blue Assay

Samples containing MopA-hp were mixed in a ratio of 1:5 (500 μ L final volume) with a mixed solution containing 5 μ L of 0.1 M $MnCl_2$, 5 μ L of 1.0 M $CaCl_2$, 1 μ L of 5 mM PQQ, 383 μ L of 50 mM HEPES pH 8.0 and 100 mM NaCl, 5 μ L of 10 mM NAD^+ , and 1 μ L of 1 mM heme, for activity testing. The EQ buffer utilized before served as the negative control. Triplicates were conducted for each sample. After 24hrs of shaking at 200 rpm at room temperature, 50 μ L of the assay mixture was mixed in a ratio of 1:5 (300 μ L final volume) with 0.04% LBB in 45 mM acetic acid. After 15 minutes of shaking at 200 rpm at room temperature, these samples were centrifuged in an Eppendorf Centrifuge 5430 R for 5 minutes at 14,000 rpm at room temperature. The supernatant (250 μ L) was placed on a 96 well plate to read their absorbance at 620 nm on a Synergy 2 (Biotek) plate reader. The absorbance readings were

compared to a standard curve produced with KMnO₄ standards where 1 μ M KMnO₄ = 2.5 μ M MnO₂.

Results

SDS-PAGE and the activity assay provide a comprehensive assessment of the success of the current MopA-hp protein purification protocol. SDS-PAGE was used to visualize the 130 kDa MopA-hp band and contaminants. The LBB colorimetric assay was used to measure Mn oxidizing activity. The combination of these techniques ensure that MopA-hp is present and active in the samples.

5 fractions are collected from the GNAC protocol. Elution 1 (E1) fraction is used for further purification as it is the most abundant and pure compared to the other fraction samples visualized by SDS-PAGE (Figure 2). The LBB activity assay was conducted and indicated that E1 had the highest Mn oxidizing activity (Table 1). This suggests that E1 has the highest concentration of active MopA-hp and is the most promising fraction for further purification.

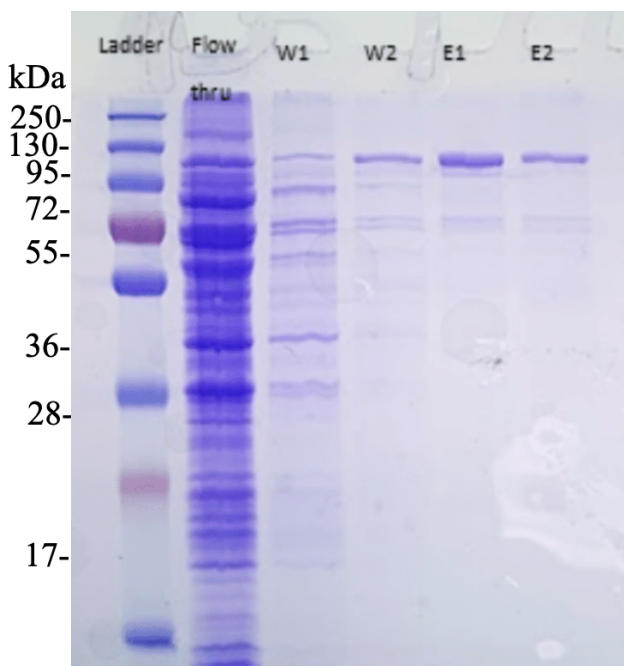


Figure 3. SDS-PAGE gel displays 5 fractions collected from GNAC. Elution 1 has the highest protein abundance and the least contaminants. Lane 1: Ladder, Lane 2: Flow through, Lane 3: Wash 1, Lane 4: Wash 2, Lane 5: Elution 1, Lane 6: Elution 2.

Specific Activity	A620/mg
CFE	0.796
Flow through	0.244
Wash 1	4.21
Wash 2	7.52
Elution 1	10.87
Elution 2	10.16

Table 1. Manganese oxidizing specific activity table displays Elution 1 (E1) has the highest activity of all the fractions. This fraction also had the most abundant protein band when visualized by SDS-PAGE (Figure 2).

Fraction	A620	GNAC Vol. (mL)	Total Activity (A*mL)	Activity Recovered (%)
CFE	1.26	4	5.04	100
FT	0.221	4.5	0.9945	19.73214286
W1	0.122	5.5	0.671	13.31349206
W2	0.124	3	0.372	7.380952381
E1	0.125	5.5	0.6875	13.64087302
E2	0.127	5	0.635	12.59920635

Table 2. Mn Oxidizing Activity Assay results indicate E1 has the highest activity recovered of the purified samples.

Three to five fractions of E1 are collected and concentrated using Amicon® Ultra-4 filter centrifuge tubes (Merck Millipore Ltd.) to about 2.5 mL. E1 is then desalted through the buffer exchange protocol using the PD-10 column. SDS-PAGE is used to analyze the resulting bands of the concentrated protein sample (Figure 3). The gel indicates that there are more contaminants in the protein sample than visible with the initial analysis.

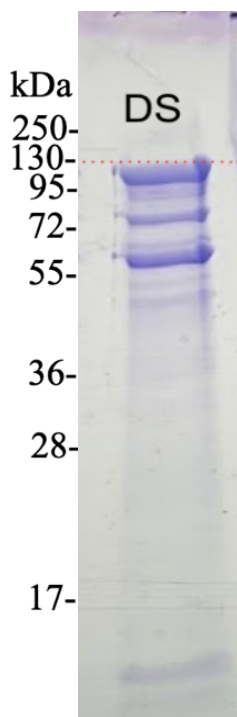


Figure 4. SDS-PAGE gel displays desalting fraction from PD-10 Column. Three bands result from the concentration of the E1 fractions.

After the buffer exchange, the anion exchange chromatography (AEC) purification step was performed next. Three fractions are collected which are the wash, bound, and elution. The elution fraction contained three bands at the same size as the starting fraction, which suggests that MopA-hp bound to the anion exchange column (Figure 4). However, the abundance of the protein is much less than the initial sample put onto the column. This indicates that some of the protein is stuck in the column or is not eluted all the way through. We have not yet obtained a sample from AEC that produces high Mn oxidizing activity. Next, we are going to combine the fractions to see if this will recover activity.

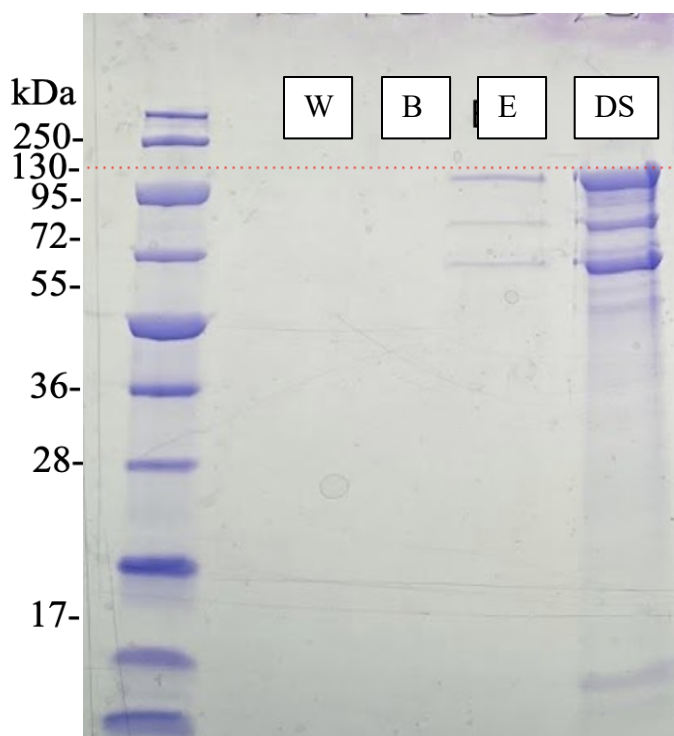


Figure 4. MopA-hp is located at 130kDa. Elution fraction contains three bands. Lane 1: Ladder, Lane 2:Wash, Lane 3: Bound, Lane 4: Elution, Lane 5: Starting Sample (DS)

Discussion

Prior discoveries and findings have helped the scientific community gain a better understanding about the manganese oxidation process, but a full understanding of the biochemical mechanism of MopA requires a pure and active protein. Current research has either provided a pure but inactive MopA-hp, or active but impure MopA-hp. This research focuses on the heterologous expression of MopA-hp in *Escherichia coli*, followed by utilizing several purification techniques to explore further purification. Mn oxides affect different biological environmental processes that require such elements in compounds and contribute to the lowering levels of toxic chemicals in the environment (Tebo et al., 2004). A final characterization of a

pure sample MopA-hp and the compounds that allow for its activity, would allow for kinetics studies of MopA-hp and thus have a better understanding of MopA-hp and its mechanism implied in Mn oxidation.

The Purification of MopA

The purpose of this project is to obtain an active and pure sample of MopA-hp after the heterologous expression in *E. coli*. To obtain a pure sample of MopA-hp, GNAC and AEC were used in attempts to purify the cell-free extract. SDS-PAGE was used to analyze the purity after each chromatography. GNAC showed that E1 had the highest protein abundance and least amount of contaminating proteins (Figure 2). This sample was concentrated and buffer exchange through the PD-10 column was done to desalt the protein solution for AEC. The PD-10 column is used for buffer exchange, AEC is done after the imidazole is removed because the salt interferes with the protein purification protocol. The fractions from AEC were visualized via SDS-PAGE. The elution fraction showed 3 bands that were very faint compared to the starting sample. These bands should be about $\frac{1}{2}$ the intensity of the desalting sample based on the volumes recovered and added to the gel, but they are only about $\frac{1}{10^{\text{th}}}$ of the intensity. This indicates that some sample is not eluted all the way by the buffer or that the protein is getting stuck in the column. More AEC samples need to be obtained to further investigate the reason why the bands are significantly less intense than the starting sample.

Protein Mn Oxidizing Activity

Mn oxidizing activity was determined in the protein samples. From the GNAC, Elution 1 had the highest activity compared to the other fractions (Table 2). We have not yet obtained a

sample from AEC that produces high Mn oxidizing activity. Next, we are going to combine the fractions from the anion exchange to see if this will recover activity.

Future works

Future works could focus on combining AEC fractions to recover activity and further analyzing the contaminants in the SDS-PAGE gels to see which bands are necessary for Mn oxidation. The characterization of a pure MopA-hp sample and its necessary compounds that conduct activity allow for kinetic studies of MopA-hp. These studies allow us to learn about the MopA-hp mechanism of Mn oxidation.

Acknowledgements

I would like to acknowledge the indispensable contributions of all those who have supported me during my time on this research project. I would first like to thank my principal investigator, Dr. Hope A. Johnson, who supported me throughout each step and guided me to be a better scientist and researcher. I would also like to thank my lab mentor, Bryan R., for his guidance, training, and mentorship. He has provided me with invaluable knowledge and taught me what it truly means to be a researcher. Thank you to all the members of our lab for lending a helping hand and supporting me throughout my research journey. I would like to recognize the unconditional moral support from family members, mentors, and close ones, as well as the different programs where I was involved: Research Careers and Preparatory Program (RCP), CSU Louis Stokes Alliance for Minority Participation (LSAMP), and Summer Undergraduate Research Academy (SUREA).

I received financial support for laboratory supplies from the CSU Fullerton Department of Biological Science, Associated Students Inc. Research Grant, and the National Science Foundation (NSF). I also received financial support for conference registration and travel from the Natural Sciences and Mathematics Inter-Club Council.

Literature Cited

- Burdige, D. J. (1993). The biogeochemistry of manganese and iron reduction in marine sediments. *Earth-Science Reviews*, 35(3), 249–284. [https://doi.org/10.1016/0012-8252\(93\)90040-e](https://doi.org/10.1016/0012-8252(93)90040-e)
- Dick, G. J., Podell, S., Johnson, H. A., Rivera-Espinoza, Y., Bernier-Latmani, R., McCarthy, J. K., Torpey, J. W., Clement, B. G., Gaasterland, T., & Tebo, B. M. (2008). Genomic Insights into mn(ii) oxidation by the marine alphaproteobacterium *aurantimonas* SP.. strain SI85-9A1. *Applied and Environmental Microbiology*, 74(9), 2646–2658. <https://doi.org/10.1128/aem.01656-07>
- Dinh, D. M. (2017). *Impurity Analysis And Characterization Of A Purified Recombinant Manganese-Oxidizing Protein From A Marine Bacterium* (thesis).
- Li Qian , K. (2020). *Function Of Heme-Binding In And Native Host Expression Of A Bacterial Manganese Oxidase* (thesis).
- Maurya, A., Singh, M. K., & Kumar, S. (2020). Biofiltration technique for removal of waterborne pathogens. *Waterborne Pathogens*, 123–141. <https://doi.org/10.1016/b978-0-12-818783-8.00007-4>
- Medina, M., Rizo, A., Dinh, D., Chau, B., Omidvar, M., Juarez, A., Ngo, J., & Johnson, H. A. (2018). MopA, the Mn oxidizing protein from *Erythrobacter* sp. SD-21, requires heme and NAD⁺ for Mn(II) oxidation. *Frontiers in Microbiology*, 9. <https://doi.org/10.3389/fmicb.2018.02671>
- Mensah, M. B., Lewis, D. J., Boadi, N. O., & Awudza, J. A. M. (2021). Heavy metal pollution and the role of inorganic nanomaterials in environmental remediation. *Royal Society open science*, 8(10), 201485. <https://doi.org/10.1098/rsos.201485>
- Nakama, K., Medina, M., Lien, A., Ruggieri, J., Collins, K., & Johnson, H. A. (2014). Heterologous expression and characterization of the manganese-oxidizing protein from *erythrobacter* sp.. strain SD21. *Applied and Environmental Microbiology*, 80(21), 6837–6842. <https://doi.org/10.1128/aem.01873-14>
- Plathe, K. L., Lee, S. W., Tebo, B. M., Bargar, J. R., & Bernier-Latmani, R. (2013). Impact of microbial Mn oxidation on the remobilization of bioreduced U (IV). *Environmental science & technology*, 47(8), 3606-3613.
- Peres, T. V., Schettinger, M. R., Chen, P., Carvalho, F., Avila, D. S., Bowman, A. B., & Aschner, M. (2016). "Manganese-induced neurotoxicity: a review of its behavioral consequences and neuroprotective strategies". *BMC pharmacology & toxicology*, 17(1), 57. <https://doi.org/10.1186/s40360-016-0099-0>
- Rizo , A. (2018). *Abiotic Interactions And Redox Analysis In Manganese Oxidation By Mopa-Hp* (thesis).

Mathematics

Combinatorial Identities Derived from the Monoid of Partial Order-preserving Injections

Tyler Kim
Advisor: Scott Annin

Abstract

The study of monoids, algebraic structures with an associative binary operation and an identity element, contributes to diverse disciplines: computer science, physics, mathematics, and others. It also reveals interesting and practical revelations on combinatorial identities. We analyze the monoid of partial order-preserving injections, or $POI(n)$, and we reveal new representations of the elements of $POI(n)$. Our work leads to combinatorial identities involving the Catalan Numbers, and in particular, we explore a proof of Touchard's Identity that relates to a sub-monoid of $POI(n)$. We also uncover a novel combinatorial identity involving the cardinality of $POI(n)$.

1 Introduction

In recent years, there has been interest in the categorization of roots of elements across various mathematical structures [1, 3, 4]. Useful applications in fields such as matrix theory and cryptography are well-known. Analyzing roots of elements in the monoid of partial order-preserving injections of an n -element set, $POI(n)$, has motivated new notational representations of the elements of $POI(n)$. In turn, these new notational devices have led to some noteworthy combinatorial insights. By counting the same objects from different perspectives, we can derive several combinatorial identities. These identities involve the sequences of Catalan numbers, C_n , and Motzkin numbers, M_n , which we will review in Section 3 below. The identities are summarized here. The first two are already known, while the third one is a new result of our research.

Touchard's Identity [12] :

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} C_k \binom{n}{2k} 2^{n-2k} = C_{n+1}$$

Relation of Motzkin to Catalan Numbers [8]:

$$M_n = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} C_k$$

Cardinality of $POI(n)$:

$$\binom{2n}{n} = 2^n + 2 \sum_{v=1}^n \sum_{k=1}^v \binom{v-1}{2k-1} C_{k-1} (2^{v-2k}) \binom{2n-2v}{n-v}$$

2 Background

2.1 Introduction to $POI(n)$

Definition The **symmetric inverse monoid** on an n -element set $S = \{1, 2, \dots, n\}$, denoted by $SIM(n)$, consists of all partial one-to-one mappings from the set to itself [5]. The elements of $SIM(n)$ are represented by a two-line notation in which $i \in S$ is placed on the top line directly above its image $\sigma(i)$. For example, consider:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & 1 & 5 & 2 & 3 & - & 7 \end{pmatrix} \in SIM(7).$$

In σ , we have that 2 maps to 1, 3 maps to 5, 4 maps to 2, 5 maps to 3, 7 is fixed, and both 1 and 6 do not get mapped to anything. The binary operation in $SIM(n)$ is composition and, as is custom, is done from right to left. The following calculation illustrates this.

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

Definition The monoid of **partial order-preserving injections**, denoted by $POI(n)$, is a submonoid of $SIM(n)$ consisting of all σ in $SIM(n)$ such that $1 \leq i < j \leq n$ implies $\sigma(i) < \sigma(j)$ [1]. Here is an example:

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

Note that if $\sigma \in SIM(n)$, then $\sigma \in POI(n)$ if and only if the elements appearing in the second row of the two-line notation are in increasing order.

Definition For an element a in a semigroup S , suppose that there exists an element, a' in S , that satisfies $a = aa'a$ and $a' = a'aa'$. Then we call a' an **inverse** of a .

Definition We call a monoid M an **inverse monoid** if for every element $a \in M$, there exists a unique inverse $a' \in M$.

Remark Both $SIM(n)$ and $POI(n)$ are inverse monoids.

Definition The **domain** of an element $\sigma \in POI(n)$ is denoted as

$$Dom(\sigma) = \{i \in S : \sigma(i) \text{ is defined}\}.$$

Definition The **range** of an element $\sigma \in POI(n)$ is denoted as

$$Rng(\sigma) = \{\sigma(i) : i \in S\}.$$

Definition The **rank** of an element $\sigma \in POI(n)$ is defined as

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

Definition Any $\sigma \in POI(n)$ such that $\sigma(i) = i$ for all $i \in Dom(\sigma)$ is a **partial identity**.

Definition For any $\sigma \in POI(n)$, the **inverse** of σ is defined as $\sigma^{-1} \in POI(n)$ such that

$$Dom(\sigma^{-1}) = Rng(\sigma) \quad \text{and} \quad Rng(\sigma^{-1}) = Dom(\sigma).$$

Since an element $\sigma \in POI(n)$ is completely determined once its domain and range are specified, the definition of the inverse is well-defined. Additionally, we can verify that this agrees with the general definition of inverses in a semigroup given above by verifying that $\sigma\sigma^{-1}\sigma = \sigma$ and $\sigma^{-1}\sigma\sigma^{-1} = \sigma^{-1}$. Also, note that both $\sigma\sigma^{-1}$ and $\sigma^{-1}\sigma$ are partial identities.

Example 1 Consider the element $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 7 & - & - & - & - \end{pmatrix} \in POI(7)$. Then $Dom(\sigma) = \{1, 2, 3\}$ and $Rng(\sigma) = \{3, 4, 7\}$. Also, $rank(\sigma) = 3$. Finally, $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & - & 1 & 2 & - & - & 3 \end{pmatrix}$; then $Dom(\sigma^{-1}) = \{3, 4, 7\}$, and $Rng(\sigma^{-1}) = \{1, 2, 3\}$.

Definition Let us define $POI_{asc}(n)$ to be the sub-monoid of $POI(n)$ consisting of all σ in $POI(n)$ such that $\sigma(i) \geq i$ for all $i \in Dom(\sigma)$. Likewise, let us define $POI_{des}(n)$ to be the sub-monoid of $POI(n)$ consisting of all σ in $POI(n)$ such that $\sigma(i) \leq i$ for all $i \in Dom(\sigma)$.

Remark The cardinality of $POI_{asc}(n)$ equals the cardinality of $POI_{des}(n)$ since we can define a bijection $\theta : POI_{asc}(n) \rightarrow POI_{des}(n)$ via $\theta(\sigma) = \sigma^{-1}$. In Example 1, note that $\sigma \in POI_{asc}(7)$ while $\sigma^{-1} \in POI_{des}(7)$.

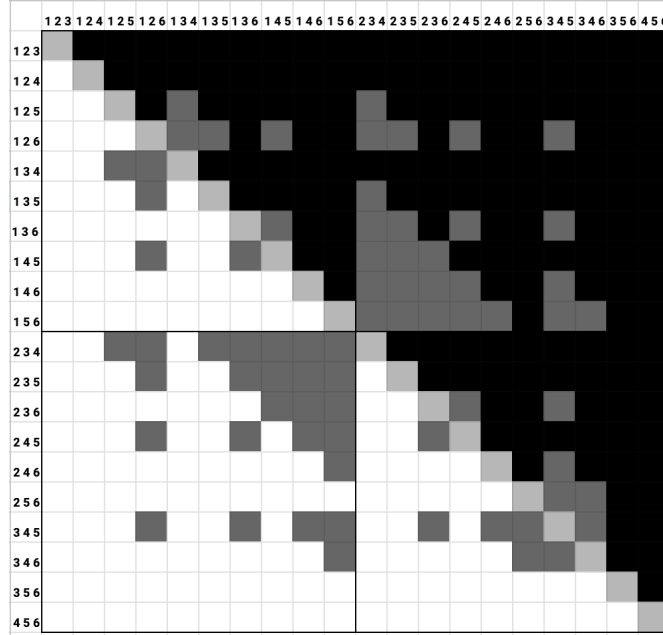


Fig. 1. A graphical representation of $POI(6)$, Rank 3. The columns represent the domain, and the rows represent the range. White squares correspond to the elements in $POI_{asc}(n)$, black squares correspond to the elements of $POI_{des}(n)$, dark gray squares correspond to elements that are in neither, and light gray corresponds to elements that are in both. Elements of $POI_{des}(n)$ are generally found in the upper-right of the grid, and elements of $POI_{asc}(n)$ are in the bottom-left. This diagram thus demonstrates the bijection between elements in $POI_{asc}(n)$ and $POI_{des}(n)$ obtained via $\sigma \rightarrow \sigma^{-1}$, which is visualized by a reflection across the main diagonal of the grid.

Remark Note that the union of $POI_{asc}(n)$ and $POI_{des}(n)$ is not $POI(n)$. For instance, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & - & - & - & 4 & 5 \end{pmatrix} \in POI(6)$ is in neither $POI_{asc}(6)$ nor $POI_{des}(6)$. Indeed, the dark gray squares in Fig. 1 identify elements of $POI(6)$ that do not belong to either subset. Also note that the intersection of $POI_{asc}(n)$ and $POI_{des}(n)$ is non-empty; all light gray squares correspond to partial identities, which are in both sets.

2.2 Path Notation

We next present a series of definitions that relate to the research we present in this paper.

Definition Let $c \in \{1, 2, 3, \dots, n\}$, and let σ be in $POI(n)$. Then if c is neither in the domain nor range of σ , we call c a **gap** of σ .

Definition Let $\sigma \in \text{POI}(n)$. If $\sigma(c) = c$, we call c a **fixed point** of σ .

Definition: Let $\sigma = (a_1 a_2 \dots a_t] \in \text{POI}(n)$ denote the partial one-to-one mapping that sends a_1 to a_2 , a_2 to $a_3, \dots, a_{t-1}$ to a_t , and does not map a_t to anything. We call $(a_1 a_2 \dots a_t]$ a **t-path**, or simply a **path**.

Definition: Let $\sigma = (a_1) = \begin{pmatrix} 1 & 2 & \dots & a_1 & \dots & n \\ - & - & - & a_1 & - & - \end{pmatrix}$ denote the partial one-to-one mapping that sends a_1 to a_1 . In other words, a_1 is a fixed point.

Definition We call the **path notation** of an element of $\text{POI}(n)$ a product of all such disjoint paths and fixed points. Such a product is uniquely determined for each element up to the order in which the paths appear in the product.

Remark Suppose that in the path notation of $\sigma \in \text{POI}(n)$ we have two paths, $a = (a_1 a_2 \dots a_t]$ and $b = (b_1 b_2 \dots b_s]$, such that $a_1 < b_1$. Then for all $r \leq \min\{s, t\}$, the ordering principle of $\text{POI}(n)$ implies that $a_r < b_r$. We then write path a before path b in the notation for consistency.

Example 2 Here is the two-line notation of an element of $\text{POI}(4)$ along with the corresponding path notation:

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

We see that 1 maps to 3, which does not have an image. Thus, in the path notation we write the path $(13]$. To represent the other path, we note that 2 maps to 4, which does not have an image. Note that we write the paths in this order according to their first values, as explained in the remark above.

Example 3 Here we present an element of $\text{POI}(6)$ in both notations:

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

Let us note two new details. First, 2 is part of a 1-path; in other words, 2 is a gap: $(2]$. However, 6 is a fixed point, which, by the convention established above, is written as (6) .

Example 4 The example below demonstrates paths of greater lengths:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & - & 5 & 6 & 8 & 9 & - & - & - & 10 \end{pmatrix} = (1\ 3\ 5\ 8](2](4\ 6\ 9](7)(10).$$

2.3 Initiate-Continue-End (ICE) Notation

Definition For $\sigma \in POI(n)$, we code σ into a string of length n , where each character in the string corresponds to $1, 2, 3, 4, \dots, n$ respectively, to form the **ICE notation** for σ . The symbols of the ICE notation are as follows:

- I**: The corresponding number is the first number of a path, initiating it;
- F**: The corresponding number is a fixed point;
- 1**: The corresponding number is a gap;
- C**: The corresponding number is in the middle of a path;
- E**: The corresponding number is the last number of a path, ending it.

Remark For convenience and utility, we will apply the ICE notation for elements in $POI_{asc}(n)$ only. However, a more generalized form can technically be used for all elements in $POI(n)$.

Definition A **slot** in the ICE notation is merely a space that a character of the string occupies. We number the slots from left to right as $1, 2, \dots, n$. For instance, slot 1 would be the space the first character occupies.

2.4 Initiate-Continue-End (ICE) Algorithm

Throughout this sub-section, σ will always refer to an element of $POI_{asc}(n)$. An **ICE algorithm** is used to translate from the ICE notation to the path notation, and reversing it translates from the path notation to the ICE notation. To explain the algorithm, we need a couple of definitions first.

Definition An **open path** is a path that is in the process of being constructed using the ICE algorithm; it is not finished. Given an open path, we must continue it by adding at least one more element to the path.

Definition An open path has a **right-most value**, which is the last number that has been added to it in the process of constructing an element of $POI(n)$ using the ICE algorithm. The right-most value of an open path changes throughout the execution of the ICE algorithm.

ICE Algorithm: Now, let us read the ICE notation from left to right. If we reach an **I** in slot k , we begin a new path with k . If we reach an **F** in slot k , we map k to itself and add (k) to the path notation. If we reach a **1** in slot k , we place k in a 1-path: $(k]$. If we reach an **E** in slot k , we place k as the last number in the open path that has the least right-most value. Finally, if we reach a **C** in slot k , we place k in the open path that has the least right-most value, and the path remains open.

Remark Each action represented by **I**, **F**, and **1** is clearly well-defined. The actions represented by **E** and **C** are also well-defined since only one open path

can have the least right-most value. Thus, every ICE notation has only one corresponding path notation, and vice versa.

Remark All elements of $POI_{asc}(n)$ that can be represented with the path notation can also be represented by the ICE notation since every element is constructed of only paths, fixed points, and gaps.

Definition We denote by $P(r)$ the number of open paths at slot r , the number of Is appearing in slots 1 through r minus the number of Es appearing in slots 1 through r . Clearly, $P(r) \geq 0$ for all r . An example is illustrated in Fig. 2.

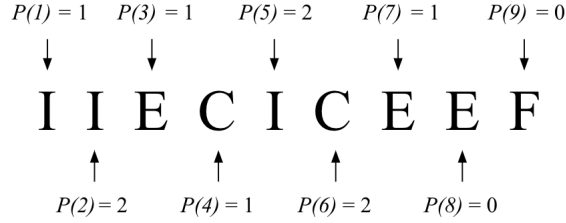


Fig. 2. An example of computing the number of open paths at each slot $r = 1$ through 9.

Example 5 Let us use the ICE notation in Fig. 2 to demonstrate a part of the ICE algorithm. After the first two Is, we have two open paths beginning with 1 and 2. Then, we place 3 in the path containing 1, since 1 is the least right-most value. We now only have one open path. Then, 4 is placed in the path containing 2, which is the only option. We continue this process to yield the path notation: $(1\ 3](2\ 4\ 6\ 8](5\ 7](9)$.

Remark We establish conditions for the slot in which each character can be placed in order for an ICE string to represent an element of $POI_{asc}(n)$:

- 1) $P(n) = 0$ for all elements in $POI_{asc}(n)$ since all paths that are initiated must eventually be closed.
- 2) 1 can be placed in any slot since the corresponding number is neither the preimage or image to any other number; it does not disrupt the ordering principle.
- 3) F can be placed in slot r only if $P(r) = 0$. **Proof:** Let us proceed by way of contradiction. Suppose that $P(r) > 0$, and we place an F in slot r . Then there exists at least one open path which contains $c < r$ as its rightmost value. This path must be continued or ended with a number greater than r . Thus, $\sigma(c) > r = \sigma(r)$, contradicting the ordering principle of $POI(n)$.

4) C can be placed in slot r only if $P(r) > 0$. **Proof:** If $P(r) = 0$, there are no open paths when slot r is reached, so it is impossible to continue any paths.

5) I can be placed in slot r only if $P(r-1) < n-r$. **Proof:** Initiating a path does not disrupt ordering. However, we note that $|P(i+1) - P(i)| \leq 1$ for all $i = 1, 2, \dots, n-1$ since the number of open paths at two consecutive slots can only differ by at most 1. It then follows that, more generally, $|P(i+j) - P(i)| \leq j$ for all $i = 1, 2, \dots, n-j$. Further, given that $P(n) = 0$, it also follows that $P(n-1) \leq 1, P(n-2) \leq 2, \dots, P(n-j) \leq j$ for all $j = 0, 1, 2, \dots, n-1$. Our result follows after we set j to $n-r$.

6) E can be placed in slot r only if $P(r) > 0$. **Proof:** One cannot end a path at slot r unless there is at least one open path at slot r .

Note that these conditions are relatively simple compared to the path notation; this fact is useful when approaching the combinatorics of $POI(n)$.

Example 6 Returning to Example 2, we can now also present the ICE notation:

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

Note that because the ICE notation begins with II , 1 and 2 are the first numbers of the paths that they belong to. Further, since the notation ends with EE , 3 and 4 have a preimage, but not an image. Note that, in the process of applying the ICE algorithm for slot 3, we must place 3 in an open path. There are two open paths at slot 3, with right-most values 1 and 2. According to the algorithm, we place 3 in the path with 1. Indeed, placing 3 in the path with 2 would result in an element that violates the order-preserving requirement of $POI(n)$.

Example 7 Here is an example in $POI(6)$ that builds from Example 3.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & - & 5 & - & - & 6 \end{pmatrix} = (1\ 4](2](3\ 5](6) = I1IEEF$$

We know that 2 is a gap. Thus, the second symbol in the ICE notation is a 1. Further, since 6 is a fixed point, the last symbol in the ICE notation is an F .

Example 8 We analyze an element containing all of the symbols 1, I , C , E , and F , which we first introduced in Example 4.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & - & 5 & 6 & 8 & 9 & - & - & - & 10 \end{pmatrix} = (1\ 3\ 5\ 8](2](4\ 6\ 9](7](10) \\ = I1CICC1EEF$$

We will now explain the construction of the ICE notation given a path notation. We note that 1 is the first number in its corresponding path, so the first symbol in the ICE notation is I . Then, 2 is a 1-path, so the second symbol is 1. Since 3 is in the same path as 1, and is thus in the middle of a path, the third symbol is a C . We continue this process for the remaining numbers.

3 Combinatorics

Now that we have established useful notations for elements of $POI(n)$, we will consider how they may be used to prove some combinatorial identities. These identities involve a famous sequence of numbers, the Catalan numbers.

3.1 Touchard's Identity

We count the number of $\sigma \in POI(n)$ such that $\sigma(c) \geq c$ for all $c \in Dom(\sigma)$, which is $|POI_{asc}(n)|$. Thus, we are counting all possible ICE strings of length n .

3.1.1 Catalan Numbers

Catalan numbers are familiar in combinatorics. For instance, the n th Catalan number is the number of non-crossing partitions of an n -element set [13]. However, the definition we use involves Catalan graphs.

Definition The n th Catalan number, denoted by C_n , is the number of routes on the upper right quadrant of the xy -coordinate plane with $x, y \in \mathbb{Z}$ from coordinate $(0, 0)$ to coordinate $(2n, 0)$ in $2n$ steps if one is allowed to move only to the right (up or down) at each step but forbidden from dipping below the x -axis. These form **Catalan graphs** [13]. By convention, we set $C_0 = 1$.

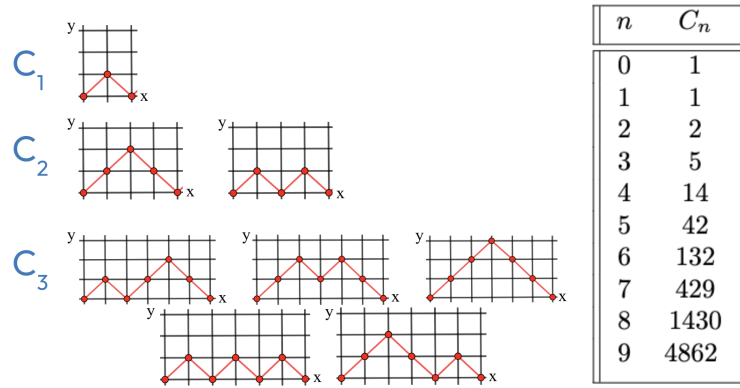


Fig. 3. Catalan graphs. Note that the graphs never drop below the x -axis.

Remark For point (x, y) on any Catalan graph, x is analogous to r , and y is analogous to $P(r)$ of the ICE notation.

Remark The conditions of $POI(n)$ state that $P(r) \geq 0$ for all r , meaning that finding all legal strings of k Is and k Es according to the ICE notation rules in the previous section is the same as finding the number of Catalan graphs of length $2k$, which is C_k .

Finally, a closed form of the n th Catalan number exists [13]:

$$C_n = \frac{1}{n+1} \binom{2n}{n}.$$

There are also a number of notable recursive identities for the Catalan numbers, such as [7, 16]:

$$C_{n+1} = \frac{2(2n+1)}{n+2} C_n$$

$$C_{n+1} = C_0 C_n + C_1 C_{n-1} + \dots + C_{n-1} C_1 + C_n C_0$$

3.1.2 Counting using Catalan Numbers

Let us count all possible ICE notations with n slots. To begin, assume the ICE notation contains exactly k Is and k Es. Given the remark above, we arrange the Is and Es in C_k ways. Then, we must place the arranged Is and Es into n slots as our second combinatorial task, with $\binom{n}{2k}$ possibilities.

Now let us define an **alternating variable**, A , where A can either be F or C , depending on the mutually exclusive situations (since an F can only be placed at slot r if $P(r) = 0$, and a C can be placed at slot r if $P(r) > 0$). Then in each remaining slot we either place an A or a 1, leading to 2^{n-2k} possibilities for the remaining $n - 2k$ slots. So the number of ICE notations with exactly k Is and k Es is $C_k \binom{n}{2k} 2^{n-2k}$. Iterating through all possible values of k , we obtain all possible ICE notations.

Theorem 1 For any $n \geq 0$,

$$|POI_{asc}(n)| = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} C_k \binom{n}{2k} 2^{n-2k}.$$

3.1.3 Motzkin Numbers

Motzkin numbers have multiple combinatorial definitions. For instance, the n th Motzkin number, M_n , is the number of different ways of drawing non-intersecting chords between n points on a circle (not necessarily touching every point by a chord) [14].

Definition The n th **Motzkin number** is the number of routes on the upper right quadrant of the xy -coordinate plane with $x, y \in \mathbb{Z}$ from coordinate $(0, 0)$ to coordinate $(n, 0)$ in n steps if one is allowed to move only to the right (up, down, or straight) at each step but forbidden from dipping below the x -axis. These form **Motzkin graphs** [15]. By convention, we set $M_0 = 1$.

Remark Since Motzkin graphs have more allowed configurations than Catalan graphs, it is clear that $M_{2n} \geq C_n$ for all n .

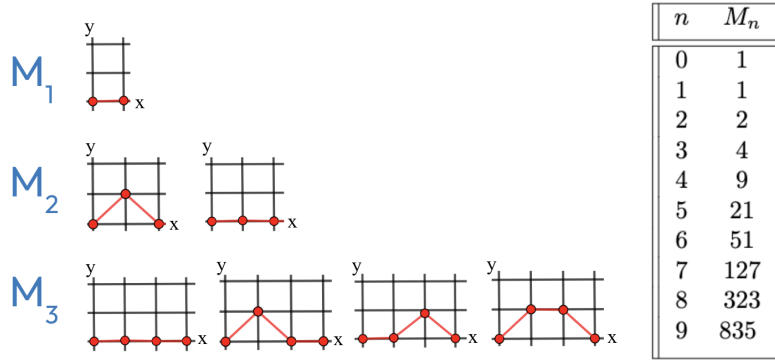


Fig. 4. The Motzkin graphs corresponding to M_1 , M_2 , and M_3 , and the first ten Motzkin numbers. We can see similarities to Catalan graphs, except for the addition of horizontal lines.

Recursive identities for the Motzkin numbers exist, for instance [15]:

$$(n + 3)M_{n+1} = (2n + 3)M_n + (3n)M_{n-1}.$$

$$M_{n+1} = M_n + M_0M_{n-1} + M_1M_{n-2} + \dots + M_{n-2}M_1 + M_{n-1}M_0$$

It is interesting to note that the latter recursion follows a similar pattern to a recursion for the Catalan numbers from Section 3.1.1.

3.1.4 Counting using Motzkin Numbers

Because a 1 in the ICE notation does not affect the number of paths that are open, it is analogous to moving straight, or horizontally, in a Motzkin graph. Thus, the number of ways to arrange a total of k I s, 1s, and E s is M_k .

Let us now assign combinatorial tasks to construct an element of $POI_{asc}(n)$. If we have a total of k 1s, I s and E s, we can arrange them in M_k ways. Then we can choose k out of n slots to place the 1s, I s, and E s. The remaining $n - k$ slots must all be A s, where A is the alternating variable as defined in Section 3.1.2. We must sum through all possible values of k , yielding the following theorem.

Theorem 2 For any $n \geq 0$,

$$|POI_{asc}(n)| = \sum_{k=0}^n M_k \binom{n}{k}.$$

In view of Theorems 1 and 2 above, we obtain

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} C_k \binom{n}{2k} 2^{n-2k} = \sum_{k=0}^n M_k \binom{n}{k}.$$

Finally, using the well-known identity [15]

$$\sum_{k=0}^n M_k \binom{n}{k} = C_{n+1},$$

we arrive to the famous Touchard's Identity:

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} C_k \binom{n}{2k} 2^{n-2k} = C_{n+1}.$$

3.2 An Equation for the Motzkin Numbers

Let us count the number of elements in $POI(n)$ with only *Is*, *Es*, and *1s*. We have established above that there are M_n ways to do so. Let us now use the Catalan numbers to count the same collection. For fixed $k \leq \lfloor \frac{n}{2} \rfloor$, let us first arrange k *Is* and k *Es* in C_k ways. Now let us place these symbols in n slots, with $\binom{n}{2k}$ possibilities. The remaining slots are all filled with *1s*. After iterating through all possible values of k , we have the following equation.

Theorem 3 For all $n \geq 0$,

$$M_n = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} C_k,$$

which relates the Motzkin numbers as a summation involving the Catalan numbers. Thus, we have discovered an identity from [8] through the novel use of the ICE notation to represent the elements of $POI(n)$.

3.3 Cardinality of $POI(n)$

3.3.1 Utilizing Original Notation

Define the set of elements of $POI(n)$ of rank k by $POI_k(n)$. Then

$$|POI_k(n)| = \binom{n}{k}^2.$$

This is because k out of n numbers are selected for the domain and the range independently. Once all the values in the domain and range are chosen, there is only one possible arrangement by the requirement that order is preserved by elements in $POI(n)$. Then, since

$$POI(n) = \bigcup_{k=0}^n POI_k(n),$$

we know that

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

The second equality can be proven in myriad ways, whether algebraic or combinatorial [12].

3.3.2 Utilizing ICE Notation

Definition Any integer $1 \leq r \leq n$ that satisfies both of the following conditions is a **splittable slot**:

- 1) For all $\ell \leq r$, either ℓ is not in $Dom(\sigma)$ or $\sigma(\ell) \leq r$.
- 2) For all $\ell > r$, either ℓ is not in $Dom(\sigma)$ or $\sigma(\ell) > r$.

Definition Let $\sigma \in POI(n)$. We denote the left-most splittable slot of σ containing E by v , if it exists.

Remark The ICE notation of any partial identity contains only 1s and Fs, for which no such v will exist.

Example 9 Let $\sigma = 1IIECEIEIE = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ - & 4 & 5 & - & 6 & - & 8 & - & 10 & - \end{pmatrix}$. We have splittable slots of 1, 6, 8, and 10, and $v = 6$.

Remark Note that n is always a splittable slot for any $\sigma \in POI(n)$.

Definition Let $1 \leq r \leq n$ be a splittable slot. For $\sigma \in POI(n)$, let $\sigma|_{\{1, \dots, r\}}$ or $\sigma_r \in POI(r)$ denote the **restricted partial order preserving injection** that maps $\{1, 2, \dots, r\}$ to itself, such that, for $i \in Dom(\sigma)$ and $i \leq r$, $\sigma_r(i) = \sigma(i)$.

Similarly, let $\sigma^r = \sigma|_{\{r+1, r+2, \dots, n\}} \in POI(n-r)$ be defined by $\sigma^r(i) = \sigma(i)$ for $i \in DOM(\sigma)$ and $i > r$.

Definition The **concatenation** of elements $\sigma \in POI(r)$ and $\tau \in POI(n-r)$ forms an element $\alpha \in POI(n)$ according to the following rule: $\alpha(i) = \sigma(i)$ for $1 \leq i \leq r$ and $\alpha(\ell + r) = \tau(\ell) + r$ for $1 \leq \ell \leq n-r$. We will denote α by $\sigma||\tau$.

Example 10 For instance, for $r = 3$ and $n = 7$, we have

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

Remark Note that r is a splittable slot of $\sigma||\tau$. For $\alpha = \sigma||\tau$, the process of taking the inverse over only σ can be written as $\alpha_r^{-1}||\alpha^r$. For instance, if $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & - & - & 6 & - \end{pmatrix}$, then 4 is a splittable slot, and

$$\alpha_4^{-1}||\alpha^4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ - & - & 1 & 2 & 6 & - \end{pmatrix} \in POI(6).$$

Cardinality of the set of Partial Identities Let us find the number of partial identities in $POI(n)$ by utilizing the ICE notation. These elements have no Is or Es. In other words, they only have gaps and fixed points in each slot, with no restrictions on placing these symbols. Thus, given n slots, there are 2^n partial identities.

Henceforth, we shall count the number of $\sigma \in POI(n)$ such that there exists at least one $c \in Dom(\sigma)$ such that $\sigma(c) \neq c$. Thus, c is neither a gap nor a fixed point.

Lemma: Assume $\sigma \in POI(n)$ and $c \in DOM(\sigma)$ such that $\sigma(c) \neq c$. Then there exists a left-most splittable slot v which contains E, such that σ_v belongs to either $POI_{asc}(v)$ or $POI_{des}(v)$, but not both.

Proof We already know that all elements that are in both $POI_{asc}(n)$ and $POI_{des}(n)$ are partial identities. Since $\sigma(c) \neq c$, σ is not a partial identity. Thus, we now must prove that σ_v is in one of $POI_{asc}(v)$ or $POI_{des}(v)$. Let us proceed by way of contradiction. Suppose that σ_v is neither in $POI_{asc}(v)$ nor $POI_{des}(v)$. Then there must exist a splittable slot u , where $u < v$, that contains E. This remark is illustrated in Examples 11 and 12 below. This contradicts the definition of v as the left-most splittable slot that contains E.

Example 11 Let $\sigma_8 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & - & - & - & - & 5 & 7 \end{pmatrix}$ for $\sigma \in POI(8)$. Then

$\sigma_8 \notin POI_{asc}(8)$ and $\sigma_8 \notin POI_{des}(8)$, and it has a splittable slot at $4 < 8$.

Example 12 Let $\sigma_7 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & - & 1 & 2 & 6 & 7 & - \end{pmatrix}$ for $\sigma \in POI(7)$. Once again, we note that $\sigma_7 \notin POI_{asc}(7)$ and $\sigma_7 \notin POI_{des}(7)$. It has a splittable slot at $4 < 7$.

Cardinality of the set of $\sigma_v \in POI(v)$ Now, we return to counting the number of possible elements $\sigma_v \in POI(v)$ with v being the left-most splittable slot containing E.

Since $|POI_{asc}(n)| = |POI_{des}(n)|$, the total number of possible σ_v is twice the total number of possible $\sigma_v \in POI_{asc}(v)$. Thus, the number of possible $\sigma_v || \sigma^v$ equals the number of $\sigma_v^{-1} || \sigma^v$.

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & - & - & 6 & 7 & - & - \end{pmatrix}}_{\sigma_v} \longleftrightarrow \underbrace{\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ - & - & 1 & 6 & 7 & - & - \end{pmatrix}}_{\sigma_v^{-1}}$$

$\swarrow v$ $\swarrow v$
 σ^v σ^v

Let us find the number of possible $\sigma_v \in POI_{asc}(v)$. We can now utilize the ICE notation, as we have restricted the domain of consideration. We have several combinatorial tasks: find the slots for Is and Es, arrange the Is and Es, and fill in the remaining slots of σ_v with Cs, 1s, and Fs.

Let us first only consider Is and Es before we place other symbols. Suppose that we have k Is and k Es in σ_v . We note that, by the definition of v , we place E in slot v . We must choose which $2k - 1$ of the remaining $v - 1$ slots in σ_v to place the remaining Is and Es. This gives us $\binom{v-1}{2k-1}$ possibilities.

Because we must always have a strictly *greater* number of Is and Es to the left of any slot in σ_v , except v itself, by the definition of v , we fix an I to the first of the chosen slots. Then, we arrange the remaining $k - 1$ Is and $k - 1$ Es according to the usual requirements on ICE strings, with C_{k-1} possibilities.

We then fill the remaining $v - 2k$ slots with Cs, 1s, or Fs, depending on the situation. In each slot, we either choose an A (as defined in Section 3.1.2) or 1, resulting in 2^{v-2k} possibilities. We have now identified the combinatorial tasks required to construct $\sigma_v \in POI_{asc}(v)$. Explicitly, there are

$$\binom{v-1}{2k-1} C_{k-1} (2^{v-2k})$$

such $\sigma_v \in POI_{asc}(v)$ with k Is and k Es. See Example 13 below for a visual representation of such σ_v . Thus, recalling that $|POI_{asc}(v)| = |POI_{des}(v)|$, there

are

$$2 \binom{v-1}{2k-1} C_{k-1}(2^{v-2k})$$

such $\sigma_v \in POI_{asc}(v) \cup POI_{des}(v)$ with k Is and k Es. Using the Lemma, this union comprises all such $\sigma_v \in POI(v)$.

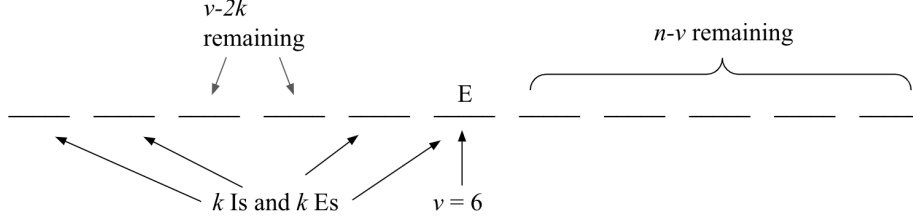


Fig. 5. An abstract construction where $v = 6$, $n = 11$, and $k = 2$. Slots 1, 2, 5, and 6 contain Is and Es, leaving two slots that can be filled independently with either a 1 or C.

Example 13 We now demonstrate possible combinations for $\sigma_6 \in POI_{asc}(6)$:

I I C C E E	I I C 1 E E	I I 1 C E E	I I 1 1 E E
I C I C E E	I C I 1 E E	I 1 1 C E E	I 1 1 1 E E
I C C I E E	I C 1 I E E	I 1 C I E E	I 1 1 I E E
F I C I E E	F I 1 I E E	1 I C I E E	1 1 1 I E E
F F I I E E	F 1 I I E E	1 C I I E E	1 1 I I E E
I C I E C E	I C I E 1 E	I 1 I E C E	I 1 I E 1 E
I I C E C E	I I C E 1 E	I 1 1 E C E	I 1 1 E 1 E
I I E C C E	I I E C 1 E	I I E 1 C E	I I E 1 1 E
F I I C E E	F I I 1 E E	1 I I C E E	1 I I 1 E E
F F I I E E	1 C I I E E	F 1 I I E E	1 1 I I E E

Note that each column has the same choice of 1s and As, and each row has Is and Es in the same places. We have four columns for the $2^{v-2k} = 2^2$ possibilities to choose between 1 and A for two slots. We have ten rows for the $\binom{v-1}{2k-1} = \binom{5}{3} = 10$ possible ways to place the Is and Es. Then to find the total number of $\sigma_v \in POI_{asc}(v)$, we multiply the number of rows with the number of columns: $2^2 \cdot 10 = 40$ such σ_v .

Cardinality of the set of $\sigma^v \in POI(n-v)$ Having found the number of possibilities for $\sigma_v \in POI(v)$, we now consider the $n-v$ slots to the right of v that remain to be filled as we construct σ^v . We know that σ^v is itself an element of $POI(n-v)$, which we concatenate with $\sigma_v \in POI(v)$ to form $\sigma \in POI(n)$. This means that there are $|POI(n-v)|$ ways to fill these remaining slots. We know that

$$|POI(n-v)| = \binom{2n-2v}{n-v}$$

as established in Section 3.3.1.

Summation Bounds The total number of $\sigma \in POI(n)$ with fixed v and k , where σ is not a partial identity, is

$$\binom{v-1}{2k-1} C_{k-1}(2^{v-2k}) \binom{2n-2v}{n-v}.$$

We now sum through all possible values of v (2 to n ; $v \neq 1$ since by the definition of v , E is required to occupy slot v , but an ICE string cannot begin with an E as we cannot end a path that was never initiated to begin with), and k (1 to $\lfloor \frac{v}{2} \rfloor$; we have k Is and k Es among the first v slots, so $2k \leq v$).

We now have bounds of the summation, but we can extend these bounds. Begin summing v with $v = 1$, but note that the corresponding value is 0. We also sum k when $\lfloor \frac{v}{2} \rfloor < k \leq v$, but once again, the corresponding values are 0. Thus, we have shown the following theorem.

Theorem 4 For any $n \geq 0$,

$$|POI(n)| = 2^n + 2 \sum_{v=1}^n \sum_{k=1}^v \binom{v-1}{2k-1} C_{k-1}(2^{v-2k}) \binom{2n-2v}{n-v},$$

or rephrased,

$$\binom{2n}{n} = 2^n + 2 \sum_{v=1}^n \sum_{k=1}^v \binom{v-1}{2k-1} C_{k-1}(2^{v-2k}) C_{n-v}(n-v+1).$$

3.4 Future Work

The study of $POI(n)$ clearly introduces a wealth of combinatorial identities to analyze. By changing axioms and focusing on specific characteristics of elements, we formulate different counting problems revolving around the same monoid. We can also construct different notations to develop new ways of counting elements. For instance, the **position notation** is an n -element string, where slot k is occupied by the number of the path that k is in. This notation reveals interesting patterns of $POI(n)$.

Acknowledgements

We would like to thank the California State University, Fullerton (CSUF) Math Department for its facilities and resources, as well as the CSUF Mathematics Summer Research Program for a mini-grant that supported this research. We would also like to thank previous researchers of $SIM(n)$ and $POI(n)$ for providing the base for our research. We are grateful for family and friends for

their continued support, and Tyler Kim is extremely thankful to Scott Annin for his invaluable guidance, mentorship, and teachings throughout this research journey.

References

- [1] Alvarez, U. (2016) - *On k th roots in semigroups of order-preserving partial permutations*, CSUF NSM Dimensions, 123-132.
- [2] Annin, S.A. (1997) - *Hierarchy of efficient generators of the symmetric inverse monoid*, Semigroup Forum, 55, 327-355.
- [3] Annin, S. A., Jansen, T., and Smith, C. (2009) - *On k th roots in the symmetric and alternating groups*, Pi Mu Epsilon Journal, 12(10), 581-589.
- [4] Annin, S.A., Cannon, T., Hernandez, C., and Torres, L. (2012) - *On k th roots in the symmetric inverse monoid*, Pi Mu Epsilon Journal, 13(6), 321-331.
- [5] Annin, S.A., and Lopez, S. (2020) - *Minimal generating sets of the monoid of partial order-preserving injections*, PUMP Journal of Undergraduate Research, 3, 190-204.
- [6] Clifford, A., H. and Preston, G.,B. - *The Algebraic Theory of Semigroups. Vol. 1*, American Mathematics Society, Providence, RI, 1961.
- [7] Davis, Tom (2016) - *Catalan Numbers*. Geometer.
- [8] Deutsch, Emeric and Sagan, Bruce - *Congruences for Catalan and Motzkin numbers and related sequences*, 2018.
- [9] GAP 2018, The GAP Group, GAP-Groups, Algorithms, and Programming Version 4.8.10; 2018. (<https://www.gap-system.org>)
- [10] Hoang, S.T., - *On k th Roots of Partial Order-Preserving Injections*, CSUF.
- [11] Hungerford, T. (2012) - *Abstract Algebra: An Introduction (3rd ed.)*. Cengage Learning.
- [12] Miklós Bóna (2016) - *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory*, Mathematical Association of America, 77.
- [13] The Online Encyclopedia of Integer Sequences - *Catalan Numbers*, OEIS Foundation.
- [14] The Online Encyclopedia of Integer Sequences - *Motzkin Numbers*, OEIS Foundation.
- [15] Wang, Yi and Zhang Zhi-Hai (2015) - *Combinatorics of Generalized Motzkin Numbers*, Journal of Integer Sequences, Vol. 18, Article 15.2.4.
- [16] Wolfram MathWorld- *Catalan Number*, Wolfram Alpha.

Efficiency of Horner's Method for Binomial Expansions

Cadence Pinkerton
Advisor: Dr. Adam Glessner

Abstract

Horner's Method is an extremely efficient tool that can be implemented in place of expansion using the binomial formula. In this paper, we review the Horner's Method algorithm and two applications: partial fraction decomposition with repeated linear factors, and the Eisenstein criterion. Within each application, we also demonstrate that Horner's Method has an organizational component that makes it easier to identify properties of the methods. Lastly, we show that Horner's Method is an order of magnitude more efficient than binomial expansion by calculating the number of computations required for each technique.

1 Introduction

Although the technique described in this paper is commonly called *Horner's method*, Horner himself attributed it to Lagrange [2], and its history dates back to at least the Han Dynasty in China. It is an efficient multipurpose tool that is known to simplify many computations. Our goal in this paper is to highlight two of these applications which are not very well known and to evaluate how efficient the method is compared to an obvious alternative, namely expansion using the binomial formula.

The plan is as follows. In Section 2, we will demonstrate one variation of Horner's method crucial to the two applications in the paper. In Section 3, we will show how to utilize Horner's method to compute the partial fraction decomposition of a rational function with repeated linear factors in the denominator, as well as how to perform shifts of polynomials in conjunction with Eisenstein's irreducibility criterion. Finally, in Section 4, we show that solving either of the above problems using the binomial theorem requires $\mathcal{O}(n^3)$ operations, while Horner's method only requires $\mathcal{O}(n^2)$ operations, thus verifying the increased efficiency of Horner's method.

2 Horner's Method

Prior to investigating the efficiency of Horner's Method, we will review the basic algorithm. Instead of giving an abstract presentation, it is simpler to give an example. Let $f(x) = x^3 - 3x^2 + 3x + 4$. Our goal is to rewrite f in terms of $x - 2$, i.e., of the form

$$f(x) = (x - 2)^3 + a_2(x - 2)^2 + a_1(x - 2) + a_0.$$

The method is organized in a table (see Figure 1) with the shift **(2)** in the top left-hand corner and the polynomial's coefficients across the top.

	1	-3	3	4
2	1	-1	1	6
	1	1	3	
	1	3		
	1			

Figure 1: Applying Horner's method to write $x^3 - 3x^2 + 3x + 4$ in terms of $x - 2$.

In the table above, we copy the leading coefficient 1 four times vertically, matching the number of coefficients. Note that the number of coefficients for the polynomial $x^3 + 1$ is still four since we would write it as $x^3 + 0x^2 + 0x + 1$ and the coefficient list in that case would be **1 0 0 1**. Let us focus on the second row of the table. We multiply the shift term, **2**, by the first 1 in the second row and add it to the coefficient above in the next column, **-3**. The resulting number is -1 which is recorded under the second coefficient. Repeating the same steps with -1 , we multiply our shift term **2** by -1 and add it to the coefficient above in the next column, **3**, yielding a 1, which is placed in our table under the third coefficient. Finally, we multiply the shift by 1 and add 4, indicating that a **6** is placed in the last column. Moving onto the second row, we are now going to treat row 2 as our new coefficients. Using the new coefficients, we follow the same steps used to create row 1. We finish row 2 after we have added to the third coefficient. This pattern continues until you are left with your first coefficient. Our new polynomial is $f(x) = (x - 2)^3 + 3(x - 2)^2 + 3(x - 2) + 6$. The coefficients of this polynomial are the diagonals from our table, read from south west to north east.

	a_3	a_2	a_1	a_0
k	a_3	$a_3k + a_2$	$a_3k^2 + a_2k + a_1$	$a_3k^3 + a_2k^2 + a_1k + a_0$
	a_3	$2a_3k + a_2$	$3a_3k^2 + 2a_2k + a_1$	
	a_3	$3a_3k + a_2$		
	a_3			

Figure 2: Using Horner's method to write $a_3x^3 + a_2x^2 + a_1x + a_0$ in terms of $x - k$.

Generalizing the prior example, we can write $f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$

in terms of $x - k$ (see Figure 2). We therefore obtain

$$f(x) = a_3(x - k)^3 + (3a_3k + a_2)(x - k)^2 + (3a_3k^2 + 2a_2k + a_1)(x - k) + (a_3k^3 + a_2k^2 + a_1k + a_0).$$

This generalized table displays makes clear that the first coefficient of our shifted polynomial will always be the same as the original polynomial. The table only includes coefficients, which is easy to implement computationally.

Let us compare this to the coefficients we get through binomial distribution. Note that the simplest way to obtain the coefficients in this manner is to evaluate $f(x + k)$.

$$\begin{aligned} a_3(x + k)^3 + a_2(x + k)^2 + a_1(x + k) + a_0 &= a_3x^3 + 3a_3kx^2 + 3a_3k^2x + a_3k^3 \\ &\quad + a_2x^2 + 2a_2kx + a_2k^2 \\ &\quad + a_1x + a_1k \\ &\quad + a_0. \end{aligned}$$

Combining like terms, we obtain:

$$f(x + k) = a_3x^3 + (3a_3k + a_2)x^2 + (3a_3k^2 + 2a_2k + a_1)x + (a_3k^3 + a_2k^2 + a_1k + a_0)$$

and hence

$$f(x) = a_3(x - k)^3 + (3a_3k + a_2)(x - k)^2 + (3a_3k^2 + 2a_2k + a_1)(x - k) + (a_3k^3 + a_2k^2 + a_1k + a_0).$$

This is the same result we obtained using Horner's method. An important question is whether one approach is more computationally efficient than the other. We will show in Section 4 that Horner's method is an order of magnitude more efficient.

3 Applications

3.1 Partial Fraction Decomposition with Repeated Linear Factors

Partial fraction decomposition is a method used to express a rational function as a sum of simpler fractions. It is often used to simplify integrals [4] and in solving differential equations via the Laplace transform [1]. An example of a partial fraction decomposition with repeated linear factors is:

$$\frac{x^3 - x + 1}{(x - 2)^5} = \frac{A}{(x - 2)} + \frac{B}{(x - 2)^2} + \frac{C}{(x - 2)^3} + \frac{D}{(x - 2)^4} + \frac{E}{(x - 2)^5},$$

where the challenge is to determine the coefficients A, B, C, D, E . This is typically done either by using the binomial theorem or differentiation, both of which

are extremely labor intensive here. Using Horner's Method makes finding the values quite simple. We construct our table in Figure 3 and obtain four coefficients **1 6 11 7**. The key insight is that we start substituting 7 for E and working our way backward.

	1	0	-1	1
2	1	2	3	7
	1	4	11	
	1	6		
	1			

Figure 3: Using Horner's method to compute the partial fraction decomposition of a rational function with repeated linear factors.

Therefore, we have $E = 7$, $D = 11$, $C = 6$, $B = 1$ and, since there is nothing left, $A = 0$. The final partial fraction decomposition is

$$\frac{x^3 - x + 1}{(x - 2)^5} = \frac{1}{(x - 2)^2} + \frac{6}{(x - 2)^3} + \frac{11}{(x - 2)^4} + \frac{7}{(x - 2)^5}.$$

Using Horner's Method also reveals some properties of partial fractions. For example, the first nonzero numerator will always be the leading coefficient of the polynomial. Also, if the degrees of the numerator and denominator are k and ℓ , respectively, then the first $\ell - k - 1$ terms of the decomposition will have coefficient 0.

3.2 The Eisenstein Criterion

The Eisenstein criterion [3] is a method used to determine whether a polynomial with integer coefficients is irreducible. The criterion states that a polynomial is irreducible if there exists a prime number such that the prime number divides all coefficients except the leading coefficient and the square of the prime number does not divide the constant coefficient. For cases where a polynomial is irreducible but the Eisenstein criterion is not satisfied, we can sometimes shift the polynomial, resulting in coefficients that satisfy the criterion.

Notation 1. For $f \in \mathbb{Z}[x]$, the set of polynomials with coefficients in $\mathbb{Z}$, let $f_{+k}(x) = f(x + k)$.

Definition 2. Let $f \in \mathbb{Z}[x]$. If $k \in \mathbb{Z}$ such that f_{+k} satisfies the Eisenstein criterion, then f_{+k} is an Eisenstein shift of f .

Lemma 3. If $f \in \mathbb{Z}[x]$, then following statements are equivalent.

1. f is irreducible.
2. f_{+k} is irreducible for some $k \in \mathbb{Z}$.
3. f_{+l} is irreducible for all $l \in \mathbb{Z}$.

Proof. If f is irreducible, then f_{+0} is irreducible. If there exists $k \in \mathbb{Z}$ such that f_{+k} is irreducible, if $\ell \in \mathbb{Z}$, and if $g, h \in \mathbb{Z}[x]$ such that $f_{+\ell} = gh$, then

$$f_{+k} = (f_{+\ell})_{+(k-\ell)} = g_{+(k-\ell)}h_{+(k-\ell)}.$$

Since f_{+k} is irreducible, we conclude that one of $g_{+(k-\ell)}$ or $h_{+(k-\ell)}$ is a unit in $\mathbb{Z}[x]$, i.e., is equal to ± 1 . Since shifting a constant does not change the constant, we conclude that g or h equals ± 1 , and hence $f_{+\ell}$ is irreducible. Finally, if $f_{+\ell}$ is irreducible for all $\ell \in \mathbb{Z}$, then $f = f_{+0}$ is irreducible. $\square$

The above lemma motivates the use of Horner's method since Horner's method gives us an efficient way to compute f_{+k} for a given polynomial f .

Consider the example $f(x) = x^2 + x + 1$. In this case, no prime divides any of the coefficients and so f fails to satisfy the Eisenstein Criterion. This is a case where we could use an Eisenstein shift to meet these conditions. However, applying Horner's method (see Figure 4) we compute the Eisenstein shift f_{+1} .

$$\begin{array}{c|ccc} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & 1 & 2 & \mathbf{3} \\ & 1 & \mathbf{3} & \\ & \mathbf{1} & & \end{array}$$

Figure 4: Using Horner's method to compute f_{+1} where $f(x) = x^2 + x + 1$.

Thus, $f_{+1}(x) = x^2 + 3x + 3$, which is irreducible by Eisenstein with $p = 3$. Had we not "guessed" correctly to use f_{+1} , we could work generally, using the Eisenstein shift f_{+k} . For example, $f_{+k}(x) = (x+k)^2 + (x+k) + 1$. Distributing using Horner's method returns $f_{+k}(x) = x^2 + (2k+1)x + (k^2+k+1)$. As the coefficient of x is odd, we immediately see that there is not Eisenstein shift after which f_{+k} will satisfy the Eisenstein criterion for the prime 2. Through trial and error (a computer can do this quite quickly!) we can check whether different values for k lead to a successful Eisenstein shift.

$$\begin{array}{c|ccc} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{k} & 1 & k+1 & k^2+k+1 \\ & 1 & \mathbf{2k+1} & \\ & \mathbf{1} & & \end{array}$$

Figure 5: Using Horner's method to compute f_{+k} where $f(x) = x^2 + x + 1$.

4 Results

To demonstrate the efficiency of Horner's Method compared to binomial expansion, we calculated the number of computations used for each. First, for

binomial expansions, we separately count the number of multiplications and additions. Using the binomial theorem, we have

$$\sum_{j=0}^n a_j (x+k)^j = \sum_{j=0}^n a_j \sum_{i=0}^j \binom{j}{i} x^{j-i} k^i.$$

Reindexing and combining like terms we obtain

$$\sum_{\ell=0}^n \left[\sum_{m=\ell}^n \binom{m}{m-\ell} a_m k^{m-\ell} \right] x^\ell$$

For each (ℓ, m) , we compute the number of multiplications needed to find $\binom{m}{m-\ell} a_m k^{m-\ell}$. This information is listed in Figure 6.

	$\binom{m}{m-\ell}$	$k^{m-\ell}$	$\binom{m}{m-\ell} \cdot a_m$	$\binom{m}{m-\ell} a_m \cdot k^{m-\ell}$	Total
$m - \ell > 1$	$2m - 4$	$m - \ell - 1$	1	1	$3m - \ell - 3$
$m - \ell \leq 1$	1	0	1	1	3

Figure 6: Computing the number of multiplications when using the binomial expansion.

Unsurprisingly, the number of operations required will depend on the degree, n , of the polynomial. As such, we pause for a moment here to recall some standard notation from mathematics and computer science for evaluating the order of magnitude of the number of operations, namely *big- O notation*.

Notation 4. Let f and g be two polynomials over $\mathbb{Z}$. We write $f \in O(g)$ if there exists $C > 0$ such that, for all $n \in \mathbb{Z}_{\geq 1}$, we have $|f(n)| \leq Cg(n)$.

For example, $5n^2 + 6n - 7 \in O(n^2)$ since, for $n \geq 1$, the triangle inequality implies that

$$|5n^2 + 6n - 7| \leq 5n^2 + 6n + 7 \leq 5n^2 + 6n^2 + 7n^2 = 18n^2.$$

We extend this notation a bit for our purposes.

Notation 5. If $\alpha, n, k \in \mathbb{Z}_{\geq 0}$ and $\alpha \leq n$, then

$$\sum_{\ell=\alpha}^n \mathcal{O}(\ell^k) = \left\{ \sum_{\ell=\alpha}^n f(\ell) \mid f(\ell) \in \mathcal{O}(\ell^k) \right\}.$$

Theorem 6. If f is a function such that $f(\ell) \in \mathcal{O}(\ell^k)$, then

$$\sum_{\ell=\alpha}^n f(\ell) \in \mathcal{O}(n^{k+1}).$$

Proof. If $f(\ell) \in \mathcal{O}(\ell^k)$ then there exists $M > 0$ such that $|f(\ell)| \leq M\ell^k$. Therefore,

$$\begin{aligned} \left| \sum_{\ell=\alpha}^n f(\ell) \right| &\leq \sum_{\ell=\alpha}^n |f(\ell)| \leq \sum_{\ell=\alpha}^n M\ell^k = M \sum_{\ell=\alpha}^n \ell^k \\ &\leq M \sum_{\ell=\alpha}^n n^k = M(n - \alpha)n^k = M(n^{k+1} - \alpha n^k) \in \mathcal{O}(n^{k+1}). \end{aligned}$$

Thus, $\sum_{\ell=\alpha}^n f(\ell) \in \mathcal{O}(n^{k+1})$. $\square$

Applying this to the resulting number of computations we have

$$\begin{aligned} \sum_{\ell=0}^n \left(3 + \sum_{m=\ell+2}^n (3m - \ell - 3) \right) &\in \sum_{\ell=0}^n \left[\mathcal{O}(1) + \sum_{\ell+2}^n \mathcal{O}(m) + \ell \sum_{m=\ell+2}^n \mathcal{O}(1) \right] \\ &\subseteq \sum_{\ell=0}^n (\mathcal{O}(1) + \mathcal{O}(n^2) + \ell \mathcal{O}(n)) \\ &\subseteq \mathcal{O}(n^2) \sum_{\ell=0}^n \mathcal{O}(1) + \mathcal{O}(n) \sum_{\ell=0}^n \mathcal{O}(\ell) \\ &\subseteq \mathcal{O}(n^3). \end{aligned}$$

Thus, our number of multiplications is of order $\mathcal{O}(n^3)$. Next, we find the number of additions. For each (ℓ, m) , we have $n - \ell + 1$ choices, and so the order of the total number of additions is:

$$\sum_{l=0}^n n - l + 1 \in \mathcal{O}(n^2).$$

Combining the number of additions and multiplications results, the total number of computations for binomial expansion is order $\mathcal{O}(n^3)$.

Now we investigate Horner's Method. Our table will look something like what is found in Figure 7.

	a_0	$\cdots$	a_n
k	a_0	$\cdots$	c_n
	a_0	$\cdots$	c_{n-1}
		$\ddots$	
	$\vdots$	c_1	
	$a_0 = c_0$		

Figure 7: A generic example of Horner's method.

A convenient property within the table is that the number of multiplications is equal to the number of additions. This implies that the number of total

computations is equal to $2(\# \text{ of multiplications})$. Each row contains a number of multiplications equal to the index of the last term, i.e.,

$$\# \text{ of multiplications} = \sum_{\ell=0}^n \ell \in \mathcal{O}(n^2).$$

Comparing the two methods, binomial expansion is order $\mathcal{O}(n^3)$ and Horner's Method is order $\mathcal{O}(n^2)$, an order of magnitude more efficient! The table also adds an organizational component that makes it visually easier to follow.

5 Conclusion

While Horner's Method has a long history, new applications for this technique continue to be discovered. This paper examined two such applications and investigated the efficiency of Horner's Method compared to expansion using the binomial formula. In Section 2, we reviewed a common variation of Horner's Method and the basic algorithm to provide motivation for our suggested applications. An important aspect of this method is the table used to visualize the computations. Using this table, we can focus directly on the polynomial's coefficients and the shift term, adding an organizational component to the analysis.

In Section 3, we introduced Horner's Method for partial fraction decomposition with repeated linear factors and the Eisenstein criterion. With partial fraction decomposition, Horner's Method provides the desired coefficients for the sum of simpler fractions while revealing properties that are not commonly observed using other methods. For instance, the first nonzero numerator is the leading coefficient of the polynomial. Additionally, we note that the first $l - k - 1$ terms of the decomposition will have coefficient 0, where the degrees of the numerator are k and the degrees of the denominator are l .

Following partial fraction decomposition, we introduced the Eisenstein criterion and the Eisenstein shift. When shifting a given polynomial f , we can apply Horner's method to easily and efficiently compute f_{+k} . Using the shift f_{+k} , Horner's method provides us with the general coefficients, which can be evaluated for different values of k to check for a successful Eisenstein shift.

In Section 4, we investigated the efficiency of Horner's Method compared to binomial expansion by calculating the number of computations required to perform each method. The calculations verify that solving problems using Horner's Method (order $\mathcal{O}(n^2)$) is more efficient by an order of magnitude than solving problems using binomial expansion (order $\mathcal{O}(n^3)$).

Overall, Horner's Method is a more efficient and organized alternative to binomial expansion that can be applied in a variety of methods.

References

- [1] Goode, Stephen W., and Annin, Scott A., *Differential Equations & Linear Algebra, 4th Edition*. Pearson. 2017.

- [2] Horner, William George, *A new method of solving numerical equations of all orders, by continuous approximation*. Philosophical Transactions. Royal Society of London. July 1819. 109: 308–335.
- [3] Hungerford, Thomas H., *Abstract Algebra: An Introduction, 3rd Edition*. Cengage Learning. 2012.
- [4] Stewart, James, *Essential Calculus: Early Transcendentals, 2nd Edition*. Cengage Learning. 2013.

A mathematical investigation of the frontogenesis function

Yulia Potyrina

Advisor: Christopher Lyons

Abstract

Inspired by its uses in meteorology, we study the *frontogenesis* function obtained from some simple scalar fields and notable vector fields. We focus especially on vector fields arising from complex-analytic functions. Based on our observations, we prove several results and offer one conjecture.

1 Introduction

The subject of meteorology, which studies weather and other atmospheric phenomena, affects all of our lives. Its impacts range from the everyday convenience of checking the forecast for tomorrow's temperature to lifesaving predictions about major storms and heat waves. Behind the scenes, meteorology is also an intensively mathematical subject. These mathematical models, derived from the physical principles of fluids, have made the endeavors of weather forecasting so successful.

Our aim in this article is to go the other way. Not to accurately model the weather using mathematics, but rather to apply mathematical tools used in meteorology to situations of purely mathematical interest. One motivation is that these mathematical situations might be regarded as overly simple, idealized “toy models” of weather. However, the study of natural phenomena has always led to interesting new mathematics, and it is in this spirit that we undertake our investigation.

In this paper we focus upon one specific meteorological concept related to *fronts*, which are boundaries between air masses of different temperatures (see Figure 1). When one passes through a front, a large temperature change occurs over a relatively short distance. The temperature at each point is represented by a type of function called a *scalar field*, and the large temperature change implies the magnitude of its *gradient vector field* is large near a front.

If, at some given point and given moment, the temperature gradient is increasing in magnitude, it might indicate that a front is beginning to form or continuing to strengthen there, a process called *frontogenesis*. Conversely, if the temperature gradient is decreasing, it could indicate the breakdown of a front;

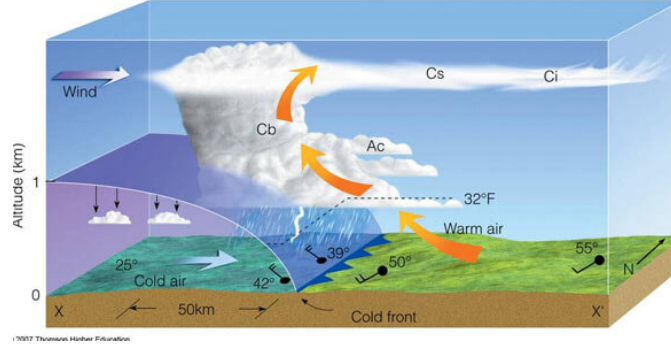


Figure 1: Visualization of a front between warm and cold air masses. (Source: Thomson Higher Education)

this is called *frontolysis*. The actions of the wind can have significant frontogenetic or frontolytic effects, since wind can move areas of hot and cold air closer together or farther apart. Wind is represented by a *vector field*, since it has a direction and magnitude at each point. The *frontogenesis function*, which is our main object of study in this paper, quantifies the frontogenetic effect of the wind on the temperature field.

Scalar fields (such as the temperature) and vector fields (such as the wind) are common objects with applications in several branches of science, but are also of interest in their own right. Thus, the frontogenesis function provides a way for a given scalar field and vector field can “interact” with each other. Our approach is start with some interesting vector fields and to see how they affect two simple temperature fields (see Figure 2). The frontogenesis function that results can then be viewed as a compelling secondary feature of the vector field.

Here is an outline of the paper. In §2, we set up our notation and define the frontogenesis function. In §3, we illustrate these concepts with a few examples, and we prove some general results about them in §4. Finally, in §5 we study frontogenesis in the context of vector fields arising from complex analytic functions, proving some results and making one conjecture about these.

2 Basic formulas and notation

We model the wind velocity (i.e., the wind speed and direction) using a vector field. On large scales, the vertical component of wind is small compared to its horizontal component, and so we take the simplified view that the wind velocity is a 2-dimensional vector field, representing the movement of air over a surface. We take this surface to be flat (much as the surface of the earth is approximately flat in small enough regions) and represent points using Cartesian coordinates (x, y) . The wind velocity vector field will then be denoted by $\mathbf{W}(x, y) = \langle u(x, y), v(x, y) \rangle$.

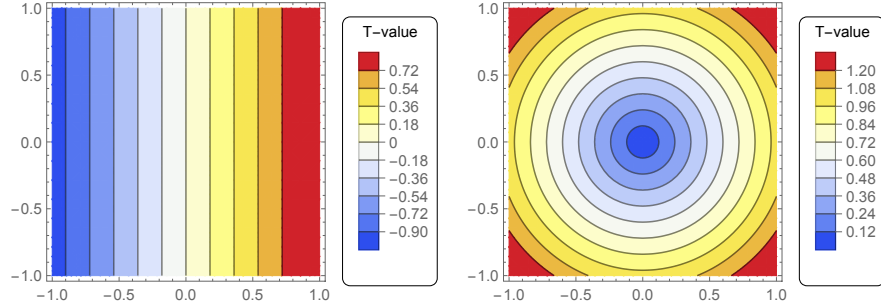


Figure 2: Contour maps of the temperature fields $T_H(x, y) = x$ (left) and $T_R(x, y) = \sqrt{x^2 + y^2}$ (right).

The temperature distribution in our 2-dimensional atmosphere is given by a two-dimensional scalar field that we denote as $T(x, y)$. Two temperature functions that we will use in this paper are

$$T_H(x, y) = x, \quad (1)$$

which represents a temperature distribution that increases uniformly in the left-to-right direction (i.e., the “horizontal” direction), and

$$T_R(x, y) = \sqrt{x^2 + y^2}, \quad (2)$$

which represents a temperature distribution that increases uniformly in the radial direction away from the origin. Contour maps for these temperature fields are shown in Figure 2.

Given a wind velocity field $\mathbf{W} = \langle u, v \rangle$ and a temperature function T , we define the **frontogenesis** function $\mathcal{F}_{\mathbf{W}, T}(x, y)$ associated to $\mathbf{W}$ and T as

$$\mathcal{F}_{\mathbf{W}, T}(x, y) = \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \frac{\partial T}{\partial y} \right) - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \frac{\partial T}{\partial y} \right) \right].$$

When it does not cause confusion, we will often denote $\mathcal{F}_{\mathbf{W}, T}(x, y)$ simply by $\mathcal{F}(x, y)$ to make the notation simpler.

The derivation and physical meaning behind the frontogenesis function is discussed in detail in [2]. Roughly speaking, $\mathcal{F}(x, y)$ represents the rate of change of $|\nabla T|$, the magnitude of the gradient ∇T , as a parcel of air at (x, y) moves along with the wind $\mathbf{W}$. As mentioned in §1, $|\nabla T|$ should be large near a front, and therein lies the interest in determining its rate of change. When $\mathcal{F}(x, y) > 0$, and we view this as a sign that the wind vector field is causing the (magnitude of the) gradient of T to increase, and we say that *frontogenesis* is occurring at (x, y) . Conversely, $\mathcal{F} < 0$ indicates *frontolysis*.

3 Some preliminary examples

To illustrate the frontogenesis function, we start with an example having a specific wind velocity field and temperature function.

Example 3.1. Let $\mathbf{W}(x, y) = \langle x^2 - y^2, -2xy \rangle$ and $T = T_R = \sqrt{x^2 + y^2}$. The left side of Figure 3 shows the vector field $\mathbf{W}$ overlaid with the contour map of T_R . First we note that

$$\frac{\partial T}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial T}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}},$$

and therefore

$$|\nabla T| = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2} = \sqrt{\frac{x^2 + y^2}{x^2 + y^2}} = 1.$$

Moreover,

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial}{\partial x}(x^2 - y^2) = 2x, & \frac{\partial u}{\partial y} &= \frac{\partial}{\partial y}(x^2 - y^2) = -2y, \\ \frac{\partial v}{\partial x} &= \frac{\partial}{\partial x}(-2xy) = -2y, & \frac{\partial v}{\partial y} &= \frac{\partial}{\partial y}(-2xy) = -2x. \end{aligned}$$

Thus,

$$\begin{aligned} \mathcal{F}_{\mathbf{W}, T_R}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \cdot \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\ &\quad \left. - \left(\frac{\partial T}{\partial y} \right) \cdot \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\ &= -\frac{1}{|\nabla T|} \left[\left(\frac{\partial T}{\partial x} \right)^2 \left(\frac{\partial u}{\partial x} \right) + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial v}{\partial x} \right) \right. \\ &\quad \left. + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \right) + \left(\frac{\partial T}{\partial y} \right)^2 \left(\frac{\partial v}{\partial y} \right) \right] \\ &= (-1) \left[\left(\frac{x^2}{x^2 + y^2} \right) \left(\frac{\partial u}{\partial x} \right) + \left(\frac{xy}{x^2 + y^2} \right) \left(\frac{\partial v}{\partial x} \right) \right. \\ &\quad \left. + \left(\frac{xy}{x^2 + y^2} \right) \left(\frac{\partial u}{\partial y} \right) + \left(\frac{y^2}{x^2 + y^2} \right) \left(\frac{\partial v}{\partial y} \right) \right] \\ &= -\frac{1}{x^2 + y^2} [x^2 \cdot 2x + xy \cdot (-2y) + xy \cdot (-2y) + y^2 \cdot (-2x)] \\ &= \frac{-2x^3 + 6xy^2}{x^2 + y^2}. \end{aligned}$$

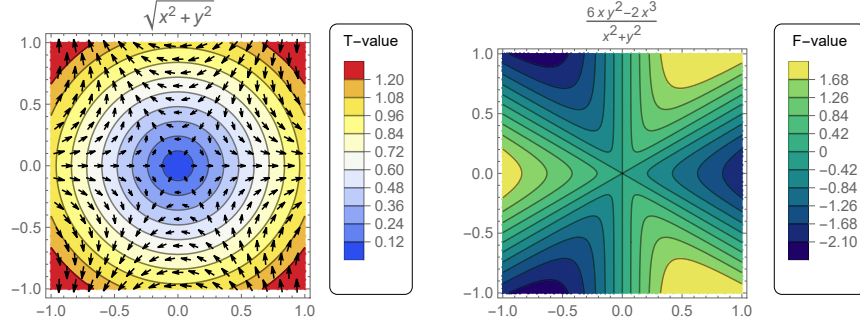


Figure 3: Plot of vector field $\mathbf{W}$ from Example 3.1 overlaid with contour map of T_R (left) and contour map of $\mathcal{F}_{\mathbf{W}, T_R}$ (right).

The right side of Figure 3 shows the contour map of $\mathcal{F}$. In comparing the left and right diagrams, one gets a sense of how $\mathbf{W}$ and T interact to create positive or negative values of $\mathcal{F}$. For instance, in the left figure along the negative x -axis, the “wind velocity” vectors move left-to-right and decrease in magnitude as one approaches the origin. This decreasing magnitude means there is a net influx of “air” of different temperature at each point on the negative x -axis. As a result, the temperature gradient at those points ought to strengthen, and this is reflected on the right side of Figure 3, which shows $\mathcal{F} > 0$ along the negative x -axis. A similar, but opposite, situation occurs along the positive x -axis. On the other hand, at points such as those on the y -axis, the wind velocity vectors are aligned with the level curves of the temperature function, and do not cause air of different temperatures to mix; in those places, we see that $\mathcal{F} = 0$.

We will revisit Example 3.1 later on, when we view the vector field $\langle x^2 - y^2, -2xy \rangle$ as arising from the complex analytic function $f(z) = z^2$; see Example 5.4. Next we consider how $\mathcal{F}$ simplifies for the special temperature functions T_H and T_R .

Example 3.2. Let $\mathbf{W} = \langle u, v \rangle$ be any wind velocity field and let $T = T_H = x$.

$$\frac{\partial T}{\partial x} = 1, \quad \frac{\partial T}{\partial y} = 0,$$

and therefore

$$|\nabla T| = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2} = \sqrt{(1)^2 + (0)^2} = 1.$$

Then,

$$\begin{aligned}
\mathcal{F}_{\mathbf{W}, T_H}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= \frac{1}{|\nabla T|} \left[(-1) \left(\frac{\partial u}{\partial x} \cdot (1) + \frac{\partial v}{\partial x} \cdot (0) \right) \right. \\
&\quad \left. - (0) \left(\frac{\partial u}{\partial y} \cdot (1) + \frac{\partial v}{\partial y} \cdot 0 \right) \right] \\
&= -\frac{\partial u}{\partial x}.
\end{aligned}$$

Example 3.3. Let $\mathbf{W}(x, y) = \langle u, v \rangle$ be a general wind velocity field and $T = T_R = \sqrt{x^2 + y^2}$. As we calculated in Example 3.1, we have

$$\frac{\partial T}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{\partial T}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}, \quad |\nabla T| = 1.$$

Therefore,

$$\begin{aligned}
\mathcal{F}_{\mathbf{W}, T_R}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= -\frac{1}{|\nabla T|} \left[\left(\frac{\partial T}{\partial x} \right)^2 \left(\frac{\partial u}{\partial x} \right) + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial v}{\partial x} \right) \right. \\
&\quad \left. + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \right) + \left(\frac{\partial T}{\partial y} \right)^2 \left(\frac{\partial v}{\partial y} \right) \right] \\
&= - \left[\left(\frac{\partial T}{\partial x} \right)^2 \left(\frac{\partial u}{\partial x} \right) + \left(\frac{\partial T}{\partial y} \right)^2 \left(\frac{\partial v}{\partial y} \right) \right. \\
&\quad \left. + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left[\left(\frac{\partial v}{\partial x} \right) + \left(\frac{\partial u}{\partial y} \right) \right] \right] \\
&= - \left[\left(\frac{x^2}{x^2 + y^2} \right) \left(\frac{\partial u}{\partial x} \right) + \left(\frac{y^2}{x^2 + y^2} \right) \left(\frac{\partial v}{\partial y} \right) \right. \\
&\quad \left. + \left(\frac{xy}{x^2 + y^2} \right) \left[\left(\frac{\partial v}{\partial x} \right) + \left(\frac{\partial u}{\partial y} \right) \right] \right] \\
&= -\frac{1}{x^2 + y^2} \left[x^2 \left(\frac{\partial u}{\partial x} \right) + y^2 \left(\frac{\partial v}{\partial y} \right) + xy \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \right].
\end{aligned}$$

4 Some general properties of the frontogenesis function

We begin by showing what happens to frontogenesis when wind directions are all reversed:

Proposition 4.1. *If we replace $\mathbf{W} = \langle u, v \rangle$ by $-\mathbf{W} = \langle -u, -v \rangle$, then the sign of the frontogenesis function changes; in symbols, $\mathcal{F}_{-\mathbf{W},T}(x, y) = -\mathcal{F}_{\mathbf{W},T}(x, y)$.*

Proof.

$$\begin{aligned}
 \mathcal{F}_{-\mathbf{W},T}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial(-u)}{\partial x} \frac{\partial T}{\partial x} + \frac{\partial(-v)}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
 &\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial(-u)}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial(-v)}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
 &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(-\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} - \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
 &\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(-\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} - \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
 &= \frac{(-1)}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
 &\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
 &= -\mathcal{F}_{\mathbf{W},T}(x, y)
 \end{aligned}$$

□

Due to Proposition 4.1, we need not expend effort calculating $\mathcal{F}$ for $-\mathbf{W}$ once we have already calculated it for $\mathbf{W}$. Next we indicate two special ways that the frontogenesis function can vanish:

Proposition 4.2. *If $\mathbf{W}$ is a constant vector field, then frontogenesis is zero.*

Proof. If $\mathbf{W} = \langle a, b \rangle$ for a, b some constants, then $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = \frac{\partial u}{\partial y} = \frac{\partial v}{\partial x} = 0$. Hence,

$$\begin{aligned}
 \mathcal{F}_{\mathbf{W},T}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
 &\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right]
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(0 \cdot \frac{\partial T}{\partial x} + 0 \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(0 \cdot \frac{\partial T}{\partial x} + 0 \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= 0.
\end{aligned}$$

□

Proposition 4.3. *If $\mathbf{W}$'s first component u does not depend on x and T only depends on x , then $\mathcal{F}_{\mathbf{W},T}(x, y) = 0$.*

Similarly, $\mathbf{W}$'s second component v does not depend on y and T only depends on y , then $\mathcal{F}_{\mathbf{W},T}(x, y) = 0$.

Proof. If $T = T(x)$ depends on x only, then $\frac{\partial T}{\partial y} = 0$. Similarly, if u does not depend upon x , then $\frac{\partial u}{\partial x} = 0$. Therefore,

$$\begin{aligned}
\mathcal{F}_{\mathbf{W},T}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= \frac{1}{|\nabla T|} \left[- \left(\frac{\partial T}{\partial x} \right) \cdot \left(0 \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot 0 \right) \right. \\
&\quad \left. - (0) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot 0 \right) \right] \\
&= 0.
\end{aligned}$$

The second part of the proposition is proved in a similar way: if $\frac{\partial v}{\partial y} = 0$ and $\frac{\partial T}{\partial x} = 0$ then

$$\begin{aligned}
\mathcal{F}_{\mathbf{W},T}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial u}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial v}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= \frac{1}{|\nabla T|} \left[(0) \cdot \left(\frac{\partial u}{\partial x} \cdot 0 + \frac{\partial v}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial u}{\partial y} \cdot 0 + 0 \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= 0.
\end{aligned}$$

□

5 Vector fields arising from analytic functions

For a complex function $f(z)$ defined on a domain D in the complex plane, it is traditional to denote the real and imaginary parts by u and v , respectively. However, to avoid confusion with the components of the wind velocity field $\mathbf{W} = \langle u, v \rangle$, we will instead put $p = \operatorname{Re}(f)$ and $q = \operatorname{Im}(f)$. There are two related vector fields on D that arise from f :

Definition 5.1. Let $f(z) = f(x + iy) = p(x, y) + iq(x, y)$ be a complex function. Let us say that the standard vector field of f is $\langle \operatorname{Re}(f), \operatorname{Im}(f) \rangle = \langle p, q \rangle$, and the Pólya vector field of f is $\langle \operatorname{Re}(f), -\operatorname{Im}(f) \rangle = \langle p, -q \rangle$.

Remark. While the standard vector field of f may seem like the most natural one to associate to it, the Pólya vector field turns out to be more useful for relating the calculus of complex functions with the calculus of vector fields. In particular, one can express complex line integrals of $f(z)$ in terms of work and flux integrals of its Pólya vector field. See the beautiful book [3] for more on these connections. It is worth noting that the Pólya vector field of f is the standard vector field of its complex conjugate function $\bar{f}$.

Proposition 5.2. Suppose that $f(z)$ is analytic on D . If $\mathbf{W}$ is the standard vector field of f , then

$$\mathcal{F}_{\mathbf{W}, T}(x, y) = |\nabla T| \left(-\frac{\partial p}{\partial x} \right) = -|\nabla T| (\operatorname{Re} f'(z)).$$

Proof. Since $f = p + iq$ is analytic on D , the Cauchy-Riemann equations [1] give

$$\frac{\partial p}{\partial x} = \frac{\partial q}{\partial y} \quad \text{and} \quad \frac{\partial p}{\partial y} = -\frac{\partial q}{\partial x}.$$

Therefore, with $\mathbf{W} = \langle p, q \rangle$ we have:

$$\begin{aligned} \mathcal{F}_{\mathbf{W}, T}(x, y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial p}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial q}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\ &\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial p}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial q}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\ &= \frac{1}{|\nabla T|} \left[-\left(\frac{\partial T}{\partial x} \right)^2 \left(\frac{\partial p}{\partial x} \right) - \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right. \\ &\quad \left. - \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial p}{\partial y} \right) - \left(\frac{\partial T}{\partial y} \right)^2 \left(\frac{\partial q}{\partial y} \right) \right] \\ &= \frac{1}{|\nabla T|} \left[-\left(\frac{\partial T}{\partial x} \right)^2 \left(\frac{\partial p}{\partial x} \right) - \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right. \\ &\quad \left. - \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial p}{\partial y} \right) - \left(\frac{\partial T}{\partial y} \right)^2 \left(\frac{\partial p}{\partial x} \right) \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{|\nabla T|} \left[- \left(\left(\frac{\partial T}{\partial x} \right)^2 + \left(\frac{\partial T}{\partial y} \right)^2 \right) \left(\frac{\partial p}{\partial x} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(-\frac{\partial p}{\partial y} + \frac{\partial p}{\partial y} \right) \right] \\
&= \frac{1}{|\nabla T|} \left[- \left(\left(\frac{\partial T}{\partial x} \right)^2 + \left(\frac{\partial T}{\partial y} \right)^2 \right) \left(\frac{\partial p}{\partial x} \right) \right] \\
&= \frac{1}{|\nabla T|} \left[-|\nabla T|^2 \left(\frac{\partial p}{\partial x} \right) \right] \\
&= |\nabla T| \left(-\frac{\partial p}{\partial x} \right).
\end{aligned}$$

Finally, for an analytic function $f = p + iq$ one has $\frac{df}{dz} = \frac{\partial p}{\partial x} + i\frac{\partial q}{\partial x}$, from which we obtain

$$\mathcal{F}_{\mathbf{W},T}(x,y) = -|\nabla T|(\operatorname{Re} f'(z)).$$

□

Proposition 5.3. *Suppose that $f(z)$ is analytic on D . If $\tilde{\mathbf{W}}$ is the Pólya vector field of f , then*

$$\mathcal{F}_{\tilde{\mathbf{W}},T_R} = \frac{1}{|\nabla T|} \left[\left(\left(\frac{\partial T}{\partial y} \right)^2 - \left(\frac{\partial T}{\partial x} \right)^2 \right) \left(\frac{\partial p}{\partial x} \right) + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right].$$

Proof. Again, we apply the Cauchy-Riemann equations to the analytic function $f = p + iq$:

$$\frac{\partial p}{\partial x} = \frac{\partial q}{\partial y} \quad \text{and} \quad \frac{\partial p}{\partial y} = -\frac{\partial q}{\partial x}.$$

With the Pólya vector $\tilde{\mathbf{W}} = \langle p, -q \rangle$, we put $u = p$ and $v = -q$ into the frontogenesis formula:

$$\begin{aligned}
\mathcal{F}_{\tilde{\mathbf{W}},T}(x,y) &= \frac{1}{|\nabla T|} \left[\left(-\frac{\partial T}{\partial x} \right) \left(\frac{\partial p}{\partial x} \cdot \frac{\partial T}{\partial x} + \frac{\partial(-q)}{\partial x} \cdot \frac{\partial T}{\partial y} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial y} \right) \left(\frac{\partial p}{\partial y} \cdot \frac{\partial T}{\partial x} + \frac{\partial(-q)}{\partial y} \cdot \frac{\partial T}{\partial y} \right) \right] \\
&= \frac{1}{|\nabla T|} \left[- \left(\frac{\partial T}{\partial x} \right)^2 \left(\frac{\partial p}{\partial x} \right) + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right. \\
&\quad \left. - \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial p}{\partial y} \right) + \left(\frac{\partial T}{\partial y} \right)^2 \left(\frac{\partial q}{\partial y} \right) \right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{|\nabla T|} \left[-\left(\frac{\partial T}{\partial x}\right)^2 \left(\frac{\partial p}{\partial x}\right) + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y}\right) \left(\frac{\partial q}{\partial x}\right) \right. \\
&\quad \left. + \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y}\right) \left(\frac{\partial q}{\partial x}\right) + \left(\frac{\partial T}{\partial y}\right)^2 \left(\frac{\partial p}{\partial x}\right) \right] \\
&= \frac{1}{|\nabla T|} \left[\left(-\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 \right) \left(\frac{\partial p}{\partial x}\right) \right. \\
&\quad \left. + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y}\right) \left(\frac{\partial q}{\partial x}\right) \right] \\
&= \frac{1}{|\nabla T|} \left[\left(\left(\frac{\partial T}{\partial y}\right)^2 - \left(\frac{\partial T}{\partial x}\right)^2 \right) \left(\frac{\partial p}{\partial x}\right) \right. \\
&\quad \left. + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y}\right) \left(\frac{\partial q}{\partial x}\right) \right].
\end{aligned}$$

□

Example 5.4. Let us illustrate these concepts with the analytic function $f(z) = z^2$. Since $(x+iy)^2 = (x^2-y^2)+i(2xy)$, the standard vector field of f is $\mathbf{W}(x, y) = \langle x^2 - y^2, 2xy \rangle$ while the Pólya vector field of f is $\tilde{\mathbf{W}} = \langle x^2 - y^2, -2xy \rangle$.

(a) Standard vector field

Let $\mathbf{W}(x, y) = \langle x^2 - y^2, 2xy \rangle$. By Proposition 5.2, the frontogenesis function of $\mathbf{W}$ on any temperature field T is

$$\mathcal{F}_{\mathbf{W}, T}(x, y) = |\nabla T| \left(-\frac{\partial p}{\partial x} \right) = -2x|\nabla T|.$$

Alternatively, since $f'(z) = 2z$, we could obtain the same answer by writing

$$\mathcal{F}_{\mathbf{W}, T}(x, y) = -|\nabla T|(\operatorname{Re}(2z)) = -|\nabla T|(\operatorname{Re}(2x + i2y)) = -2x|\nabla T|.$$

For the special temperature fields T_H and T_R we have $|\nabla T| = 1$ everywhere, and so

$$\mathcal{F}_{\mathbf{W}, T_H} = \mathcal{F}_{\mathbf{W}, T_R} = -2x.$$

(b) Pólya vector field

Now we let $\tilde{\mathbf{W}}(x, y) = \langle x^2 - y^2, -2xy \rangle$. By Proposition 5.3, the frontogenesis function of $\tilde{\mathbf{W}}$ on any temperature field T is

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_R} = \frac{1}{|\nabla T|} \left[\left(\left(\frac{\partial T}{\partial y}\right)^2 - \left(\frac{\partial T}{\partial x}\right)^2 \right) (2x) + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y}\right) (2y) \right].$$

Since for the special temperature fields T_H and T_R we have $|\nabla T| = 1$ everywhere,

$$\begin{aligned}
\mathcal{F}_{\tilde{\mathbf{W}}, T_H} &= (0 - 1)(2x) + 2(0 \cdot 1)(2y) \\
&= -2x \\
\mathcal{F}_{\tilde{\mathbf{W}}, T_R} &= \frac{1}{1} \left[\left(\frac{y^2}{x^2 + y^2} - \frac{x^2}{x^2 + y^2} \right) (2x) + 2 \left(\frac{xy}{x^2 + y^2} \right) (2y) \right] \\
&= \left(\frac{y^2 - x^2}{x^2 + y^2} \right) (2x) + \left(\frac{4xy^2}{x^2 + y^2} \right) \\
&= \frac{2xy^2 - 2x^3 + 4xy^2}{x^2 + y^2} \\
&= \frac{-2x^3 + 6xy^2}{x^2 + y^2}.
\end{aligned}$$

We remark that this is the same answer that we obtained in Example 3.1, before we had developed any of the propositions used in this example.

Example 5.5. The function $f(z) = \frac{1}{z}$ is not entire, but only analytic on the punctured complex plane $\mathbb{C} \setminus \{0\}$. We have $f(z) = \frac{1}{x+iy} = \frac{x-iy}{(x+iy)(x-iy)} = \frac{x}{x^2+y^2} + i \left(\frac{-y}{x^2+y^2} \right)$.

(a) Standard vector field

The standard vector field of $f(z) = \frac{1}{z}$ is $\mathbf{W} = \langle \text{Re}(f), \text{Im}(f) \rangle = \left\langle \frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2} \right\rangle$. Applying Proposition 5.2, the frontogenesis function of $\mathbf{W}$ on any temperature field T is

$$\mathcal{F}_{\mathbf{W}, T}(x, y) = |\nabla T| \left(-\frac{\partial p}{\partial x} \right) = |\nabla T| \left(-\left(\frac{-x^2 + y^2}{(x^2 + y^2)^2} \right) \right) = |\nabla T| \left(\frac{x^2 - y^2}{(x^2 + y^2)^2} \right).$$

In particular, for the special temperature fields T_H and T_R we have $|\nabla T| = 1$ everywhere, and so

$$\mathcal{F}_{\mathbf{W}, T_H} = \mathcal{F}_{\mathbf{W}, T_R} = \frac{x^2 - y^2}{(x^2 + y^2)^2} = \frac{(x - y)(x + y)}{(x^2 + y^2)^2}.$$

(b) Pólya vector field

The Pólya vector field of $f(z) = \frac{1}{z}$ is $\tilde{\mathbf{W}} = \langle \text{Re}(f), -\text{Im}(f) \rangle = \left\langle \frac{x}{x^2+y^2}, \frac{y}{x^2+y^2} \right\rangle$. By Proposition 5.3, the frontogenesis function $\mathcal{F}_{\tilde{\mathbf{W}}, T_R}$ for any temperature field T is

$$\mathcal{F}_{\tilde{\mathbf{W}}, T} = \frac{1}{|\nabla T|} \left[\left(\left(\frac{\partial T}{\partial y} \right)^2 - \left(\frac{\partial T}{\partial x} \right)^2 \right) \left(\frac{\partial p}{\partial x} \right) + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right].$$

Applying this to $T_H = x$ gives

$$\begin{aligned}
\mathcal{F}_{\tilde{\mathbf{W}}, T_H} &= (1) \left[\left((0)^2 - (1)^2 \right) \left(\frac{\partial p}{\partial x} \right) + 2(0) \left(\frac{\partial q}{\partial x} \right) \right] \\
&= -\frac{\partial p}{\partial x} \\
&= -\left(\frac{-x^2 + y^2}{(x^2 + y^2)^2} \right) \\
&= \frac{x^2 - y^2}{(x^2 + y^2)^2}
\end{aligned}$$

and applying it to $T_R = \sqrt{x^2 + y^2}$ gives

$$\begin{aligned}
\mathcal{F}_{\tilde{\mathbf{W}}, T_R} &= \frac{1}{|\nabla T|} \left[\left(\left(\frac{\partial T}{\partial y} \right)^2 - \left(\frac{\partial T}{\partial x} \right)^2 \right) \left(\frac{\partial p}{\partial x} \right) + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right] \\
&= (1) \left[\left(\frac{y^2 - x^2}{x^2 + y^2} \right) \left(\frac{y^2 - x^2}{(x^2 + y^2)^2} \right) + 2 \left(\frac{xy}{x^2 + y^2} \right) \left(\frac{2xy}{(x^2 + y^2)^2} \right) \right] \\
&= \frac{(y^2 - x^2)^2}{(x^2 + y^2)^3} + \frac{4x^2 y^2}{(x^2 + y^2)^3} \\
&= \frac{y^4 - 2x^2 y^2 + x^4 + 4x^2 y^2}{(x^2 + y^2)^3} \\
&= \frac{y^4 + 2x^2 y^2 + x^4}{(x^2 + y^2)^3} \\
&= \frac{(y^2 + x^2)^2}{(x^2 + y^2)^3} \\
&= \frac{1}{x^2 + y^2}.
\end{aligned}$$

Calculations such as Examples 5.4 and 5.5 lead us to two observations. The first observation is one that holds for all analytic functions, and is one that we are able to prove:

Proposition 5.6. *Let $f(z)$ be an analytic function on D , let $\mathbf{W}$ denotes its standard vector field, and $\tilde{\mathbf{W}}$ denote its Pólya vector field. Then,*

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_H} = \mathcal{F}_{\mathbf{W}, T_R} = \mathcal{F}_{\tilde{\mathbf{W}}, T_H} = -\operatorname{Re} f'(z)$$

Proof. For the standard vector field $\mathbf{W}$ of $f(z)$, this follows immediately from Proposition 5.2 and the fact that $|\nabla T_H| = |\nabla T_R| = 1$. Moreover, Proposition 5.3 gives

$$F_{\tilde{\mathbf{W}}, T_H} = \frac{1}{1} \cdot \left((0^2 - 1^2) \left(\frac{\partial p}{\partial x} \right) + 2(1 \cdot 0) \left(\frac{\partial q}{\partial x} \right) \right) = -\frac{\partial p}{\partial x} = -\operatorname{Re} f'(z).$$

□

Our second observation involves the interaction Pólya vector field of f and the radial temperature distribution T_R . While not as simple as the formula in the previous proposition, the frontogenesis function in this case takes on an intriguing form that relates the analytic functions f' and z^2 :

Proposition 5.7. *Let $f(z)$ be an analytic function on D , $\tilde{\mathbf{W}}$ be the Pólya vector field of $f(z)$, $\tilde{\mathbf{V}}$ be the Pólya vector field of $f'(z)$, and $\mathbf{U} = \langle x^2 - y^2, 2xy \rangle$ be the standard vector field of $g(z) = z^2$. Then*

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_R} = -\frac{1}{x^2 + y^2} (\mathbf{U} \cdot \tilde{\mathbf{V}}). \quad (3)$$

Proof. By Proposition 5.3 we have

$$\begin{aligned} \mathcal{F}_{\tilde{\mathbf{W}}_{(z^n)}, T_R}(x, y) &= \frac{1}{|\nabla T|} \left[\left(\left(\frac{\partial T}{\partial y} \right)^2 - \left(\frac{\partial T}{\partial x} \right)^2 \right) \left(\frac{\partial p}{\partial x} \right) + 2 \left(\frac{\partial T}{\partial x} \frac{\partial T}{\partial y} \right) \left(\frac{\partial q}{\partial x} \right) \right] \\ &= \frac{1}{1} \left[\left(\frac{-x^2 + y^2}{x^2 + y^2} \right) \left(\frac{\partial p}{\partial x} \right) + \left(\frac{2xy}{x^2 + y^2} \right) \left(\frac{\partial q}{\partial x} \right) \right] \\ &= -\frac{1}{x^2 + y^2} \left[(x^2 - y^2) \left(\frac{\partial p}{\partial x} \right) + (2xy) \left(-\frac{\partial q}{\partial x} \right) \right] \\ &= -\frac{1}{x^2 + y^2} \left[\langle x^2 - y^2, 2xy \rangle \cdot \left\langle \frac{\partial p}{\partial x}, -\frac{\partial q}{\partial x} \right\rangle \right]. \end{aligned}$$

The first factor in the dot product above is $\mathbf{U}$. Remembering the formula $\frac{df}{dz} = \frac{\partial p}{\partial x} + i \frac{\partial q}{\partial x}$, we see that the second factor is the Pólya vector field $\tilde{\mathbf{V}}$ of $f'(z)$. This establishes (3). $\square$

Remark. With notation as in Proposition 5.7, the Cauchy-Riemann equation applied to f allows gives

$$\tilde{\mathbf{V}} = \left\langle \frac{\partial p}{\partial x}, -\frac{\partial q}{\partial x} \right\rangle = \left\langle \frac{\partial p}{\partial x}, \frac{\partial p}{\partial y} \right\rangle = \nabla p = \nabla(\operatorname{Re} f(z)).$$

So we can alternatively write (3) as

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_R} = -\frac{1}{x^2 + y^2} (\mathbf{U} \cdot \nabla(\operatorname{Re} f(z))).$$

Finally, when we look at the power functions $f(z) = z^n$, for an integer n , we have observed an interesting pattern. Since we have so far not found a proof of this, we state it as:

Conjecture 5.8. *Let $f(z) = z^n$, for $n \in \mathbb{Z}$, and let $\tilde{\mathbf{W}}$ be the Pólya vector field of f . Then*

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_R} = -\frac{n}{x^2 + y^2} \operatorname{Re}(z^{n+1}).$$

We observe this in Example 5.4 when $f(z) = z^2$ and

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_R} = \frac{-2x^3 + 6xy^2}{x^2 + y^2} = -\frac{2}{x^2 + y^2}(x^3 - 3xy^2) = -\frac{2}{x^2 + y^2} \operatorname{Re}(z^3),$$

and in Example 5.5 when $f(z) = z^{-1}$ and

$$\mathcal{F}_{\tilde{\mathbf{W}}, T_R} = \frac{1}{x^2 + y^2} = -\frac{(-1)}{x^2 + y^2} \cdot 1 = -\frac{(-1)}{x^2 + y^2} \operatorname{Re}(z^0).$$

In Figure 4, we illustrate vector fields and resulting frontogenesis functions appearing in this conjecture for $n = -3, -4, \dots, -7$. We note that the contour map of $\mathcal{F}$ in these plots yields the shape of a “flower” in the center, where the computer automatically cuts off its plot of values that increase without bound as one approaches the origin. For $f(z) = z^{-k}$, k a positive integer, Conjecture 5.8 predicts that

$$\begin{aligned} \mathcal{F}_{\tilde{\mathbf{W}}, T_R} &= -\frac{(-k)}{x^2 + y^2} \operatorname{Re}(z^{-k+1}) = \frac{k}{x^2 + y^2} \operatorname{Re}\left(\frac{1}{z^{k-1}}\right) \\ &= \frac{k}{x^2 + y^2} \operatorname{Re}\left(\frac{\bar{z}^{k-1}}{(|z|^2)^{k-1}}\right) = \frac{k}{(x^2 + y^2)^k} \operatorname{Re}(z^{k-1}) \end{aligned}$$

The number of “petals” of the flower (which, for z^{-k} , is $2(k-1)$) related to the regions of alternating sign for the function $\operatorname{Re}(z^{k-1})$ as one circles the origin.

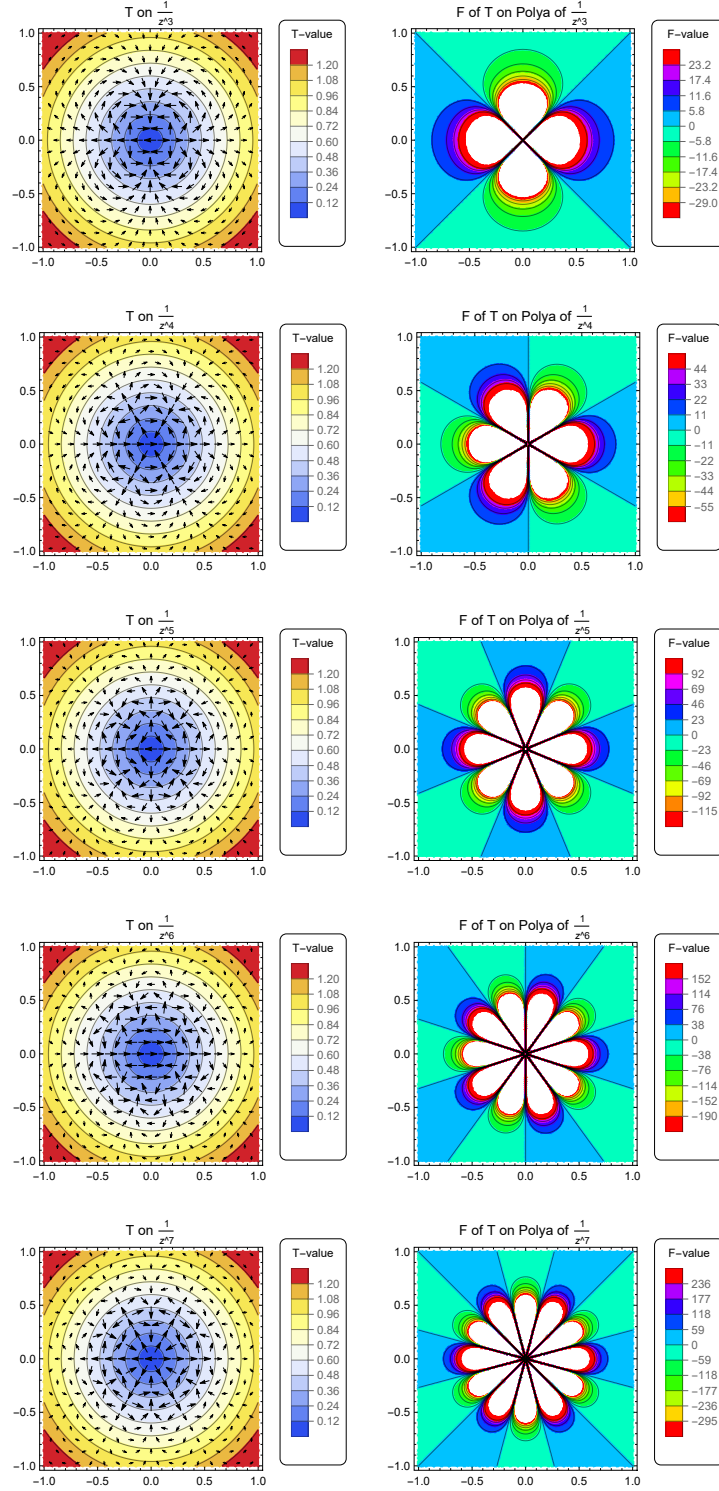


Figure 4: Left images: Plot of the Pólya vector field of $f(z) = z^n$ for $n = -3, -4, -5, -6, -7$ (from top to bottom) overlaid with contour map of T_R . Right images: Contour map of frontogenesis function resulting from setup in image to its left.

6 Conclusion

We have explored mathematical aspects of the frontogenesis function $\mathcal{F}(x, y)$ arising from 2-dimensional meteorological models of the interaction of wind and temperature. From the mathematical standpoint, this function arises when we pair a vector field $\mathbf{W}(x, y)$ with a scalar field $T(x, y)$. By restricting our focus to the simple scalar fields T_H and T_R , our viewpoint has been to see what kinds of frontogenesis functions arise when we pair these simple scalar fields with interesting, more complicated sorts of vector fields. In §4 we gave some results that apply to a wide class of vector fields $\mathbf{W}$, while in §5 we studied the vector fields that come from analytic functions. In the latter section, we found several interesting patterns, were able to verify most of them with proofs, and left one observed pattern as a conjecture.

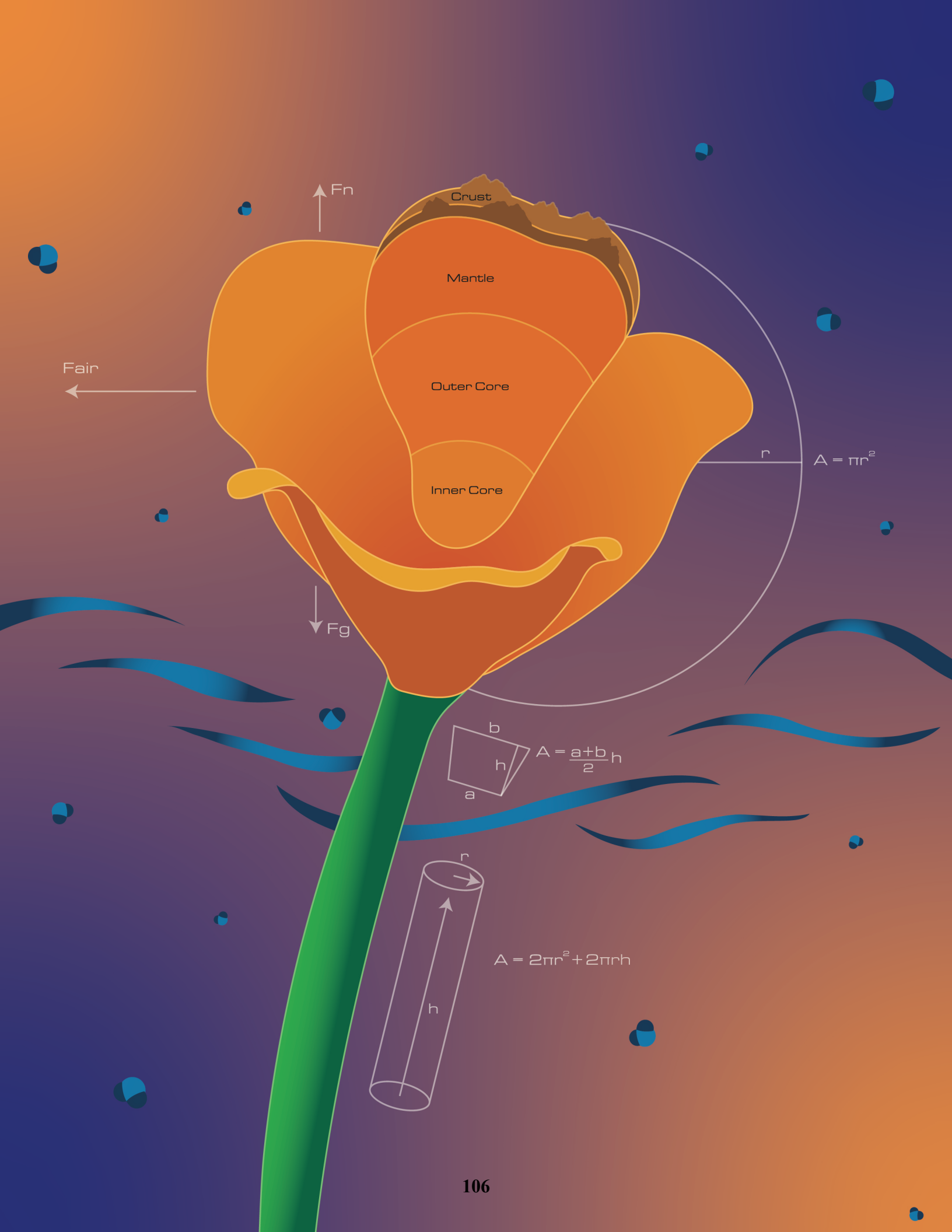
In addition to proving this conjecture and others like it, there are many ways that this work can be extended. As just one example, one could instead start with simple vector fields $\mathbf{W}$ and choose interesting scalar fields (e.g., two-variable harmonic functions, which are the real and imaginary parts of complex analytic functions) and see what shape $\mathcal{F}(x, y)$ takes.

Acknowledgements

The author was generously supported during this research by the LSAMP program as an LSAMP research scholar.

References

- [1] Gamelin, T. *Complex Analysis*. Springer, New York, 2001
- [2] Martin, J. *Mid-Latitude Atmospheric Dynamics: A First Course*. John Wiley & Sons, Hoboken, 2006.
- [3] Needham, T. *Visual Complex Analysis*. Oxford University Press, Oxford, 1997.



F_n

Crust

Mantle

Outer Core

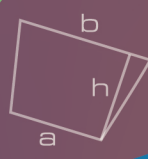
Inner Core

F_{air}

F_g

r

$$A = \pi r^2$$



$$A = \frac{a+b}{2} h$$



$$A = 2\pi r^2 + 2\pi r h$$